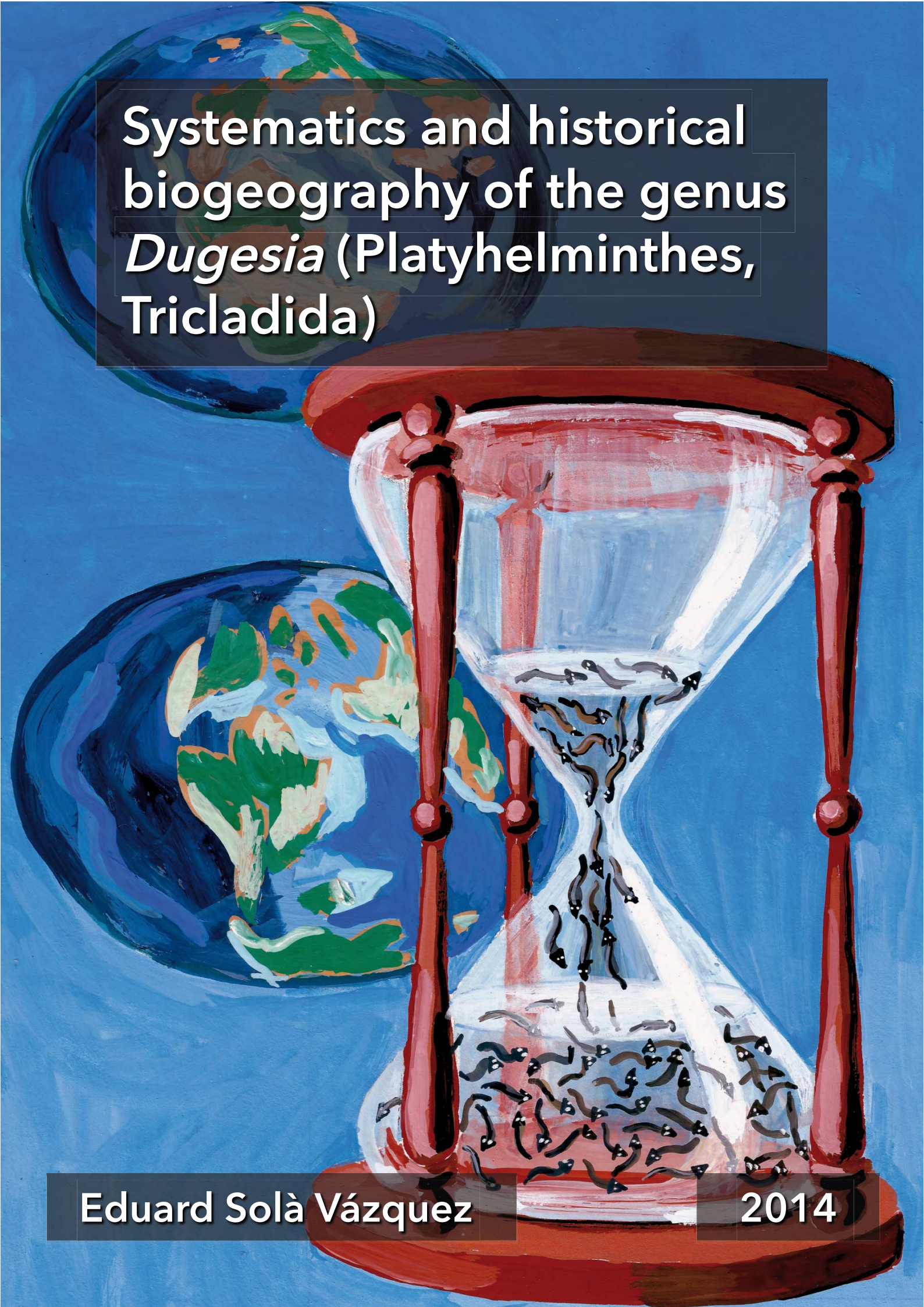
Systematics and historical biogeography of the genus *Dugesia* (Platyhelminthes, Tricladida)

Eduard Solà Vázquez

Systematics and historical biogeography of the genus *Dugesia* (Platyhelminthes, Tricladida)



Eduard Solà Vázquez

2014

Cover illustrations: Joan Solà Lamelas

Universitat de Barcelona

IRBio
Institut de Recerca de la Biodiversitat

Departament de Genètica

Programa de Doctorat de Genètica

Systematics and historical biogeography of the genus Dugesia (Platyhelminthes, Tricladida)

Sistemàtica i biogeografia històrica del gènere Dugesia (Platyhelminthes, Tricladida)

Memòria presentada per Eduard Solà Vázquez per a optar al grau de Doctor per la Universitat de Barcelona

Eduard Solà Vázquez

EDUARD SOLÀ VÁZQUEZ
El doctorand

DR. MARTA RIUTORT LEÓN
La directora

Barcelona, maig de 2014.

Systematics and historical biogeography of the genus *Dugesia* (Platyhelminthes, Tricladida)

Eduard Solà Vázquez

Universitat de Barcelona

IRBio
Institut de Recerca de la Biodiversitat

Acknowledgments

Pràcticament he perdut la capacitat d'exclamar amb alegria i despreocupació "Oh, quin riu!" o "Mireu quin rierol més bonic!" quan en veig un, com faria qualsevol persona amb un cert entusiasme per la natura. Aquestes expressions de joia han sigut substituïdes per la no menys joiosa "Planàries! Aquí segur hi ha planàries!".

Tot aquest deliri u obsessió pels platihelmints de vida lliure (i d'aigua dolça) possiblement no s'hauria donat si en primer lloc en **Julio Rozas** no m'haguès presentat na Marta Riutort. El primer agraïment va per ell.

Després, majestuosament, vull mostrar el meu agraïment a na pròpia **Marta Riutort** que va assumir el risc de reclutar un espècimen com jo. A ella li dec tots aquests anys entre planàries i tot el que això ha comportat, que és moltíssim.

Però qui més m'han sofert tots aquests anys han sigut les companyes i amigues de laboratori **Eva Lázaro**, **Marta Álvarez**, **Laia Leria** i **Àngels Tudó**. És tan complicat agrair tantes coses en tan poc espai! Hauré de resumir donant-vos sencillament les gràcies. Ha sigut fantàstic compartir tot aquest temps amb vosaltres. Em quedo amb una pila de grans records.

Nogensmenys, aquelles persones que han passat pel laboratori en períodes curts han contribuït a fer el desenvolupament de la tesi més amè i interessant. Muito obrigado **Cláudia** i gràcies **Jose**! També a les noves incorporacions que han arribat cap al final de la meva tesi, i amb els que, de ben segur, hauria agraït compartir més temps al laboratori: **Oleguer**, **Paula**, **Raquel** i **Santi**.

Muy agradecido también a **Eduardo Mateos** que tuvo el coraje de llevarme a mi primer muestreo, a Andalucia, y con el que es fácil hacer amigos en los bares del sur de España. El siguiente paso fue la exploración de las islas y las ensaladas griegas, y algun que otro muestreo más.

Una altra gran coneixença ha sigut en **Miquel Vila**, a qui vaig trobar per primera vegada a Amsterdam i amb el que he tingut el privilegi de mostrejar diverses vegades. Mai oblidaré aquelles llegums tan extraordinàries que vam cuinar a Holanda.

Ringrazio a **Giaccinta Stocchino**, una grande colaboratrice e splendida persona. Ho avuto la fortuna di incontrarla più di una volta. Grazie Giacinta!

I no oblidar l'**Enric** i la **Marga**, que en més d'una ocasió m'han portat planàries dels seus viatges a Grècia. Gràcies! També vull fer una menció especial per les companyes oficials d'esmorzar **Gema**, **Eva**, **Eli** i **Paola**. Sou magnífiques!

La vida a la tercera planta del departament de genètica no seria el mateix sense els seus habitants, **Bàrbara**, **Jenny**, **Isaac** i **Marta**. Ah! I els congressos a Madrid tampoc serien el mateix sense **Cristina**, **Pablo**, i **Roser**! Gràcies també a l'**Álex** per fer-me veure la llum en un moment crític d'aquestes últimes setmanes.

I'm specially indebted to **Ronald Sluys** who teached me in the painstaking process of the planarian sectioning and who was a very nice host during my two visits in Amsterdam. He has also been helping me or solving any doubt by mail faster than any other human being. Dank u Ronald!

Southwards, in the beautiful Greek lands I met **Konstaninos Gritzalis,** who was also a fantastic host. We spent many car hours together, sampling planarians all across Greece. I still remember the enthusiasm with which he talked about phylosophy and other interesting social stuff about Greek life. It was one of my greater experiences. Σας ευχαριστούμε! Kostas!

The stay in Michigan was one of the other fantastic experiences during the thesis development in which I was lucky to learn a lot. I'm very grateful to **Lacey L. Knowles** who had no problem in accepting me in her lab. I really enjoyed the events at her home such as pumpkin carving, thanksgiving day, cookies workshop, and random barbacues. I'm also extremely happy to have had the opportunity to meet fascinating and cheerful people in the Ecology and Evolutionary Department of the University of Michigan: **Melisa**, **Pavel**, **Qixin**, **Lucy**, **Juan Pablo**, **Déa**, **Pamela**, **Diego**, **Raquel**, **Carlos**, **JP**, **Hayley**, **Mauricio**, **Tristan**, and many others.

Fora del món acadèmic nombroses persones m'han donat el seu suport i preguntat i fet broma al voltant de les planàries. El meu agraïment més gran per a la meva família; la meva mare **Dora**, el meu pare **Joan** i la meva germana **Blanca**, sense els quals aquesta tesi no haguès sigut possible, literalment. Han sigut col·laboradors entusiastes responsables de part del mostreig a Grècia. D'altra banda, mon pare ha ideat i executat diverses il·lustracions de planàries per a presentacions i per a aquesta mateixa tesi. I sense sortir de l'àmbit personal, hi ha algú que s'ha carregat de paciència i afecte durant gran part de la meva tesi i a qui també li dec moltíssim, mil gràcies **Noelia**. Moltíssimes gràcies per tot a tots.

El meu últim i més que sincer agraïment va per cadascuna de les ***Dugesia***, ***Recurva*** i altres planàries (***Polycelis***, ***Crenobia***, ***Dendrocoelum***, ***Phagocata***, ***Microplana***) que s'han creuat pel camí de la meva tesi. A elles els dec les meves primeres passes pel món científic.

Thesis summary in Catalan language

Dugesia és un gènere de platihelmints triclàdides de vida lliure que habita a l'aigua dolça, trobant-se a fonts, rierols, rius i llacs entre d'altres. Les espècies d'aquest gènere es caracteritzen per presentar un cap triangular amb dos ulls i un cos allargat i aplanat dorsoventralment. Aquesta forma tan característica les fa reconeixibles per persones no expertes que també les solen identificar per les seves capacitats de regeneració. Quan les planàries d'aigua dolça són ferides o bé seccionades, aquestes tenen la capacitat de regenerar el tros que els hi manca gràcies a l'activitat dels neoblasts, que actuen com a cèl·lules mare. Aquesta capacitat sembla ser més accentuada en la família dels dugèsids i especialment en aquells individus que es reprodueixen asexualment per fissiparitat. A part de la reproducció asexual per fissió, les *Dugesia* també poden reproduir-se per partenogènesi (ponen ous que no han sigut fecundats) o bé sexualment per fertilització creuada. En estat salvatge es poden trobar individus reproduint-se d'una de les tres maneres, però es desconeix si poden canviar el mètode de reproducció a la natura. Tanmateix, en condicions de laboratori s'ha observat recentment com progenitors asexuals triploides engendraven descendència sexual diploide, suggerint que aquest canvi també es pot donar a la natura. L'asexualitat s'acostuma a relacionar en *Dugesia* amb individus amb cariotips triploides mentre que la reproducció sexual és típica d'animals diploides.

Els aspectes biogeogràfics de les *Dugesia* són els que han captat majoritàriament el nostre interès pel desenvolupament d'aquesta tesi. Les espècies d'aquest gènere es troben distribuïdes a Àfrica, Europa, Orient Mitjà, Àsia Meridional, Extrem Orient i Australàsia. En contrast amb l'àmplia distribució de les *Dugesia*, les planàries d'aigua dolça es caracteritzen per tenir una capacitat de dispersió reduïda, limitada a la continguïtat dels rierols, rius i llacs que habiten. Es tracta d'organismes fràgils de desenvolupament directe que no poden sobreviure en aigua salada i, per tant, es considera que no poden dispersar a través de mar i oceans. Tampoc es considera probable la dispersió aèria per ocells o per sobre terra. És per això que alguns planariòlegs han considerat que els patrons filogenètics de les espècies de planària haurien de reflexar els esdeveniments d'aïllament i contacte de les conques fluvials o cossos d'aigua dolça i de les masses de terra que les contenen. Així doncs, sembla ser

que les planàries serien uns organismes adequats per a realitzar estudis de biogeografia històrica. Aquest és l'enfocament que hem donat a dos dels quatre articles principals presentats en aquesta tesi, combinant filogènies moleculars i dades paleogeogràfiques per tal d'esbrinar de quina manera els processos històrics han afectat a la diversificació i distribució del gènere *Dugesia*. Els estudis biogeogràfics duts a terme s'han centrat en primer lloc a la zona de l'Egeu, part de l'àrea que avui en dia ocupa Grècia i part de la regió més occidental de Turquia (Capítol 1), i posteriorment sobre tota la distribució del gènere, incloent espècimens distribuïts des de Sud Àfrica i Madagascar fins a Austràlia, passant per Europa, Orient Mitjà i l'Extrem Orient (Capítol 2).

La motivació de l'estudi de biogeografia històrica centrat a Grècia es basava en la gran diversitat d'espècies de *Dugesia* ja descrites a la zona, així com en la complexa i força ben coneguda història geològica de la regió. Aquests dos factors convertien Grècia en un bon model per a testar hipòtesis de biogeografia històrica en *Dugesia*. Vam testar aquestes hipòtesis fent servir una aproximació molecular multilocus, utilitzant mètodes bayesians i de màxima versemblança en la reconstrucció d'àrbres filogenètics. D'altra banda, vam realitzar estimacions dels temps de divergència dels diferents llinatges analitzats emprant un rellotge molecular relaxat i vam inferir les possibles àrees geogràfiques que ocupaven els ancestres utilitzant un mètode bayesià. La topologia dels arbres filogenètics d'aquest treball presentaven una estructura que podria indicar una certa correlació amb la història geològica de l'Egeu. Així, les espècies de Creta resultaren ser el grup germà de la resta de *Dugesia* gregues, un fet que seria coherent amb l'esdeveniment de separació d'aquesta illa en primer lloc de l'antiga massa de terra unificada anomenada Ägäis entre fa uns 11 i fa uns 9 milions d'anys (Ma). Una altra estructura topològica interessant en l'arbre de *Dugesia* era la separació d'espècies pròpies de l'oest i de l'est de l'Egeu en dos grups, amb una o dues excepcions. Aquesta divergència va ser probablement resultant d'un esdeveniment geològic concret, l'obertura de la "trinxera central de l'Egeu" (en anglès *Mid-Aegean trench*), de la qual se'n coneix l'impacte en la diversificació d'altres grups de fauna que es trobaven a la mateixa zona. Segons l'anàlisi de datació realitzat en aquest estudi, les espècies del centre de l'Egeu (*Dugesia ariadnae* i *D. improvisa*) probablement haurien creuat aquesta trinxera durant l'anomenada crisi salina del Messinià, que fa uns 5 Ma va assecar del tot o parcialment el Mediterrani, permetent que conques fluvials aïllades confluïssin. Altres esdeveniments històrics interpretats a partir dels resultats obtinguts contemplen la possibilitat d'una extinció de *Dugesia* a la Grècia occidental seguida per

una recolonització des del nord i l'expansió geogràfica d'una o unes poques poblacions que havien perdurat a la península del Peloponès. D'altra banda, hem trobat evidències de dispersió per humans d'animals d'aquesta mateixa península cap a les illes de Creta i Cefalònia. La taxa de substitució obtinguda a partir d'aquest estudi va resultar ser comparable a la d'animals d'altres grups (p.e. artròpodes) en ser de 0.0173 per posició per milió d'anys.

L'objectiu del segon treball de biogeografia històrica de *Dugesia* era el d'incloure el màxim possible de representats de *Dugesia* al llarg de la seva distribució coneguda per tal de descobrir patrons biogeogràfics que ajudéssin a explicar l'àmplia distribució del gènere i si aquesta està relacionada amb la seva antiguitat. Per a aquesta recerca també vam emprar mètodes de reconstrucció filogenètica bayesians i de màxima versemblança. Tanmateix, vam realitzar una datació emprant un rellotge molecular relaxat i vam dur a terme la reconstrucció de les àrees ancestrals dels diferents llinatges amb una metodologia basada també en la màxima versemblança. Fins a l'inici d'aquest treball la proposta principal sobre l'origen i dispersió de les *Dugesia* suggeria Gondwana com a bressol d'aquest gènere. Gondwana era un superterreny que incloïa tots els continents i subcontinents de l'actual hemisferi sud i que va formar part del supercontinent Pangea fins fa aproximadament uns 185 Ma. Al final del període Triàsic (fa uns 210 Ma), Europa va quedar coberta per l'extensió cap a l'oest de l'oceà de Tetis. D'aquesta manera, Àsia va quedar aïllada de la resta de Gondwana. Més tard, fa uns 160 Ma Gondwana va començar a fracturar-se, iniciant el trencament pel que avui en dia és la costa de Somàlia. Aquesta fractura va dur en primer lloc a la separació de Madagascar, l'Índia, Austràlia i Antàrtida respecte d'Àfrica i Sud-Amèrica. Fa uns 88 Ma l'Índia va separar-se de Madagascar i va migrar cap el nord fins a impactar amb Àsia fa uns 35−20 Ma. Aquesta és una de les vies proposades per la dispersió a Euràsia d'ancestres procedents de Gondwana. L'altra via proposada contempla la dispersió a través de l'impacte de la Península Aràbiga amb l'Orient Mitjà fa uns 20 Ma. D'acord amb la topologia dels nostres àrbres filogenètics, l'hipòtesi més plausible és en realitat un origen més antic, sobre Pangea, fa almenys uns 220 Ma, probablement vora els 240 Ma. Després del seu trencament, dos llinatges de *Dugesia* que han perdurat fins el present haurien diversificat un al nord i un altre al sud. Les espècies que es troben avui en dia a Madagascar haurien sobreviscut i continuat sobre l'illa des del moment del seu trencament de la present costa Somalí fa entre 160 i 130 Ma. En aquest treball també hem trobat evidències de dispersió, així per exemple des de Madagascar a Oman i des

del sud-est asiàtic a Nova Guinea i Austràlia. Una altra conclusió interessant d'aquest treball suggereix que si realment es tracta d'un gènere tan antic com 240 Ma, ens trobem davant un cas d'estasi morfològica molt antiga. Les espècies de *Dugesia* s'assemblen totes externament i la morfologia interna es manté molt homogènia, fins el punt que es coneixen força casos de convergència o paral·lelisme evolutiu dels estats dels escassos caràcters morfològics de l'òrgan copulador.

Un altre aspecte diferent a la biogeografia històrica però relacionat amb la diversificació i estudi evolutiu de *Dugesia* és la delimitació i descripció d'espècies d'aquest gènere. En el context de la crisi de biodiversitat actual és important catalogar el màxim d'espècies abans no desapareguin. Tot i això, la descripció formal d'espècies d'aquest gènere és quelcom problemàtic per diversos motius. El primer és el procés laboriós i no sempre exitós que implica la preparació de les seccions histològiques necessàries per a una correcta anàlisi i diagnosi dels espècimens. El segon és l'existència d'espècies amb morfologies extremadament similars, fet que dificulta la seva distinció. Per últim, l'extensa presència de poblacions fissípares, que no presenten òrgan copulador. Aquesta part de les *Dugesia* és la única que ofereix caràcters morfològics rellevants per a distingir espècies. Tot i que és possible induir la formació de l'òrgan copulador en condicions de laboratori, no sempre és fàcil o infal·lible i sovint requereix una cura i atenció prolongada en el temps dels animals en captivitat. Donades les diverses dificultats ennumerades, l'ús de dades moleculars en la delimitació d'espècies emprant diversos mètodes de delimitació d'espècies és prometedora per tal de solventar-les o fer-les menys feixugues. En el treball de delimitació i descripció d'espècies inclós en aquesta tesi (Capítol 3) s'ha aplicat un mètode basat en dades moleculars conegut com a *General Mixed Yule-Coalescent* (GMYC) sobre poblacions de *Dugesia* de Grècia. Aquesta informació es va combinar amb anàlisis morfològiques quan disposavem d'elles, resultant en la descripció de quatre espècies noves d'aquest gènere. També es van proposar nombroses espècies noves candidates en base a resultats únicament moleculars i/o sobre resultats morfològics incomplets (preparacions morfològiques danyades). D'altra banda es va descriure un nou gènere exclusivament en base a la morfologia (*Recurva*), però que vam situar en un arbre filogenètic molecular, resultant ser el gènere germà de *Schmidtea*. En l'aproximació adoptada de delimitació d'espècies pel mètode GMYC vam emprar un únic gen mitocondrial, el Cox1. Tanmateix, l'ús de més marcadors moleculars i diversos mètodes de delimitació d'espècies basats en molècules probablement permetrà resoldre i descriure amb més

precisió els casos problemàtics en la descripció i identificació d'espècies de *Dugesia*. Aquest sistema serà especialment útil quan es tracti amb espècies críptiques o poblacions asexuals, prioritzant la resexualització d'aquelles bèsties que resultin menys similars molecularment (o considerades com a llinatges moleculars independents) respecte a les ja descrites.

Finalment, pel desenvolupament de la present tesi també es va treballar en la seqüenciació i annotació de genomes mitocondrials complets de diverses espècies de tricl àdides amb l'objectiu primer d'obtenir nous marcadors mitocondrials (Capítol 4). Tot i no tenir èxit en l'obtenció del genoma mitocondrial d'una espècie d'interès de *Dugesia*, es van aconseguir els genomes mitocondrials complets de dues espècies pertanyents a dues famílies diferents de planàries tricl àdides (*C. alpina*, Planariidae; *Obama* sp., Geoplanidae). D'aquesta manera s'ha enxamplat l'escassa disponibilitat de mitogenomes de platihelmints de vida lliure de tres (dos publicats i un tercer accessible a GenBank) a cinc. De la major disponibilitat de genomes mitocondrials en vam treure profit duent a terme unes anàlisis de tipus evolutiu, tot comparant els mitogenomes dels tricl àdides amb aquells de platihelmints paràsits. L'objectiu era el de trobar evidències de possibles diferències selectives entre aquests dos grups degut als diferents cicles vitals dels tricl àdides (vida lliure) i dels neodermats (paràsits). Esperavem trobar una pressió selectiva més relaxada en el segon grup, ja que aquest tipus de vida implica mides poblacionals efectives petites. Sorprenentment, els resultats mostren que els tricl àdides (concretament els geoplanoïdeus) presenten una major relaxació en la selecció dels nucleòtids del mitogenoma en comparació amb els paràsits.

Una mirada general sobre els resultats d'aquesta tesi indiquen la utilitat i la conveniència d'emprar dades moleculars en estudis de tipus biogeogràfic i sistemàtic en planàries d'aigua dolça, millorant i resolent antigues hipòtesis o incerteses. De fet, l'ús d'aquest tipus de dades hauria de ser inseparable de qualsevol estudi evolutiu o de diversitat de tricl àdides. D'altra banda, l'ús de més marcadors moleculars com genomes mitocondrial sencers o seqüències nuclears obtingudes per tecnologies de seqüenciació de *next-generation* és prometedor i necessari per tal d'aconseguir més informació i en conseqüència una major potència per aplicar tests estadístics o obtenir una major resolució en la recerca evolutiva de les planàries d'aigua dolça.

Contents

Section I

Introduction

1

The genus *Dugesia* (Girard, 1850)

Dugesia (Girard, 1850) is a widespread genus of free-living Platyhelminthes which representatives dwell in freshwater habitats of Africa, Eurasia and Oceania. It is one of the most popularly known planarian due to its usual depiction in textbooks, that explain their regenerative capabilities after being chopped off or wonded. *Dugesia* external appearance makes it easily recognizable for non-specialists, who are familiar with their head of triangular shape with two eyes and its flattened and elongated body (Fig. 1.1).

Dugesia is the most specious genus among the dugesiids, at the beginning of this work it included 73 species (Annex 1 − Table 1). Due to their external similarities, every species is described on the basis of its inner morphology, particularly on features of the copulatory apparatus. The combination of different diagnostic characters allows the erection of new species or the assignment of individuals to those species already described (Sluys *et al*., 1998).

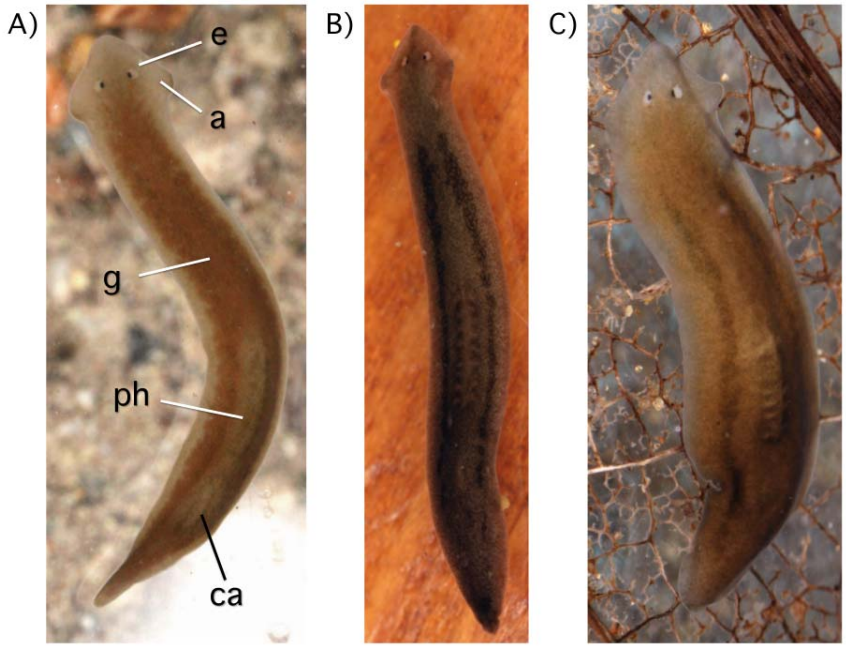


Fig. 1.1 External appearance of living specimens of *Dugesia*. A) *D. elegans* from Rhodes; B) *D. cretica* from Crete; C) *D. ariadnae* from Naxos. Abbreviations: a, auricle; ca, copulatory apparatus; e, eye; g, gut; ph, pharynx. Photographies: Eduardo Mateos.

1.1 Phylogeny and evolution

1.1.1 The phylogenetic position of the genus *Dugesia*

Up to the start of this thesis, *Dugesia* was one out of the eleven formally described genera within the Dugesiidae family (Sluys *et al*., 2009). According to molecular phylogenetics, its sister taxon is the genus *Schmidtea* (Álvarez-Presas *et al*., 2008).

Molecular data have shown the dugesiids to be the sister group of the land planarians (Geoplanidae), together constituting the superfamily Geoplanoidea. The monophyly of this group relies solely on the support of the molecular data after no morphological synapomorphies have been successfully found (Carranza *et al*., 1998; Sluys, 2001). On the other hand, the rest of freshwater families, Planariidae, Dendrocoelidae and Kenkiidae, are encompassed in the superfamily Planarioidea, the sister clade of the Geoplanoidea. These two superfamilies are included in the suborder Continenticola within the order Tricladida. The other triclad suborders are the predominantly saltwater inhabitants Maricola and the cave dwellers Cavernicola (Fig. 1.2).

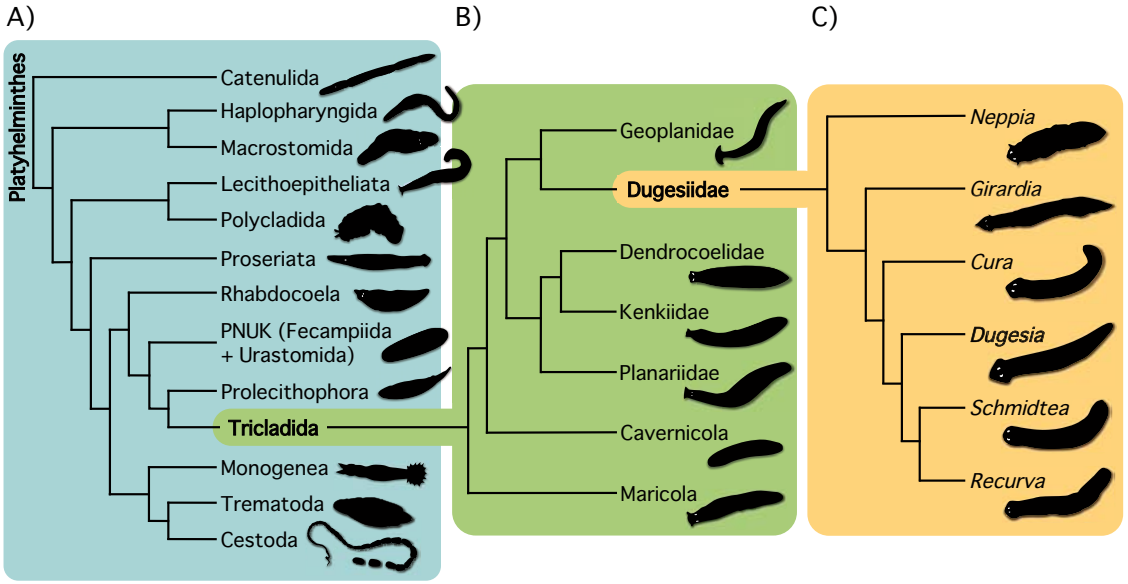


Fig. 1.2 Phylogeny of *Dugesia* at different levels: A) Platyhelminthes phylum phylogeny (based on Riutort *et al*., 2012); B) Tricladida order phylogeny (based on Sluys *et al*., 2009); C) Dugesiidae family phylogeny (based on Álvarez−Presas *et al*., 2008 and Sluys *et al*., 2013).

The relationships of the Tricladida with the other Platyhelminthes are still quite uncertain, as long as most of the phylum phylogeny is not yet fully resolved (reviewed by Riutort *et al*., 2012). Nonetheless, molecular studies suggest that most likely triclads are the sister group of either Fecampiida or Prolecithophora. In the tree of life, Platyhelminthes are clustered with little doubt within the Lophotrochozoa. However, its position within this group is still unclear (e.g. Giribet, 2008).

1.1.2 The Platyhelminthes through geological time

It is not known with certainty when Platyhelminthes split from its sister group. One of the few indirect evidences of the phylum antiquity may be found looking at its relationship with other major phylums for which there exist a richer fossil record. As most phyla, it is generally thought that platyhelminths were already present during the Cambrian period, about 541−485.4 million years ago (Mya). Supporting this hypothesis, molecular dating analyses have pointed an early Cambrian, Ediacaran or even a Cryogenian origin (Peterson *et al*., 2004; Peterson *et al*., 2008; Blair, 2009; Edgebombe *et al*., 2011).

The fossil record of the platyhelminths is extremely sparse because its representatives have no hard body parts prone to fossilize, excepting the hooks and eggs of parasitic flatworms, that can be preserved in certain conditions (Dentzien-Dias *et al*., 2013 and references therein). The oldest putative platyhelminth fossil dates back to the late Devonian (382−373 Mya). It consists of hooks belonging to parasitic flatworms preserved in acanthodian fishes remains (Upeniece, 2001). The only fully preserved fossil of a free-living platyhelminth is about 40 My old, from the Eocene. It is a specimen of a Typhlopanoida rhabdocoel species called *Micropalaeosoma balticus* Poinar, 2003 preserved in baltic amber (Poinar, 2003; 2004). Among the sparse fossil record of the Platyhelminthes, the only fossils putatively attributed to the Tricladida order are few rare pieces from the Miocene preserved in calcareous nodules, including six silicifed specimens and numerous cocoons or egg capsules (Pierce, 1960). These fossil individuals were attributed to undetermined species of Rhabdocoela, Planariidae and Rhynchodemidae, while the cocoons were identified as belonging to the 'Turbellaria' group. Later, the assignment of the specimen described as a rhynchodemid was challenged by Ogren and collaborators (1993).

1.2 Characteristics

1.2.1 General features

As all triclad representatives, *Dugesia* species are metazoans of bilateral symmetry, with a dorso-ventral flattened body, and an anterior-posterior polarity. They are triblastics (three tissue layers), acelomates and unsegmented.

Externally, *Dugesia* is characterized by a head of triangular form with two eyes in the middle and a flattened and elongated body. They are generally inconspicuously coloured animals. The dorsal surface colouration of *Dugesia* ranges from some shade of grey, brown or black to the creamy white of the unique cave-dweller *D. batuensis* BALL, 1970. The dorsal surface is usually plain but some species present pigments mottles (e.g. dark yellow, dark reddish brown, to brown in *D. siamana* KAWAKATSU, 1980 or brown-black in *D. capensis* SLUYS, 2007). Additionally, some species also present stripes, such as *Dugesia neumanni* (NEPPI, 1904). The ventral surface of the *Dugesia* species is always paler than the dorsal and in some species it can also show indistinct pigment spots (e.g. *D. siamana*) conferring a granular or mottled appearance (De Vries, 1988a).

The Tricladida order name comes from the Ancient Greek (*tri/τρι-*, 'three'; and *klados/κλάδος*, 'branch') and describes the main inner characteristic of the group, its digestive system comprised by three main intestinal trunks (Fig. 1.3). From these trunks, many diverticula are projected. The intestine starts at the end of the pharynx and from this point, two branches go backwards along each side of the body. The third branch goes forward along the middle line of the body till just behind, or leveled with, the eyes or the brain. Because triclads have a blind gut (i.e. they lack an anus), the indigestible remains are flushed out through the pharynx. This is a retractile tubular structure located approximately at the middle of the body, housed in a cavity when retracted. It can be protuded from the ventral mouth. In *Dugesia* the pharynx inner structure is constituted by two main musculature layers which are made up of sublayers. The internal sublayer is consisting of two distinct layers, one thick circular adjacent to the epithelium of the pharynx lumen and a thinner one of longitudinal fibres. The external sublayer may consits of three layers. The inner and outer wall of the pharynx are covered by a predominantly glandular epithelium. The tip of this structure holds digestive glands that help in the swallowing of the meal (Ball and Reynoldson, 1981).

In the dugesiids, excepting many *Girardia* species and *Bopsula evelinae* MARCUS, 1946, the pharynx is unpigmented (Sluys, 2001).

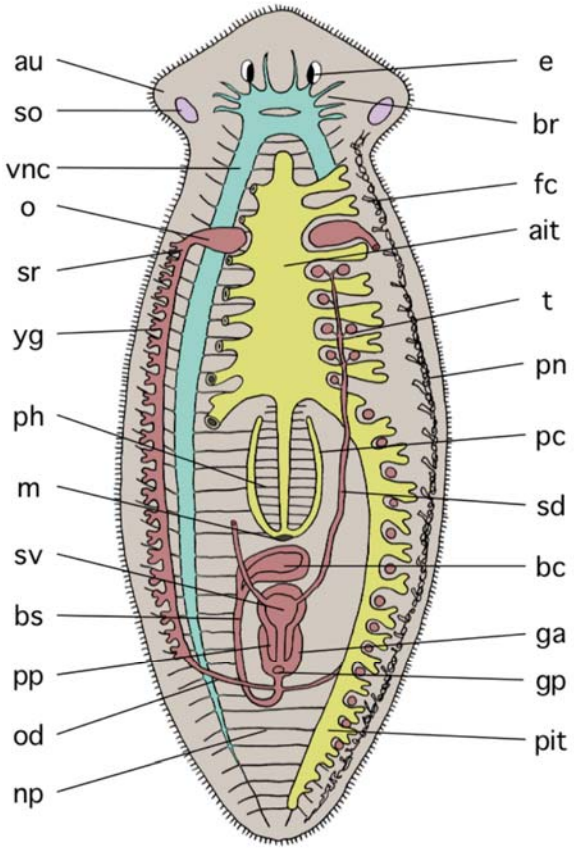


Fig. 1.3 Schematic figure showing different internal systems of a *Dugesia*. In blue: nervous system; In yellow: digestive system; In red: reproductive system. Abbreviations: ait, anterior intestinal trunk; au, auricle; bc, bursa copulatrix; br, brain; bs, bursal stalk or bursal canal; e, eye; fc, flame cell; ga, genital atrium; gp, genital pore; m, mouth; np, nervous plex; o, ovary; od, ovovitelline duct; pc, pharyngeal chamber; ph, pharynx; pit, posterior intestinal trunk; pn, protonephridium; pp, penis papilla; sd, sperm duct or vas deferens; so, sensory organ (auricular grooves); sr, seminal receptacle; sv, seminal vesicle or bulbar cavity; t, testis; vnc, ventral nerve cord; yg, yolk gland. Based on Kawakatsu and Mitchell, 2004. Illustration: Joan Solà.

As the rest of triclads, *Dugesia* species present an excretory system involved in the elimination of waste products (Ishii 1980a; 1980b). It is consisting in a network of flame cells connected by protonephridial ciliated ducts beneath the epidermis on each side of the body. The nephridiopores open from the dorsal and ventral surface of the animal. This system is also involved in the organism osmoregulation (Hyman, 1951) (Fig. 1.3).

The nervous system is relatively rudimentary. It basically includes a central nervous system consisting of a bilobed 'brain' or cerebral ganglions and two main nerve

cords starting from that 'brain' and running along the ventral side of the body (Agata, 1998). A nervous plexus that connects with the main nerve cords runs beneath the body-wall musculature (Fig. 1.3).

Dugesia species do not have either a circulatory or a respiratory system. Therefore, they depend on diffusion through the monolayered epidermis to obtain oxygen. Their skin is covered with cilia restricted to the ventral surface, the auricles, and the surfaces surrounding them (Skaer, 1961; MacRae, 1967; Best *et al*., 1968). It also bears rhabdites, small rod-like structures which can be extruded. These enigmatic bodies are thought to have a protective function, being defensive structures against attacks or generating a protective envelope against adverse physical or chemical conditions. Additionally, in the epidermal and subepidermal layers there are many glands of different kinds some of them involved in mucus production (Török and Röhlich, 1959; Klima, 1961; Skaer, 1961; Spiegelman and Dudley, 1973). The epithelium is penetrated by ducts of many of these glands.

In *Dugesia* the subepidermal musculature underneath the skin is constituted by four layers. It is thicker on the ventral surface and it is involved in functions such as locomotion and waste excretion. There is a thin layer of transverse fibres between the ventral nerve cord and the guts throughout the body. However, its development is not constant between the *Dugesia* species (De Vries, 1988a). The body is filled with the parenchyma or mesenchyme, a diffuse connective tissue (Ball and Reynoldson, 1981).

Dugesia species never swim, they move gliding on firm substrats by the activity of the ventral longitudinal muscles plus cilia of the epidermis. Muscle contractions give more power than cilia in the animal motion. Their locomotion is facilitated by a mucus 'carpet' secreted by themselves on which they glide. Interestingly, the secretion of this mucus implies a major expediture of energy in the flatworm economy (Calow and Woolhead, 1977).

Dugesia species can react to external stimulus through sensory organs located at the anterior end of the body. They have both chemoreceptors (MacRae, 1967) and photoreceptors (Carpenter *et al.*, 1974), that are connected through nervous projections with the cephalic ganglia that process the external stimuli. The photoreceptors consist in two eyes situated on the dorsal side of the head in conspicuous pigment-free patches. These eyes are consituted by a multicellular pigment cup containing many retinal cells

(Hesse, 1897; Ball, 1974b; Sluys, 2001 and references therein). The presence of supernumerary eyes may occur, and in other cases the eyes are reduced (e.g. *Dugesia myopa* DE VRIES, 1988A).

Dugesia present two types of chemoreceptive sense organs, the sensory organs of the auricles or auricular grooves and sensory organs marginally placed (sensory fossae) (De Vries and Sluys, 1991). The sensory organs of the auricles or auricular grooves are the principal organs of chemoreception. They are on each side of the head, one on either side at the dorsal surface of the body, marginally placed at the eyes level or in a position slightly posterior (Fig. 1.3). These unpigmented organs are constituted by a strip of modified sensory epithelium that is richly supplied with nerve endings, covered with long cilia, and rhabdites-free. They are widely present in the Tricladida order and they have been considered a putative plesiomorphy of the dugesiids (Wilhelmi, 1908).

The sensory fossae are located at the anterior margin of the body. They are small patches of modified sensory epithelium, shallow and inconspicuous small invaginations in the body wall (De Vries and Sluys, 1991). *Dugesia* species have between 5 and 10 pairs. Their number can even depend on the individual (cf. De Vries, 1988a).

1.2.2 Reproductive system characteristics

As it has been already mentioned at the beginning of the present introduction, *Dugesia* reproductive system is of capital importance in the species erection and recognition. The copulatory apparatus description is included in all formal proposals of new species.

The reproductive system of *Dugesia* includes two paired ovaries located in the anterior part of the body, on the ventral side, and close to the cephalic ganglia (Fig. 1.3). From the ovaries the oviducts run ventrally to the level of the copulatory apparatus, then they turn dorsal and open to the vaginal area of the bursal canal, above the openings of the shell glands (Fig. 1.4). The testes are dorsal, follicular and numerous, distributed in rows throughout the two sides of the body from the ovaries to the posterior end of the animal. The sperm is released into the single seminal vesicle through the vasa deferentia or sperm ducts. They usually enlarge before entering separately in the seminal vesicle located in the penis bulb and surrounded by bulbar muscles, forming a sort of pseudo-

seminal vesicle. The copulatory apparatus is located at the posterior half of the body. The penis is constituted by a penis bulb located in the atrial wall and of variable muscularity, and by the free intromittent penis papilla that projects into the atrium (De Vries, 1988a). The atrium is connected with the bursa copulatrix through a bursal canal which runs to the left of the copulatory apparatus. In the bursa copulatrix the excedent of the sperm is stored and digested. The bursal canal is surrounded by a thin subepithelial inner layer of longitudinal muscles overlain by circular fibres, and an ectal reinforcement is present in the vaginal area (i.e. distal section of the bursal canal). In many cases, the ectal reinforcement extends further anterior to the bursa copulatrix. The intrabulbar seminal vesicle is separated from the ejaculatory duct by a diaphragm that can vary in shape, size and position. There is a penial glandular region separating the seminal vesicle from the ejaculatory duct concentrated in that diaphragm (De Vries and Sluys, 1991). The openings of the penial glands are concentrated in the diaphragm. The 'adenodactyls' are additional structures of various types often present, but they have no known function. The copulatory apparatus opens to the exterior through the gonopore or genital pore, situated ventrally on the midline of the body. There use to be additional cement glands discharging around such gonopore (Fig. 1.4).

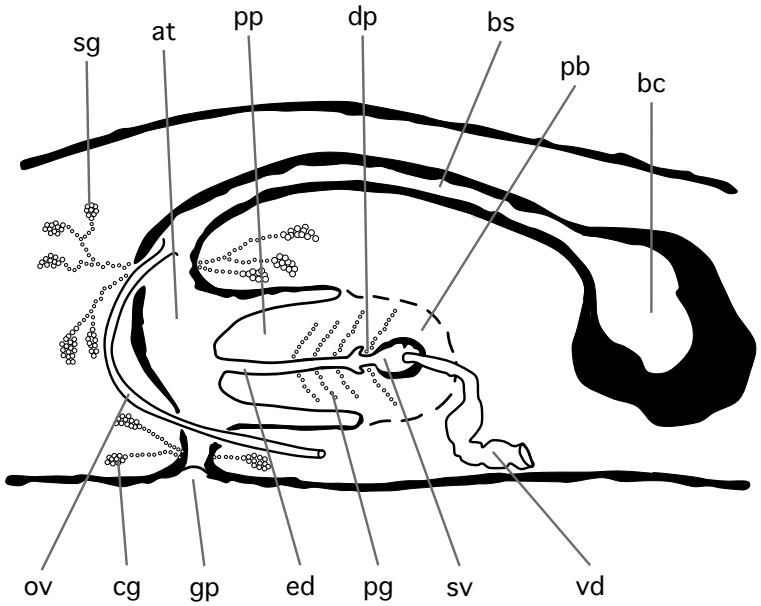


Fig. 1.4 Schematic drawing of a generalistic *Dugesia* copulatory apparatus. Abbreviations: at, atrium; bc, bursa copulatrix; bs, bursal canal; cg, cement glands; dp, diaphragm; ed, ejaculatory duct; gp, gonopore; ov, oviduct; pb, penis bulb; pg, penial glands; pp, penis papilla; sg, shell glands; sv, seminal vesicle; vd, vas deferens.

1.2.3 Differential inner features of the genus *Dugesia*

The genus *Dugesia*, contrarily to other dugesiid genera, is a well-defined and homogeneous group. All its species are characterized by the presence of a diaphragm between the seminal vesicle and the ejaculatory duct. A second synapomorphy suggested for the whole genus *Dugesia* is the emergence of the oviducts from the dorsal surface of the ovaries (Sluys, 2001). A third proposed but doubtful synapomorphy was the extension of the ectal reinforcement (i.e. third layer of longitudinal muscles) along the bursal canal, not confined to the region where the oviducts open into the bursal canal but extending further anteriorly, often reaching as far as the bursa copulatrix (De Vries, 1988a; De Vries and Sluys, 1991). Yet, this may not be among the strongest apomorphies for the genus because some species lack of it (e.g. *Dugesia afromontana* STOCCHINO & SLUYS, 2012 or *D. aethiopica* STOCCHINO, CORSO, MANCONI & PALA, 2002). On the other hand, the other genera *Neppia* and *Romankenkius* have also an ectal reinforcement confined to the vaginal area and to the zone around the openings of the oviducts (i.e. distal section of the bursal canal) (Sluys, 2001). Furthermore, species such as *N. jeanneli* (DE BEAUCHAMP, 1913) also present an extension of an ectal reinforcement along the bursal canal (Sluys, 2007). Additionally, phylogenetic molecular analyses have shown *Neppia* not to be closely related with *Dugesia* (Álvarez-Presas *et al*., 2008), suggesting a case of evolutionary parallelism. The lack of this character in some *Dugesia* species could be due to a secondary loss in these taxons.

1.3 Development, reproduction and regeneration

1.3.1 Development

All triclads present a direct development, with no larval stages. They have vitellaria and a quite complicated embryonic development. After fertilization, planarians generate an ectolecithal egg or cocoon that is formed in the genital atrium. The egg contains yolk-poor egg cells among many thousands helper yolk cells. These external yolk cells will be ingested by the embryos. The cocoons have a hull membrane that contains several offsprings, between 1 and 20 embryos develop depending on the species (e.g. Cardona

et al., 2005). Their shell is formed by the secretion of cells surrounding the genital atrium. *Dugesia* cocoons are over 2 mm in diameter, round and stalked. They are left attached to the substratum (e.g. aquatic plants and stones) by a pedicel cemented at the base. The whole cocoon splits when *Dugesia* hatchs, leaving the empty shell with the edges curled back. The hatchlings are small replicas of the parents, but they are sexually immature and less pigmented (Ball and Reynoldson, 1981).

1.3.2 Reproduction

There are two different types of reproduction in *Dugesia*, they reproduce either sexually or asexually.

The sexual reproduction is done by adult producing eggs. When sexual, *Dugesia* species do cross-fertilization during the copulation, the male apparatus transfer a spermatophore to the female apparatus of the partner (cf. Sluys, 1989). Because the sperm can be stored for several months, the fertilization may occur much after the copulation (cf. Benazzi and Gremigni, 1982). The breeding cycle of the *Dugesia* is predominantly iteroparous, they breed repeatedly over several seasons (Calow and Read, 1986). Self-fertilization in planarians is rare (Ullyott and Beauchamp, 1931; Benazzi, 1952; Anderson and Johann, 1958).

Asexual reproduction occurs either by parthenogenesis or by fissiparity. Parthenogenetic reproducing animals are sperm-dependent (pseudogamy) and they produce eggs that will hatch clonic offsprings (Beukeboom *et al*., 1996; Beukeboom and Vrijenhoek, 1998). The fissiparous asexual reproduction produces a new generation by transverse fission of the adults. It is not preceeded by any differentation of the new individual (i.e. architomy). The new clones regenerate the missing part of the body thanks to neoblasts recruited to the wound (cf. Baguñà, 1998). Fissiparous populations present no trace of reproductive organs, or they appear underdeveloped.

It is assumed that *Dugesia* sexual reproducing populations present a diploid karyotype while asexual reproducing populations (either fissiparous or parthenogenetic) are triploid and/or tetraploid. Nonetheless, triploid fissiparous *D. ryukyuensis* specimens have been induced to shift from asexual to sexual reproduction when fed with sexual planarians in laboratory conditions (Kobayashi *et al.*, 1999; Chinone *et al.*, 2014). The

originally fissiparous specimens developed both gonads and reproductive organs and started reproducing by copulation. *D. ryukyuensis* is capable to overcome the meiotic problems of chromosomal pairing and segregation characteristic of triploid organisms by different meiotic systems in female and male germ lines. This species can form haploid gametes and reproduce sexually, producing diploid offsprings. It is interesting the idea that this shifting process may be the same in wild populations.

The development of reproductive organs in populations that usually reproduce by fission may be controlled by a neurosecretory process. The agent responsible of the resexualization has been traditionally called 'sex-inducing' or 'sexualizing' substance (Grasso and Benazzi, 1973; Grasso *et al.*, 1975; Benazzi and Grasso, 1977; Sakurai, 1981; Teshirogi, 1986; Hauser, 1987). Those individuals that resexualize are called ex-fissiparous and are characterized by the presence of hyperplasic ovaries and an increased body size.

1.3.3 Regeneration

Planarians are best-known to non-specialists because of their ability to regenerate after injuries and even after being chopped off. This fact was first noticed by Pallas (1774), who described the regeneration capabilities of two species of dendrocoelids (from Brøndsted, 1969). Since then, a great amount of regeneration research has been carried out. The interest on planarians regeneration has raised the two dugesiid species *Schmidtea mediterranea* BENAZZI, BAGUÑÀ, BALLESTER, PUCCINELLI & DEL PAPA, 1975 (Newmark and Sánchez-Alvarado, 2002) and *Dugesia japonica* ICHIKAWA & KAWAKATSU, 1964 (Agata and Watanabe, 1999) to become model organisms in development and regeneration research. Recently, *Dendrocoelum lacteum* (MÜLLER, 1774) has been proposed as a regeneration-impaired planarian model species, after its regeneration abilities are not as good as in other species (Liu *et al*., 2013). There is a different degree of regeneration capability among the triclads, being the Dugesiidae the most regenerative triclad family. In the taxa were asexual reproduction have an important role, the regenerative capabilities are better (or the other way around). Therefore, the ability to regenerate is often linked to the asexual reproduction (Brøndsted, 1969; Sánchez-Alvarado, 2000). Other groups than dugesiids within the Platyhelminthes, such as the macrostomids

(Macrostomidae), have also excellent regeneration capabilities (Egger *et al*., 2007 and references therein).

The regeneration capability of freshwater flatworms when wonded or asexually reproducing is due to the proliferative activity of the neoblasts through the body. These undifferentiated cells or stems cells are responsible for the generation of all germ line cells by mitosis, they can produce all known differentiated cell types (Keller, 1894; Reddien and Sánchez-Alvarado, 2004). The neoblasts give to triclads a great plasticity at the cellular level (Baguñà *et al*., 1989).

1.4 Distribution and ecology

1.4.1 Distribution and dispersal

Dugesiidae representatives are worldwide distributed excepting Antartica, Greenland, Iceland and some oceanic islands. Some genera have a disjunct distribution such as *Romankenkius* (Australia, Tasmania and South America), *Cura* (Australia, Tasmania, New Zealand, North America and South Africa) or *Girardia* (Australia − doubtful, Tasmania and American continent) (Grant *et al*., 2006). On the other hand, the distribution of some dugesiid genera is very restricted, such as the monotypic genus *Bopsula* (São Paulo, Brazil). In contradistinction to the dugesiids, the planariids and dendrocoelids have an exclusive Holarctic distribution.

The representatives of the genus *Dugesia* are widely distributed, being present in Africa, Madagascar, Europe, Middle East, South Asia, Far East and Australasia (Fig. 1.5). This wide geographic coverage has suggested an old origin of the genus, being it possibly contemporary or anterior to the breakage of Gondwana (Ball, 1974b; 1975).

Freshwater planarians spend their entire life cycle into an aquatic environment. Adults are very fragile and very few freshwater planarians have some kind of resting stages resistant to extremes of temperature or desiccation conditions. Although some species have been found living in brackish waters, freshwater flatworms are not able to survive in salt water. Therefore, transoceanic dispersal has been considered very improbable (Ball, 1974a).

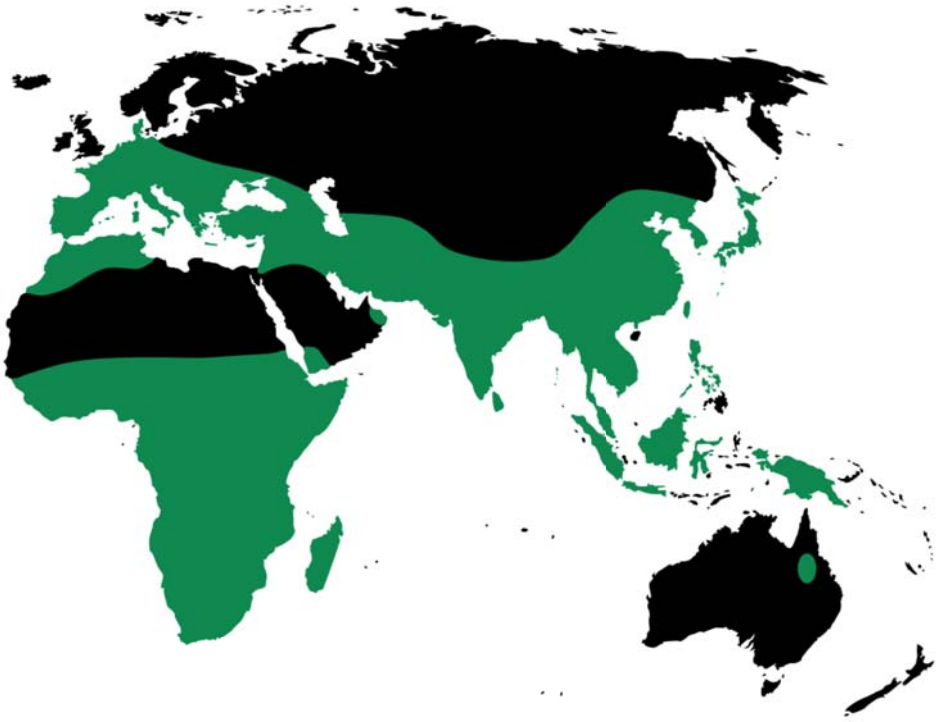


Fig. 1.5 Distribution range map of the *Dugesia* genus species (shaded in green).

One resistant example to desiccation is the planariid species *Hymanella retenuova* CASTLE, 1941, capable to produce a thick-shelled cocoon that can survive dry periods (Ball, 1969a). Specimens of another planariid, *Polycelis nigra* (MÜLLER, 1774) also were found in laboratory conditions to envelope themselves in a gelatinous capsule formed by their own mucus and remaining inactive within the capsule. This has been suggested as a form of resitance to desiccation or starvation (Vila-Farré *et al*., 2011). On the other hand, the dugesiids *Cura pinguis* (WEISS, 1909) and *D. sicula* LEPORI, 1948 are capable of remaining in a moist environment within the stream bed surviving when it has been dried-up (Grant *et al*., 2006; Ribas, 1990).

Ian R. Ball (1974a), in addition to some other previous authors (Ullyott, 1936; Leloup, 1944), suggested that freshwater planarians are poor dispersers because they mainly disperse by their own activities along contiguous freshwater bodies. They are not able to disperse overland (Reynoldson, 1966), but it is possible that they can move through groundwater when soil conditions are suitable (Ball, 1974a). Therefore, freshwater flatworms have been considered as generally slow to colonize new areas (Reynoldson, 1966; Ball and Fernando, 1970; Ball, 1974a). Such low vagility would explain the restriction to particular geographical areas of many freshwater flatworms species. On the other hand, the wide distribution range of the triclad freshwater families

and some genera may be explained by tectonism and/or by different processes of freshwater bodies contact and severing, also proposed as a dispersal way for freshwater fishes (Durand *et al*., 1999; Waters *et al*., 2001). These processes may include river capture (Bishop, 1995), river reversal, or river confluence after sea level lowering.

Passive dispersal of freshwater flatworms has been considered very unlikely. However, the dispersal of freshwater planarians, both cocoons and adults, by floodwaters is documented (Leloup, 1944). A different way of passive dispersal considered by some researchers is dispersal by birds. Specifically, it has been reported that the two planariid species *Crenobia alpina* (DANA, 1766) and *Polycelis felina* (DALYELL, 1814) have dispersed eventually by this way on short distances in northwestern Europe (Dahm, 1958; Reynoldson, 1966). Nonetheless, it is still considered a very improbable way to disperse, so of little significance on a wider scale (Reynoldson, 1966). However, the bird dispersal of cocoons may sound more realistic, as they may be more resistant to aerial journies than adults. Yet, cocoons are frequently layed on the ventral side of rocks and in the case of the dugesiids they are attached to the substratum. Still, this fact makes the aerial dispersal very unlikely.

Human-mediated dispersion has been proven for some planarian species. There are many events of triclad translocation leading to introduced species, such as the freshwater planarians *Girardia tigrina* GIRARD, 1850 from America to Europe and to Japan (Gourbault, 1969; Kawakatsu *et al*., 1993), and *Schmidtea polychroa* SCHMIDT, 1861 from Europe to America (Ball, 1969b). Also *Girardia dorotocephala* WOODWORTH, 1897 from North America can now be found in Hawaii (Schockaert *et al*., 2008) and *D. sicula* in the Canary Islands (De Vries, 1988b; Lázaro and Riutort, 2013), both archipelagos of volcanic origin. These introductions have been proposed to occur because of trade in aquarium plants, exotic fishes or other exotic aquatic animals, for instance (Ball, 1969b; Sluys *et al*., 2010). Moreover, there is a bunch of reports on introduced land planarians in the United Kingdom (UK), mainland Europe (Justine *et al*., 2014), and in North America (Ducey *et al*., 2006). The introduction in Scotland had lead to agricultural problems, after the voracious invasive flatworms had reduced dramatically the earthworm population, those leading to a reduction of drainage and consequently to agricultural production losses (Haria, 1995).

1.4.2 Ecology

Dugesiids inhabit any type of freshwater body, including streams, rivers, lakes or caves, and human created habitats (e.g. channels). Some species have been found living in brackish waters such as *Schmidtea polychroa* in the Baltic Sea (Ball and Reynoldson, 1981).

Because freshwater planarians are negatively phototrophic, they tend to aggregate under rocks and other debries on the bedriver or among the vegetation in response to light. They also show responses to water currents and simple reactions to stimuli such as heat, magnetic fields, and gamma-radiation (Ball and Reynoldson, 1981).

Dugesia species can survive under a wide range of temperatures. For instance, the Circum-Mediterranean *Dugesia sicula* is known to be able to live at temperatures between 10ºC and 25°C (Charni *et al*., 2004; Vila-Farré, 2011) and *D. subtentaculata* from Western Mediterranean is found between 12ºC and 19ºC (L. Leria personal communication). On the other tip of *Dugesia* geographical distribution, *D. ryukyuensis* KAWAKATSU, 1976 from Japan also present a similar tolerance, from 9ºC to 24ºC in the wild (Kawakatsu and Mitchell, 2004).

The feeding of freshwater flatworms is based on a wide variety of invertebrates such as dipters or nematodes (Lischetti, 1919; Stage and Yates, 1939; Koy and Plotnick, 2008). They are essentially predators, but they can feed on damaged or recently dead prey. Planarians are attracted to wonded preys by the chemosensory cells in the auricular grooves on the head. They are also known to feed on frog eggs (M. Vences personal communication).

In general, freshwater planarians have few predators. Dragonfly and damselfly nymphs, and some fish and adult and larval newts are known to eat them (Davies and Reynoldson, 1971). Moreover, stonefly nymph and trichopteran larvas are also known to predate upon freshwater planarians (Wright, 1974). It has also been observed cannibalism and feeding of one species of planarian upon the other (Ball and Reynoldson, 1981). It is thought that flatworms are able to defend themselves against predators, particularly fishes, by extruding the rhabdites, which provide chemical protection.

The food availability has a strong impact on the species diversity, population and individual size. It can also impact on the reproductive strategy in freshwater flatworms, being predominant the fissiparity in less productive places (Reynoldson, 1961). In those habitats with enough energetic resources available, the sexual reproducing species are favored in front of the fissiparous specimens (Calow, 1979; Romero, 1987).

2

Molecular approaches in evolutionary biology

2.1 Molecular phylogenetics

The aim of phylogenetics is to determine the evolutionary relationships of organisms. These relations are depicted as trees that show information on the evolutionary history of the genes and species under study on their topologies (pattern of diversification) and branch lengths (rates of change). Phylogenetic trees present the pattern of descent amongst a group of species or molecules, showing which genes or organisms have a more recent common ancestor (Pagel, 1999). For the obtention of the evolutionary trees for organisms, different kinds of data are used, such as morphological and molecular characters (e.g. nucleotides or amino acid sequences). Those datasets using only molecular information are nowadays predominant because they are able to give more reliable relationships due to their putative objectivity and the higher number of characters available. On the other hand, they are relatively easier to get and analyze. Nonetheless, datasets combining both molecular and morphological information are also widely used (e.g. Lemey *et al.*, 2004; Glenner *et al.*, 2004).

2.1.1 The discordance of gene trees

The phylogenetic trees obtained from the analyses of different genes separately are not necessarly synonyms of the actual species trees. While there is a unique species tree, there are differing genealogical histories (gene trees) for the different DNA sequences used in the analysis. However, although the history of a specific locus could differ from the species history, the similarities among different DNA sequences contain information about the species relationships within a group of related organisms that

have a common ancestor (Maddison, 1997; Slowinski and Page, 1999). The gene trees are embedded within the species lineages.

A common procedure in multilocus analyses is the concatenation of the different genes into a single contiguous sequence of DNA, then it is used in different phylogenetic inference analysis methods. Although some works have concluded that this may result in robust and well-supported phylogenies (Chen and Li, 2001; Rokas *et al*., 2003; Gadagkar *et al*., 2005; Rokas and Carroll, 2005), other have demonstrated that such procedure could fail (Carstens and Knowles, 2007; Kolaczkowski and Thornton, 2004; Kubatko and Degnan, 2007; Mossel and Vigoda, 2005). Concatenated datasets in traditional inference methods may be a problem because they assume that all the data assembled follow a unique gene history, while actually every gene has its own history arising within the common species tree. However, this seems to be more problematic when dealing with recent diverging lineages, prone to incomplete lineage sorting (Carstens and Knowles, 2007).

The no-corresponding genealogies within a species tree may be explained by any of the following processes: Horizontal gene transfer, hybridization, gene duplication, or incomplete lineage sorting. Recombination within a fragment of a gene under study will also have an impact on phylogenetic inference (Maddison, 1997).

The horizontal gene transfer occurs when genetic material of one species is transferred to another, different species. It mostly and commonly take place in bacteria.

The gene or gene cluster duplication process within the genome leads to the generation of paralogous sequences. These sequences should not be confused with orthologous sequences, those that split and evolve independently when a speciation event occur. This means that a single copy of the gene is inherited by each species, they are originated by vertical descent. The inclusion of paralogous genes provides information of the duplication but will be equivocal for the speciation analysis, leading to a gene tree-species tree discordance (Fitch, 1970; Goodman *et al.*, 1979). Therefore, it is of paramount importance to have certainty on the usage of orthologous genes in a phylogenetic analysis, avoiding to mix them with paralogous sequences.

The hybridization is a process that occurs when two distinct species interbreed and a hybrid organisms is generated. This event will ultimately have an impact on the inferred phylogenies because descendants of an hybridization event share some genetic

material from each of the parental organisms. The hybridism has been estimated to occur in about 25% of the plants and 10% of animals (Mallet, 2007). The hybridzation may lead to introgression, that is the stable integration of genetical material from a different species by repeated back-crossing (Rieseberg and Wendel, 1993).

The lineage sorting or *deep coalescence* occurs when multiple gene lineages (ancestral polymorphisms) persist through speciation events. Thus, the ancestral gene copies do not coalesce into a common ancestral copy until much before previous speciation events. *Deep coalescence* is more likely to take place if the population have a bigger effective sample size and the branches of the species tree are short, consisting in few generations (Pamilo and Nei, 1988; Maddison, 1997). On the other hand, when population sizes have been small in comparison with the length of the branches of the phylogenetic tree, then it is more probable for a gene tree to match the species tree.

The gene trees can also present differences in their topologies as a consequence of their different evolution rates. For instance, protein coding mitochondrial genes evolve at higher rates than nuclear ones (Moore, 1995). Thus, mitochondrial genes give more reliable information at shallow diversification events while nuclear genes are more able to solve deeper nodes. Therefore, it is convinient to use molecular sequences evolving at different rates, giving resolution at different levels of the phylogenetic tree. The use of data from multiple genes will allow obtaining a certain estimation of the species tree (Pamilo and Nei, 1988; Takahata, 1989; Wu, 1991; Doyle, 1992).

2.1.2 Evolutionary models

When a molecular phylogenetic analysis is carried out using probabilistic inference methods (Bayesian and maximum likelihood approaches), it is mandatory to set an evolutionary model for the dataset. These models consider the natural process by which one sequence mutates to another over time by taking into account the substitution rates and the nucleotide frequencies (Rosenberg and Kumar, 2003). In addition, two extra parameters could be contemplated in the evolutionary models. First, the discrete gamma approximation (G or Γ) that allows to model for variation in the rate of evolution across sites (Yang, 1994) and second, the parameter invariant sites (I) that considers some sites within a sequence to be unchanging.

The evolutionary models include a wide range of substitution models that use different parameters to describe the relative rates of nucleotide replacement during evolution. The different evolutionary models that result from combining the substitution models, gamma distribution, and presence or not of invariable sites can be evaluated in order to find out which of them fits better our dataset. This evaluation is usually performed using two statistical criteria based on the AIC (Akaike Information Criterion; Akaike, 1974) or BIC (Bayesian Information Criterion; Schwarz, 1978). The evolutionary model that best fits the data has to be calculated for each gene separately before the phylogenetic analysis.

The simplest nucleotide substitution model is the Jukes-Cantor (JC), it assumes that all nucleotide changes occur at the same rates. On the other hand, the most general model, the GTR (general time-reversible substitution model), allows variation in the rates of all possible nucleotide changes. It also allows the condition of time-reversibility. The models of DNA substitution were reviewed by Goldman (1993).

2.1.3 Phylogenetic inference methods

The phylogenetic inference methods are pivotal in the reconstruction of the evolutionary histories of all living creatures. Phylogenetic reconstruction methods perform a tree evaluation by the use of an 'optimality' criterion and by the examination of different tree topologies for a given number of taxa searching for the tree that optimizes this criterion. When comparing sequences in an aligment, each sequence position is a 'character' and the nucleotide or amino acid at the position is a 'state'. All the character positions are analyzed independently.

The most popular phylogenetic inference methods are Maximum Parsimony (MP), Maximum Likelihood (ML) and Bayesian inference (BI). They all work on discrete character-states (e.g. morphological characters or DNA sequence data).

Maximum Parsimony is a non-parametrical statistic method that considers that a tree topology for a given alignment of sequences must be explained with the smallest evolutionary change (substitutions) (Fitch, 1971). Some disadvantatges are associated with it. MP is prone to generate 'long branch attraction' (clustering of those lineages with more changes) because it does not correct for homoplasic states (Felsenstein,

1978). This is because it assumes that a common character state is inherited directly from a common ancestor.

Maximum likelihood algorithms work with numerical optimization techinques in the search for the tree that gives the maximum probability of observing the character states given that tree topology and a model of evolution (Felsenstein, 1981; Pagel, 1999). The likelihood calculation implies the sum over all possible nucleotide or amino acid states in the internal nodes for a particular tree. The tree that yields the highest likelihood is chosen as the best one.

The Bayesian methods, based on the Bayes' Theorem, are conceptually different from MP and ML. They do not search for the single best tree, but for the probability distribution of many inferred trees. Thus, these methods explore for a set of arguable trees or hypotheses for the given data (Huelsenbeck *et al*., 2001). The BI need the specification of prior beliefs, given by the researcher. This is formalized as a prior distribution for the model parameters, like the substitution model parameters or the branch lengths. The posterior probabilities are obtained exploring the tree space with the Markov chain Monte Carlo (MCMC) technique in two independent runs. This technique starts by simulating a random set of parameters and then proposes a new 'state', which is a new set of parameters or a new tree topology. In each step the likelihood ratio and prior ratio are calculated relative to the current step. If the combined product is better, the parameters are accepted and a next step is proposed. Eventually, worse parameters are also accepted. When convergence between the two independent runs is reached a set of probable model/tree solutions is obtained. The initial trees of the chain generated during the initial phase have low likelihood values, because they are influenced by the starting point. In order to use those trees with higher likelihoods that have reached a 'plateau', the initial trees of the chain are discarded (burn-in). The posterior probability for a particular node or tree is proportional to the frequency with which it has been sampled.

2.1.4 Molecular phylogenetics of *Dugesia*

Hitherto, few attempts to settle the phylogenetic relationships among the *Dugesia* species have been carried out (De Vries, 1987; Kawakatsu and Mitchell, 1989; Sluys *et*

al., 1998; Lázaro *et al*., 2009). The first exhaustive phylogenetic analysis ever done included all the species described up to 1998 (68 species). It was based on morphological and karyological information. The result of this analysis came out with a poorly resolved phylogenetic tree with several polytomies but showing big *Dugesia* clades (Fig. 1.6). The resulting lack of resolution in *Dugesia* morphology-based phylogenies is probably due to the low number of morphological characters available for the genus. Moreover, some morphological characters in *Dugesia* do not contain phylogenetic information since the same morphological state can be found in far related species. Therefore, the combination of diagnostic morphological characters is considered sufficient to identify species, but not suitable to find out the evolutionary relationships of the different *Dugesia* species (Sluys *et al*., 1998).

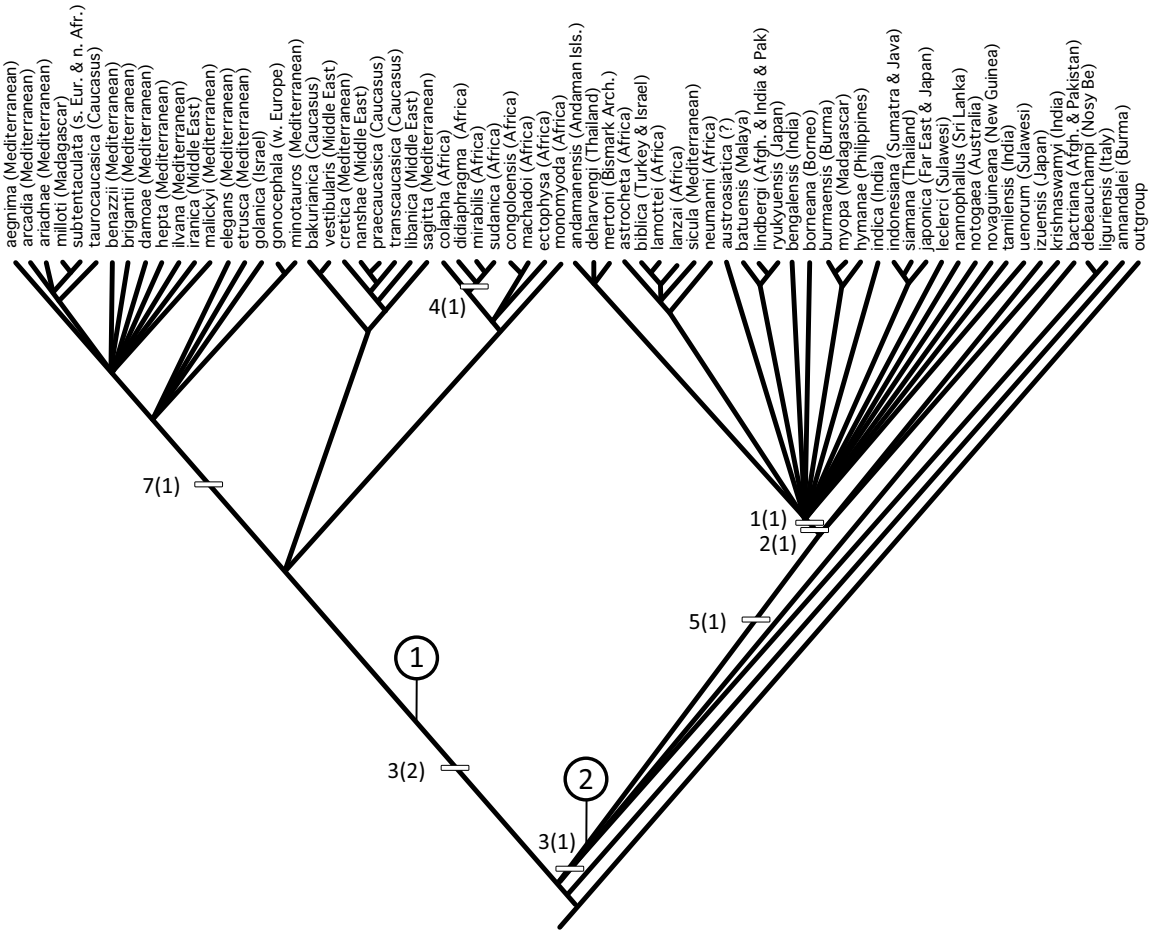


Fig. 1.6 Morphology-based phylogenetic tree based on Sluys *et al.*, 1998. The numbers in circles indicate the major phyletic groups. White bars show the postulated synapomorphies, accompanied by character number and state (see the description in the Fig. 1.3 in the Discussion Section). The major biogeographic regions for each species are also shown.

The premier molecular phylogenetic analyses including *Dugesia* representatives were encompassed in Continenticola focalized works (Carranza *et al*., 1998; Álvarez-Presas *et al*., 2008). The first molecular phylogeny focused on *Dugesia* species was published in 2009, it included many species from the western part of the Mediterranean (Lázaro *et al*., 2009). Later, another work dealing with Chinese *Dugesia* populations was published (Zhang *et al*., 2010). Although these works used a limited number of *Dugesia* representatives, they gave reliable information on the relationships of the specimens under study, supporting the avantatges of using this information on the genus.

2.2 Historical biogeography and phylogeography

Biogeography can be either defined as the study of the geographical distribution of living organisms (Spellerberg and Sawyer, 1999) or as the study of the present and past distribution of animals, plants and other organisms (MacDonald, 2003). Focusing on the second definition, the biogeography is interpreted as an interdisciplinary field that studies the patterns of species distribution in a geographical space through geological time and also identifies natural biotic units (Ball, 1975; Hausdorf and Henning, 2007).

Biogeographers hypothesize about the historical processes that may have shaped the current organisms' distribution patterns. Such observed geographic patterns might be explained by three different kinds of biogeographical processes: dispersal, vicariance and extinction. Biogeography also addresses how ecological factors have determined these distributions (e.g. climatic tolerance and dispersal limitation), as underlying deep historical events (Wiens *et al*., 2004; Riddle *et al*., 2008).

The so-called historical biogeography is considered a sort of biogeography subdiscipline that has long played a key role within the evolutionary biology. Its aim is the reconstruction of species patterns and processes (e.g. speciation, dispersal, and extinction of lineages) that happen over long periods of time in the context of a dynamic Earth history. Historical biogeography uses information from Earth geological sciences such as timing of climate change, orogenies and plate movements, basic phenomena to understand many distribution patterns (Milllington *et al*., 2011). Specifically, the plate tectonics or continental drift theory has impacted heavily on the historical biogeography

since its advent in the 1960's (Dietz, 1961; Hess, 1962). This theory has had a great influence on the causal associations between the geological historical processes and the geographical distribution and divergence of organisms (Brundin, 1966).

The modern historical biogeography is using molecular phylogenetic hypotheses in order to reconstruct the history of one or many taxa (Cox and Moore, 2005; Lomolino *et al*., 2006). The growing availability of molecular data and molecular phylogenies altogether with the increasing accuracy of the knowledge of the history of the Earth and methods to date lineage divergences have provided robustness and have made more attractive the study of the historical diversification of life on Earth in a biogeographical context (Riddle *et al*., 2008). This has led to an increasing number of studies and methodologies with the aim to infer the history of various taxa (e.g., Ree *et al*., 2005; Wojcicki and Brooks, 2005; Ree and Smith, 2008).

2.2.1 Historical biogeography of *Dugesia*

The freshwater triclads are considered suitable organisms to perform biogeographical analyses (Ball, 1974a; 1983) because they are organisms of low vagility. The planariologist Ian R. Ball wrote on the biogeography of freshwater planarians that 'a causal explanation of their distribution must take careful consideration of historical events. Further, since the history of a taxon in nature is reflected by both its morphology and its distribution, a causal explanation of distribution is intimately concerned with the evolutionary relationships of its members.' An updated version of this sentence would include a reference to 'a reflection by its molecules' apart from the morphology.

However, the first problem when doing research on freshwater flatworm historical biogeography is the lack of fossils that could shed light on the evolutionary history of the group, giving a minimum age for certain planarian clades in certain geographic areas.

Some planariologists have speculated with the origin of *Dugesia* and its family Dugesiidae, taking into account their present distribution range (Fig. 1.5). Kawakatsu (1968) placed the origin of the dugesiids in the Balkan Peninsula, because this area was considered an evolutionary center (Stanovic, 1960). From there, the genus would have dispersed until covering its present distribution. However, this hypothesis was

invalidated by Ball (1974a), who suggested instead the origin of the dugesiids to be in south of the present-day equator. These opposite proposal placed the origin of the family Dugesiidae in Gondwana, in what is Antarctica today. However, he also suggested that possibly over 220 Mya in the early Mesozoic times or even earlier in the aftermath of the Permo-Carboniferous glaciations, dugesiids diversified on the Pangaea supercontinent. According to Ball, the genus *Dugesia* arose in Africa when the breakage of Gondwana already had started. Once the Tethys Sea was closed, the genus dispersed northwards in Eurasia probably using a route from Africa through Middle East to Europe and Asia.

Sluys and collaborators updated Ball's hypothesis in 1998. They agreed with the Pangaean Dugesiidae family origin hypothesis and proposed two possible ways of *Dugesia* dispersal in Eurasia from Gondwana former lands. The first proposal explained a release in Asia after India collided with the continent. The Indian subcontinent split from Madagascar 88 Mya and rapidly drifted northwards, colliding about 40−20 Mya. The second explanation for *Dugesia* dispersal in Eurasia was proposed to have happened through the impact of the Arabian plate with it, around 20 Mya. Finally, Sluys and collaborators explained the presence of *Dugesia* on Northern Australia from Southeastern Asian populations that probably dispersed during the Pleistocene.

2.3 Divergence time estimation

The first attempts to estimate the times of lineages divergence were based on the assumption that gene sequences accumulate mutations at a roughly constant rate over time (Zuckerkandl and Pauling, 1962; 1965). This so-called molecular clock hypothesis was in agreement with the neutral theory of evolution (Kimura, 1968; 1983). Its rate could be estimated looking at the fossil record.

However, real molecular data often does not behave in such a 'clock way' (e.g. Britten, 1986). The evolution rate is dependent on many factors, being neither constant along time nor between lineages. These factors influencing the rate may include the underlying mutation rate, metabolic rates in a species, generation times, population sizes, and selective pressure (Bromham and Penny, 2003).

Since the advent of the molecular clock hypothesis, the divergence time estimation has become much more sophisticated, taking into account its uncertainties. To consider the non-clock behaviour, the relaxed molecular clock models can accommodate variation in the evolution rate when estimating divergence dates, allowing any number of local molecular clocks (Drummond *et al*., 2006; Yoder and Yang, 2000). Moreover, the new 'relaxed' methods of dating and tree reconstruction have lead generally to a better concordance between molecular and paleontological dates (e.g. Smith *et al*., 2006).

Approaches developed to the present allow the researchers to use dates to calibrate nodes on a phylogenetic tree taking into account their uncertainties, assessing an upper and lower bound, or a probability distribution (Drummond *et al*., 2006; Kishino *et al*., 2001; Yang and Rannala, 2006). At present a common procedure is the use of multiple calibration points instead of a single point to estimate the clock rates, each point being associated with a probability distribution that summarizes the available information (Yang and Rannala, 2006).

The phylogenies are often dated using the fossil record to calibrate interior nodes (e.g. Ronquist *et al*., 2012). But some well-documented problems are associated to this method (Bromham *et al*., 1999; Pérez-Losada *et al*., 2004): the incomplete fossil records and the difficulty to stablish the actual relationships between the fossils and the extant organisms. Additionally, fossils just give a minimum time back to the common ancestor of a particular taxon (Benton and Ayala, 2003) and there is an error associated with their process of datation (Magallon, 2004). However, fossil record is poor or non-existant for many living lineages, including the whole Platyhelminthes phylum. In these cases, the alternative to the fossil record for calibrating phylogenetic trees is the use of mutation rates obtained for other groups, secondary calibration points, or geological or paleogeographical information.

These geological events would include tectonic drift, island formation, or mountain range uplifts, among others. Calibrations based on such events may provide a method for formulating and testing evolutionary hypotheses and help in the understanding of biodiversification timeframes (Rambaut and Bromham, 1998; Bromham, 2003; Bromham and Penny, 2003; Sanmartín and Ronquist, 2004). However, much contention has been also held on the use of geological events in calibrating trees. The main problem is the difficulty to know how well a geological date is corresponding to the time at which lineages split (Heads, 2005; Magallon, 2004). For instance, some

studies had used the formation of islands to estimate substitution rates, such as the rise of Hawaii and the Canary Islands (e.g. Price and Clague, 2002; Gubitz *et al*., 2000). However, the lineages on these islands may have diverged at a different time than their geological formation. The same situation would apply on lineages distributed on former united landmasses that split before or after their geological breakage.

2.3.1 Divergence time estimation of Dugesiidae

Up-to-date, the only work that has carried out a divergence time estimation of a freshwater flatworm phylogenetic tree is one dealing on *Schmidtea mediterranea* (Lázaro *et al*., 2011). This study used a short fragment of the Cox1 gene in order to obtain a phylogenetic tree including both planariids and dugesiids. The tree was calibrated using the drift of Africa from South America about 100 Mya, considering it as the causal event of the *Girardia* split from its sister group, as Ball proposed (Ball, 1974a). The mitochondrial gene Cox1 substitution rate was found to be very slow, 0.0027 mutations per site per million years (0.27% substitutions per million years). Later, this rate was applied to a phylogenetic tree including only specimens of *Schmidtea mediterranea*. The divergence time estimation result showed a putative old origin (∼20−4 Mya) for this species.

2.4 Molecular species delimitation

The definition of the *species* concept is still a matter of debate (De Queiroz, 1998; 2005; 2007; Hausdorf, 2011). However, it has fairly reached a sort of consensus; a *species* is described as a lineage of populations or metapopulations that evolves indepenently from the others through time (Simpson, 1951; Wiley, 1978; De Queiroz, 2005), and it is commonly interpreted as evolutionary significant units (Moritz, 1994). The main disagreement is on where along the divergence continuum two different lineages should be recognized as two different species (Hey, 2006).

De Queiroz (2005) argued that all species concepts defined until then, such as the genealogical, morphological, or reproductive concepts, are indicators of species-level differentiation, and they have not to be considered independently but altogether.

These different traditional criteria that have been taken separately until then, are all informative attributes that accumulate during the process of lineage diversification. This idea is called the Generalized Lineage Concept (GLC) and it has indirectly promoted many recent approaches to species delimitation (e.g. Knowles and Carstens, 2007). However, not all the researchers working on species delimitation have adopted the GLC concept or it is an unanimous prerequisite for species delimitation (e.g. Rosell *et al*., 2010; Barrett and Freudenstein, 2011; Duminil *et al*., 2012).

The traditional species description, based essentially on morphological data, may ignore such independent lineages of populations and metapopulations by being unable to detect processes such as reticulated evolution or cryptic species, specially in those cases of diversification at an early stage. Now, molecular data is used as a tool to detect these overlooked cryptic or problematic lineages. Notably, molecular species delimitation approaches have increased the rate of candidate species delimitation (e.g. Morando *et al*., 2003; Mayer, 2007; Vieites *et al*., 2009), the identification of cryptic species (reviewed by Bickford *et al*., 2007) and the identification of 'young' species (Knowles and Carstens, 2007). Interestingly, many of the studies using a molecular-based species delimitation approach are focused on taxonomically understudied organism groups and they usually analyze data from a single genetic locus in order to get a preliminary estimation of the species diversity (Carstens *et al*., 2013). Thus, even the simpler molecular-based approaches are convenient tools to boost the taxonomic knowledge of many poorly studies or complicated groups.

In recent years, the number of methods available for molecular based species delimitation have experienced a great increase (Carstens *et al*., 2013; Leaché *et al.*, 2014) and at the present they are widely used. Such methods of molecular species delimitation range from non-parametric (e.g. Wiens and Penkrot, 2002) to highly parameterized (e.g. Yang and Rannala, 2010). One of the main reasons that explains their popularity is the great advantatge attributed to these methods to delineate species objectively, bringing a statistical framework to detect independent evolving lineages.

The methodologies of molecular species delimitation that are based on the coalescent theory may include either single locus or multiple loci approaches. Among those dealing with just one locus the most popular is probably the General Mixed Yule-Coalescent (GMYC; Pons *et al*., 2006). This method is able to distinguish the coalescent from the speciation processes by looking at the branching pattern of an

ultrametric phylogenetic tree obtained from mitochondrial markers. GMYC plots a threshold when it identifies a dramatic increase of such branching pattern, pointing those entities beyond this limit to be different species. It is specially useful when dealing with understudied groups as it may lead further investigation to more detailed species delimitation on such unknown organisms.

For instance, those entities delimited by the GMYC may be also used for individual assignment to putative species and thus be further used in multilocus species delimitation methods that require previous individual assignments, like those implemented in bpp or spedeSTEM (Satler *et al*., 2013). Other multilocus-based programs do not need *a priori* information assigning individuals to putative species, for instance STRUCTURAMA (Huelsenbeck *et al*., 2011) and BROWNIE (O'Meara, 2010).

At the moment, the most popular species tree-based method for species delimitation is that implemented in the program bpp. On the basis of the genetic alignments, the assignment of individuals to putative species, and the input of a guide species tree, it performs statistical estimations testing if the assigned putative species are distinct (Yang and Rannala, 2010). Thus, bpp keeps or lump together the different putative species used as an input. A different program working in a similar way is spedeSTEM (Carstens and Dewey, 2010).

The species delimitation is a matter of great interest in evolutionary biology. The molecular methodologies and approaches dealing with this challenge use information from phylogenetics and population genetics in order to distinguish when the processes at population level start to produce phylogenetic patterns indicating speciation. However, all the existing molecular methods to delimit species are only capable to delineate evolutionary lineages with accurancy under some plausible set of conditions. Therefore, the most advisable way to delimit species may be to analyze molecular data using different delimitation methods and to delimit lineages that are consistent or not exclusive across results and data sources (e.g. DNA, morphology, ecology or behaviour).

2.4.1 The integrative taxonomy

Since the advent of the Linnean nomenclature in 1758 comparative morphology has been predominant in the species discovery. Now, new methodologies and the integrative usage of different data sources are reshaping and revitalizing the taxonomic field.

The ability to assign individuals to species and detect species limits is highly impacted by the evolutive processes underlying the speciation of different groups. Therefore, the species delimitation approaches must be conducted taking into account the life history, geographical distribution, morphology and behaviour among other data sources when possible (Knowles and Carstens, 2007; Schlick-Steiner *et al*., 2010). Those approaches based purerly on genetic methods could be prone to an inadequate description of the diversity (Harrington and Near, 2012). Therefore, it is necessary to consider data types in a wider context, using non-genetic data sources along various concepts of species (Carstens *et al*., 2013; Edwards and Knowles, 2014). This idea is expressed in the framework of the integrative taxonomy. For instance, such integrative approaches would be appropriate in many cases where morphological evidences provide taxonomic clues of a new species, while molecular-based methods are therefore used to validate or reject such hypothesis (e.g. Carstens and Dewey, 2010; Welton *et al*., 2013) or the other way around.

Few statistical methodologies of species delimitation beyond the genetic data are available. However, it seems probable that they will become more popular and visible in the coming years. Not many integrative studies are using methodologies such as ecological niche modelling (Peterson, 2001; Hugall *et al*., 2002; Bond and Stockman, 2008; Zhou *et al*., 2012), a promising but still understudied framework. Other new integrative approaches such as a statistical framework for species delineation combining ecological, morphological and molecular data are now appearing (Edwards and Knowles, 2014). This method considers statistically the inclusion of the variance in intrinsic characters and allows to avoid the overdescription of species or taxonomic inflation in species delineation thanks to the use of different species concepts in qualitative taxonomic frameworks (Issac *et al*., 2004).

If morphological and genetic evidences are not congruent, it is still common to adopt a conservative approach, preventing a new species description (e.g. Leliart *et al*.,

2009; Barrett and Freudenstein, 2011) in order to avoid biodiversity inflation, but considering the possiblity of morphologically cryptic species (e.g. Salter *et al*., 2013). In such cases further data type is desirable in order to obtain more evidences rejecting or supporting the species hypothesis. In this way the description of new species will not falsely delimit entities that do not match actual evolutionary lineages. Indeed, this is the considered the essence of the integrative taxonomy, taxonomic inference should be based on congruence among multiple analyses and data sources (e.g. Padial *et al*., 2010; Schlick-Steiner *et al*., 2010). However, some researchers are now claiming for molecular data to be enough for formal species description, using it as a diagnostic character for species erection (Jörger and Schrödl, 2013).

2.4.2 Integrative taxonomy on *Dugesia*

Hitherto, only one 'integrative taxonomy' work has been carried out involving triclads in general, and *Dugesia* in particular. It is a paper by Stocchino and collaborators (2013), that used morphological, molecular, karyological, and cytogenetic data available from the literature to describe a new species, *Dugesia superioris* STOCCHINO & SLUYS, 2013. However, such molecular method consisted in the recognition of a distinct branch in a phylogenetic tree published by Lázaro and collaborators (2009). No statistical approach was carried out.

A preliminary species delimitation by GMYC method was also done on different *Dugesia* species of the Western region of the Mediterranean (Lázaro, 2012). This analysis recognized most of those species already described and it only oversplit one case, *D. benazzii*.

Section II

Objectives

Objectives

General Objective

The main and general objective of the present thesis is to widen and shed new light on the knowledge of the diversity, historical biogeography and/or evolution of the freshwater planarian genus *Dugesia* in the Northeastern Mediterranean in particular and across its whole distribution in general. Thus, we aimed to understand how have these organisms been so successful, covering a very large distribution area and diversifying in a relative high number of species.

Particular Objectives

To perform a wide sampling of *Dugesia* including Eurasia, Africa, Madagascar and Australasia, and use molecular data and up-to-date phylogenetic, biogeographic and taxonomical methodologies in order:

- To find out a putative origin age of the genus *Dugesia*.
- To infer the impact of historical processes on the distribution patterns and diversification of the genus *Dugesia* in the northeastern Mediterranean and across its whole distribution.
- To obtain molecular substitution rates to be used in future studies on *Dugesia* and other triclad species.
- To test a molecular-based species delimitation method on *Dugesia* populations and use the results together with morphological data to describe new species found.

To obtain complete mitochondrial genomes of triclads representatives in order to facilitate the use of further molecular markers in future studies on this complicated group and to explore the evolutionary history of the Platyhelminthes phylum.

Section III

Publications

1

Supervisor report

Supervisor report

Dr. Marta Riutort León, supervisor of the doctoral thesis prepared by Mr. Eduard Solà Vázquez, entitled "*Systematics and Historical Biogeography of the genus Dugesia*" reports that the thesis is made as a compendium of four publications with original data (1-2-3-4 items in the main part of the thesis):

Article 1

Solà E, Sluys R, Gritzalis K, Riutort M (2013). Fluvial basin history in the northeastern Mediterranean region underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae). *Molecular Phylogenetics and Evolution*, 66:877−888.

Impact factor 4.066 (2012). Rank **15** (of **47, Q2**) in the category *Evolutionary Biology*.

Article 2

Solà E, Stocchino GA, Manconi R, Leria L, Harrath H, Riutort M (en preparació). *Dugesia* (Platyhelminthes, Continenticola), a widespread and morphologically homogeneous living genus from the Mesozoic.

Article 3

Sluys R, Solà E, Gritzalis K, Vila−Farré M, Mateos E, Riutort M (2013). Integrative delineation of species of Mediterranean freshwater planarians (Platyhelminthes: Tricladida: Dugesiidae). *Zool J Linn Soc* 169:523−547.

Impact factor 2.583(2012). Rank **18** (of **151, Q1**) in the category *Zoology*.

Article 4

Solà E, Álvarez-Presas M, Frías-Lóprez C, Littlewood DTJ, Rozas J, Riutort M (en preparació). Evolutionary analysis of mitogenomes from parasitic and free-living flatworms.

Contributions of the candidate to the articles. The doctoral student participated in two of the three sampling trips in Greece for publications 1 and 3. Has obtained molecular data for articles 1-3, and was in charge of the PCR amplification and sequencing of

Crenobia alpina mitogenome for article 4. Performed the phylogenetic, biogeographic and species delimitation analyses of articles 1-3. Annotated *Crenobia alpina* mitogenome, and performed the statistical analyses comparing mitochondrial genomes from parasites and triclads for article 4. Wrote the initial draft of the manuscripts of articles 1, 2 and 4, and participated in writing the final version of all articles. The work presented in this thesis has not been used, implicitly or explicitly, for the preparation of another thesis.

Barcelona, 9 May 2014

Signed: Marta Riutort

2

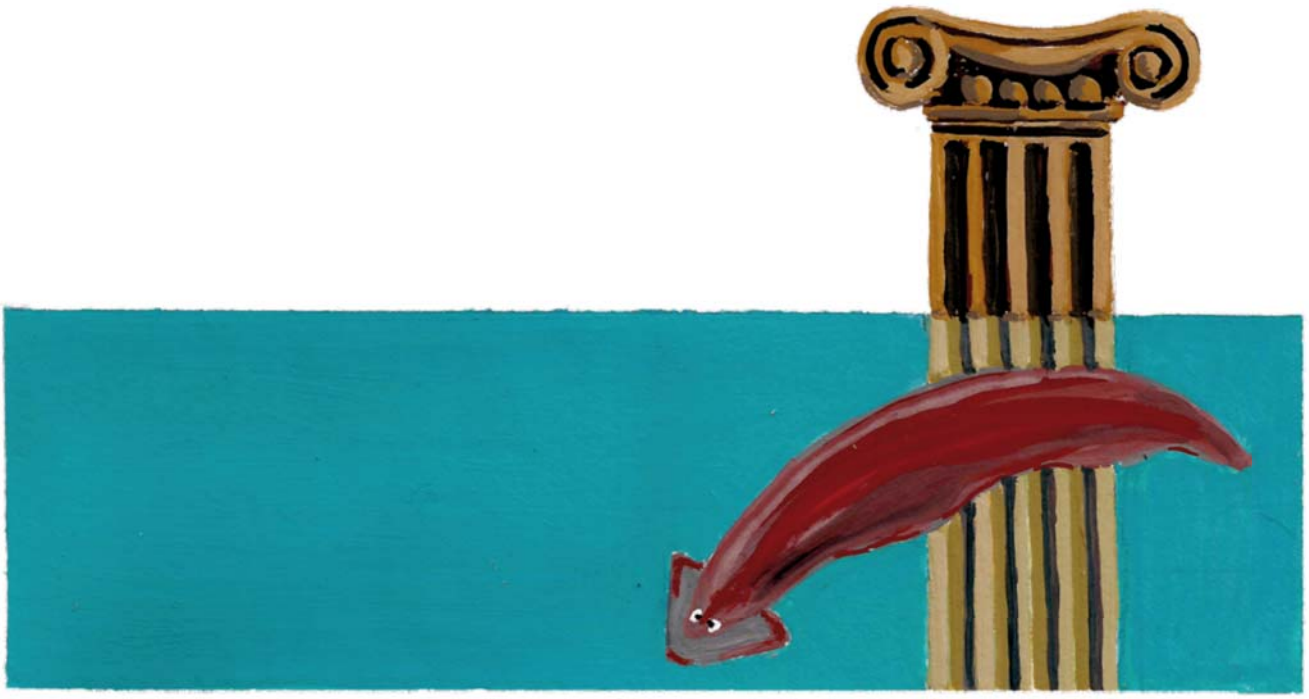
Historical biogeography and systematics

Chapter 1

Fluvial basin history in the northeastern Mediterranean region underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae)

Reference

Solà E, Sluys R, Gritzalis K, Riutort M (2013). **Fluvial basin history in the northeastern Mediterranean underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae).** *Mol Phylogenet Evol* **66**:877–888.



Summary

The aim of this paper was to make the first attempt to obtain an estimation of divergence times for the genus *Dugesia*. The molecular biogeographical work was focused on a geographical region which paleogeological history is reasonably well-established and which is known to harbor a relatively rich diversity of species of the genus *Dugesia* (9 species when we started this work). Therefore, we considered this area as a suitable place to carry out this approach. On the other hand, we aimed to expand the former knowledge of the phylogenetic relationships of the different species of the genus *Dugesia*.

The results showed a quite well-defined biogeographical structure of the genus representatives on the area. Interestingly, specimens of the species *Dugesia cretica* from three sampling localities on Crete appeared to be the sister group of the rest of species in the Aegean region. Crete was the first island to become isolated from the former united landmass called Ägäis about 11−9 Mya, which pointer to this split as a good calibration point for the estimation of the divergence times inthe phylogenetic tree. Another event apparently mirrored in the topology of the Greek *Dugesia* species was the advent of the Mid-Aegean trench (c. 12−9 Mya). This event split the region in a Western and an Eastern part and had an impact on the fauna of the region. Evidences of dispersal during the Messinian Salinity crisis were also found (c. 5.6−5.3 Mya), as well as possible extinctions in Western Greece followed by colonizations from the north and geographical expansions within the Peloponnese peninsula. We also found evidences of human-mediated dispersal from this peninsula to the island of Crete and Cephalonia.

This is the first attempt to use paleogeographical events to obtain diversification times and substitution rates of genes for *Dugesia*. This information is added to the same previous kind of approach of *S. mediterranea* by Lázaro and collaborators (2011).

Molecular Phylogenetics and Evolution 66 (2013) 877–888

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Fluvial basin history in the northeastern Mediterranean region underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae)

Eduard Solà [a], Ronald Sluys [b], Konstantinos Gritzalis [c], Marta Riutort [a,*]

[a] Departament de Genètica, Facultat de Biologia and Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Catalonia, Spain
[b] Naturalis Biodiversity Center, Leiden and Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, The Netherlands
[c] Hellenic Centre for Marine Research, Institute of Inland Waters, Anavyssos, Greece

ARTICLE INFO

Article history:
Received 2 May 2012
Revised 8 November 2012
Accepted 10 November 2012
Available online 24 November 2012

Keywords:
Aegean
Biogeography
COI
Divergence times
Freshwater planarians
ITS-1
Molecular clock
Phylogeography

ABSTRACT

In this study we analyzed the phylogenetic relationships of eastern Mediterranean freshwater planarians of the genus *Dugesia*, estimated divergence times for the various clades, and correlated their phylogeographic patterns with geological and paleoclimatic events, in order to discover which evolutionary processes have shaped the present-day distribution of these animals. Specimens were collected from freshwater courses and lakes in continental and insular Greece. Genetic divergences and phylogenetic relationships were inferred by using the mitochondrial gene subunit I of cytochrome oxidase (COI) and the nuclear ribosomal internal transcribed spacer-1 (ITS-1) from 74 newly collected individuals from Greece. Divergence time estimates were obtained under a Bayesian framework, using the COI sequences. Two alternative geological dates for the isolation of Crete from the mainland were tested as calibration points. A clear phylogeographic pattern was present for *Dugesia* lineages in the Eastern Mediterranean. Morphological data, combined with information on genetic divergences, revealed that eight out of the nine known species were represented in the samples, while additional new, and still undescribed species were detected. Divergence time analyses suggested that *Dugesia* species became isolated in Crete after the first geological isolation of the island, and that their present distribution in the Eastern Mediterranean has been shaped mainly by vicariant events but also by dispersal. During the Messinian salinity crisis these freshwater planarians apparently were not able to cross the sea barrier between Crete and the mainland, while they probably did disperse between islands in the Aegean Sea. Their dependence on freshwater to survive suggests the presence of contiguous freshwater bodies in those regions. Our results also suggest a major extinction of freshwater planarians on the Peloponnese at the end of the Pliocene, while about 2 Mya ago, when the current Mediterranean climate was established, these Peloponnese populations probably began to disperse again. At the end of the Pliocene or during the Pleistocene, mainland populations of *Dugesia* colonized the western coast, including the Ionian Islands, which were then part of the continent.

1. Introduction

The Mediterranean Basin has a long and complex geological history and is therefore generally considered as an excellent laboratory region to study the effect of paleogeographic events on the evolutionary diversification of fauna and flora. This has resulted in a wealth of studies that focus on the biogeography and evolution of taxa in the western or in the eastern Mediterranean (e.g. de Jong, 1998; Veith et al., 2004; Lázaro et al., 2009; Lymberakis and Poulakakis, 2010).

Especially the northeastern Mediterranean region is well suited for phylogeographic studies to unravel the historical processes that underlie present-day species distributions and current levels of diversity and endemism (Sfenthourakis and Legakis, 2001). This area has been subjected to tectonism, volcanism and sea level changes since the Miocene (Dermitzakis, 1990; Perissoratis and Conispoliatis, 2003), resulting in a complex geological history. The major events in the geological history of the Aegean area are relatively well known. The Aegean archipelago started to form c. 16 million years ago (Mya), when the single landmass Ägäis started to fragment (Dermitzakis, 1990) as a consequence of the collision of the African/Arabian tectonic plates with the Eurasian plate

* Corresponding author. Address: Departament de Genètica, Facultat de Biologia and Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Av. Diagonal 643, 08028 Barcelona, Catalonia, Spain. Fax: +34 934 034 420.
E-mail address: mriutort@ub.edu (M. Riutort).

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2012.11.010

(Krijgsman, 2002). The opening of the mid-Aegean trench (MAT) started at c. 12 Mya when the sea invaded the land from south to north, starting between Crete and Kasos–Karpathos; at c. 9 Mya the previously uniform landmass became divided into an eastern and a western Aegean sections (Dermitzakis and Papanikolaou, 1981). At about 10 Mya Crete was the first island to become separated from the mainland (Dermitzakis, 1990; Cosentino et al., 2007). Apart from tectonic fragmentation events, the Hellenic area also experienced several sea level changes, such as during the Messinian salinity crisis (MSC) (5.96–5.33 Mya; Krijgsman et al., 1999) and during the Pleistocene glaciations (2.58 Mya–11.7 kya; Perissoratis and Conispoliatis, 2003), thus leading to contact between previously isolated landmasses and ancient river drainage systems or to the severance of single landmasses and river basins (Maurakis et al., 2001).

In the past few years an increasing number of studies have carried out historical biogeographic analyses on a wide range of organisms in this region, such as snails (e.g. Parmarkelis et al., 2005; Kornilios et al., 2009), arthropods (e.g. Poulakakis and Sfenthourakis, 2008; Papadopoulou et al., 2009; Parmakelis et al., 2006), reptiles (e.g. Kasapidis et al., 2005; Poulakakis et al., 2003, 2005), frogs (Akın et al., 2010), and plants (Bittkau and Comes, 2005). Most of these studies conclude that the evolutionary diversification of organisms in the northeastern Mediterranean has been driven by vicariance induced by geological and marine barriers. In general, the three divergence patterns proposed by Lymberakis and Poulakakis (2010) can be recognized among the taxa in this region: (1) species already present before breakup into several component areas, (2) species that reached the area after the formation of the MAT (after c. 9 Mya), and (3) much more recent, human-mediated arrivals. Nevertheless, differences in the organisms' biology and ability to disperse can result in different responses to the geological history of the area and, therefore, to differences in current patterns of distribution (Douris et al., 2007).

In this study we used freshwater planarians of the genus *Dugesia* Girard, 1850 as a model to examine the effect of the paleogeography of the Hellenic region on the evolutionary diversification of its component fauna. For this purpose, the genus *Dugesia* is an ideal model group, in view of the fact that (1) the Mediterranean region is a hotspot of biodiversity, with over 20 species from a world total of about 75 species, (2) freshwater planarians do not possess larval dispersal stages and do not tolerate salt water and thus need contiguous freshwater bodies to survive and disperse (Ball and Fernando, 1969; Ball, 1975). A recent study on Mediterranean *Dugesia* species revealed a clear correspondence between phylogenetic relationships and paleogeography (Lázaro et al., 2009). Unfortunately, virtual absence of planarian fossils prevents absolute dating of divergence times and neither did paleogeographic information facilitate calibration of a molecular clock, thus impeding precise dating of the phylogeographic patterns. Further, that study concentrated on species in the western Mediterranean, in contrast to our present focus on the eastern Mediterranean region.

For the present study we sampled numerous *Dugesia* populations distributed across the northeastern Mediterranean region, comprising populations from Greek islands as well as the mainland (Fig. 1). We generated a calibrated phylogenetic tree for these populations, with the aim to examine the effects of geological processes, paleoclimatic events, and anthropogenic dispersal on the historical diversification and current distribution of these planarians in this region. Furthermore, we also set out to examine the correlation between molecular and morphological markers in species determination.

2. Materials and methods

2.1. Sample collection

Dugesia specimens were collected from the type localities of eight Greek species (de Vries, 1984, 1988) and from other localities on the mainland and some islands during the spring seasons of 2009 and 2010. For each locality some specimens were fixed and preserved in absolute ethanol for molecular analysis. Other animals were fixed with Steinmann's fluid (cf. Sluys, 1989) for morphological analyses and were, subsequently, preserved in 70% ethanol. For information on sampling localities, see Table 1 and Fig. 1.

2.2. Morphological analysis

Specimens that had been preserved for morphological analysis were cleared in clove oil and then embedded in paraffin wax, sectioned at intervals of 6 or 8 μm (depending on the size of the animals) and mounted on albumen-coated slides. Sections were stained in Mallory-Cason/Heidenhain (Humason, 1967; Romeis, 1989) and mounted in DPX. Reconstructions of the copulatory complex were obtained by using a camera lucida attached to a compound microscope. All material has been deposited in the collections of the Netherlands Center for Biodiversity Naturalis, Leiden, Netherlands.

2.3. Sequencing procedure

Total genomic DNA extraction was performed on two individuals fixed in absolute alcohol per sample locality, using the commercial reagent DNAzol (Molecular Research Center Inc., Cincinnati, OH) following the manufacturer's instructions.

Specific primers were used to amplify a fragment of the mitochondrial gene cytochrome c oxidase subunit I (COI) and the nuclear ribosomal internal transcribed spacer-1 (ITS-1) sequences. Sequences and annealing temperatures for each pair of primers are given in Table 2. Final PCR reaction volume for all molecules was 25 μl. To 1 μl of DNA sample to amplify we added: (1) 5 μl of Promega 5× Buffer, (2) 1 μl of dNTP (10 mM), (3) 0.5 μl of each primer (25 μM), (4) 2 μl of $MgCl_2$ (25 mM), (5) 0.15 μl of *Taq* polymerase (GoTaq® Flexi DNA Polymerase of Promega). Double-distilled and autoclaved water was added to obtain the final PCR volume. In order to obtain amplification of the sequences it was necessary in many cases to vary the annealing temperatures or the amount of $MgCl_2$ or DNA.

PCR products were purified before sequencing using the purification kit illustra™ (GFX™ PCR DNA and Gel Band of GE Healthcare) or by using a vacuum system (MultiScreen™$_{HTS}$ Vacuum Manifold of Millipore). Sequencing reactions were performed by using Big-Dye (3.1, Applied Biosystems) with the same primers used to amplify the fragment, except for the forward COI sequence that was obtained with a more internal primer (COIEF3), due to sequencing problems when using BarT. Reactions were run on an automated sequencer ABI Prism 3730 (Unitat de Genòmica of Serveis Científico-Tècnics of the Universitat de Barcelona). Obtained chromatograms were visually checked.

2.4. Sequence alignment and genetic divergence

An approximate 750 bp fragment of the mitochondrial gene COI and an approximately 700 bp fragment of ITS-1 were sequenced. Additionally, sequences of other *Dugesia* species available in GenBank were retrieved (Table 1). Alignments of the sequences were obtained with the online software MAFFT version 6 (Katoh and

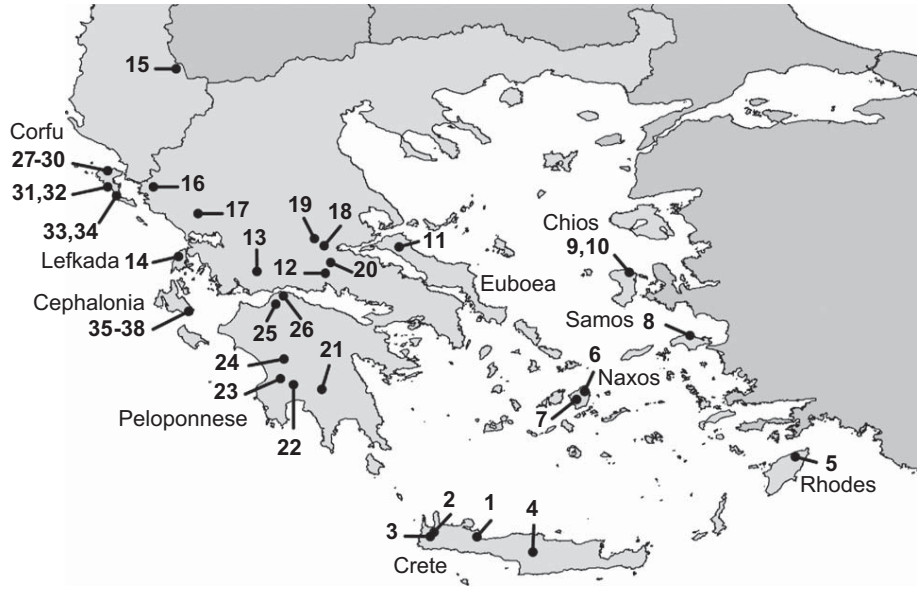


Fig. 1. *Dugesia* localities sampled in Greece; numbers correspond to the locality codes reported in Table 1. The Albanian population (15; cf. Lázaro et al., 2009) is also shown in the map.

Toh, 2008) and were manually edited with the software BioEdit (version 7.0 for PC) (Hall, 1999). Prior to analyses, the COI sequences were translated into amino acids showing no stop codons. Equivocal positions of ITS-1 alignment were removed with the software Gblocks (Talavera and Castresana, 2007), allowing half gap positions in the alignment. Genetic divergences among individuals were calculated with MEGA 5.0 computer package (Tamura et al., 2011) using the Kimura 2-parameters correction.

2.5. Phylogenetic analysis

Level of sequence saturation was analyzed by plotting observed transitions and transversions against the divergence for COI and ITS-1 under the TN93 nucleotide substitution pattern model with the program DAMBE (Xia and Xie, 2001). Three data sets were analyzed: ITS-1, COI, and an alignment with both molecules concatenated. Phylogenetic analyses were performed using two inference methods: Maximum Likelihood (ML) and Bayesian inference (BI). In all the likelihood and Bayesian analyses we set GTR + I + Γ as evolutionary model, leaving the inference programs to estimate all the parameter values and hence the best model. In the analyses of the concatenated data set we set the parameters estimation as unlinked. ML analyses were performed with the program RaxML 7.0.0 (Stamatakis, 2006). 1000 replicates were calculated to obtain bootstrap supports. BI was conducted using the program MrBayes (v. 3.2: Ronquist et al., 2012). Given the high number of terminals we ran 1 cold and 4 heated chains for two runs to ensure a better sampling in the tree space. 1,000,000 generations were performed for each gene, saving a tree every 100 generations. The convergence of the topologies and model parameters of both runs was surveyed by checking that the standard deviation of the split frequencies reached a value below 0.01 (default burn-in = 25%). In order to infer the topology and the posterior probabilities we discarded the first 25% of trees for COI, ITS-1, and concatenated data sets in order to avoid inclusion of those trees obtained before likelihood values had stabilized, which were checked by plotting likelihoods against generations, and both runs had converged.

In a preliminary analysis, the genus *Schmidtea* Ball, 1974 was used as the outgroup (sister group of *Dugesia*; cf. Álvarez-Presas et al., 2008) to determine the root for the genus *Dugesia*. The results showed that *D. sicula* and *D. aethiopica* form a monophyletic clade that is the sister group of all other *Dugesia* species used in this study. Therefore, these two species were used to root all subsequent analyses.

2.6. Molecular clock calibration

In the absence of planarian fossils, only paleogeographic events of known age can be used to calibrate a molecular clock. However, in the case of planarians this is also not straightforward. Fortunately, the complex and well-known geological history of the eastern Mediterranean enables one to find such calibration points. In particular, the well-supported split in our phylogenetic trees between Cretan species and all other Greek species suggested this node as the best point to calibrate the phylogenetic tree. In order to assign a divergence date to this calibration point we considered two alternative scenarios, corresponding to the two times in its history that Crete became isolated. The first isolation of Crete took place c. 11–9 Mya (Dermitzakis, 1990), when it was separated from the mainland Ägäis. During the MSC, between 5.96 and 5.33 Mya, the Mediterranean dried out because of the closure of the Strait of Gibraltar (Hsü, 1972), reconnecting Crete to the mainland. Subsequently, the second Cretan isolation event occurred after the MSC, when the Mediterranean reflooded. In order to test, by using Bayes Factors (BFs), which of the two datings better explained our data we compared three temporal scenarios using a second calibration point, since a single calibration point does not provide a powerful test. As second calibration point we used the separation between eastern and western regions in the Aegean Sea (in our case corresponding to the Aegean islands, east, versus the rest, west), albeit that the clusters presumably correlated with that event have low support in our trees. For this splitting, we also considered two other possible datings. The first calibration point is the opening of the MAT (c. 12–9 Mya), as used in other studies (cf. Lymberakis and Poulakakis, 2010 and references therein). However, given the topology of the tree obtained, a scenario with Cretan lineages diverging at the end of the Messinian salinity crisis (5.3 Mya) and the east–west split occurring between 12 and 9 Mya was impossible because the east–west split occurs after

Table 1
Sampling localities of *Dugesia* populations used in this study (see also Fig. 1).

Locality code	Species	Sampling site	Coordinates	GenBank accession no.	
				COI	ITS-1
	Outgroups				
	D. aethiopica	Lake Tana, Ethiopia	Lázaro et al. (2009)	FJ646932 + FJ646976	FJ646889
	D. benazzii	R. Lernu, Sardinia, Italy	Lázaro et al. (2009)	FJ646933 + FJ646977	FJ646890
	D. etrusca	Tuscany, Italy	Lázaro et al. (2009)	FJ646939 + FJ646984	FJ646898
	D. gonocephala	Vijlen, Limburg, Netherlands	Lázaro et al. (2009)	FJ646941 + FJ646986	FJ646900
	D. hepta	R. S. Lucia, Sardinia, Italy	Lázaro et al. (2009)	FJ646943 + FJ646988	FJ646902
	D. ilvana	I. Elba, Tuscany, Italy	Lázaro et al. (2009)	FJ646944 + FJ646989	FJ646903
	D. sicula	S. Antioco, Sardinia, Italy	Lázaro et al. (2009)	FJ646947 + FJ646994	U84356
	Dugesia sp.	Vernár, Slovak Republic	48°55′21.06″N 20°18′34.45″E	KC007033	KC007104
				KC007017	KC007110
		Ludrová, Slovak Republic	49°1′46.18″N 19°19′49.07″E	KC007013	KC007114
		Prosiek, Slovak Republic	49°9′15.18″N 19°29′53.64″E	KC007030	KC007113
	Ingroups				
1	*D. cretica*	Georgioupoli, Crete, Greece	35°21′37.94″N 24°15′6.51″E	JN376141	KC007051
				KC006976	KC007050
2		Kakopetros, Crete, Greece	35°24′29.34″N 23°45′19.23″E	KC006974	KC007054
				KC006973	KC007053
3		Sasalos, Crete, Greece	35°24′9.86″N 23°42′42.39″E	KC006975	KC007052
				KC006977	KC007055
4	*Dugesia* sp.	Rouvas Gorge, Crete, Greece	35°9′48.66″N 24°54′34.71″E	KC007032	KC007102
				KC007012	KC007091
5	*D. elegans*	Petaloudes Valley, Rhodes, Greece	36°20′13.51″N 28°3′44.90″E	KC006985	KC007062
				KC006984	KC007063
6	*D. ariadnae*	Apollonas, Naxos, Greece	37°9′53.96″N 25°32′42.94″E	JN376142	KC007048
				KC006972	KC007049
7	*D. improvisa*	Melanes, Naxos, Greece	37°5′3.38″N 25°26′59.40″E	KC006987	KC007065
				KC006986	KC007064
8	*D. damoae*	Manolates, Samos, Greece	37°47′21.26″N 26°49′17.80″E	KC006979	KC007057
				KC006978	KC007056
9	*D. effusa*	Nagos, Chios, Greece	38°33′27.73″N 26°4′28.26″E	KC006983	KC007058
				KC006981	KC007061
10		Nagos, before the opening to the sea, Chios, Greece	38°33′34.73″N 26°4′56.86″E	KC006980	KC007060
				KC006982	KC007059
11	*Dugesia* sp.	Prokopi, Euboea, Greece	38°49′45.72″N 23°16′53.48″E	KC007026	KC007112
				KC007010	KC007089
12	*Dugesia* sp.	Eleonas – Gravia, Phocis, Greece	38°34′29.21″N 22°23′38.50″E	KC007018	KC007090
				KC007014	KC007101
13	*Dugesia* sp.	Varia, Aetolia-Acarnania, Greece	38°35′34.87″N 21°35′11.02″E	KC007011	KC007108
				KC007020	KC007092
14	*Dugesia* sp.	Vafkeri, Lefkada, Greece	38°43′31.41″N 20°39′46.59″E	KC007034	KC007088
				KC007009	KC007093
15	*Dugesia* sp.	Pogradec, Albania	~40°53′44.05″N 20°37′52.32″E	FJ646970 + FJ647015	FJ646930
16	*Dugesia* sp.	Filiates, Thesprotia, Greece	39°38′16.09″N 20°23′41.48″E	KC007028	KC007103
				KC007035	KC007107
17	*Dugesia* sp.	Potamia, Preveza, Greece	39°22′37.42″N 20°52′38.41″E	KC007037	KC007109
				KC007036	KC007105
18	*D. malickyi*	Gorgopotamos, Phthiotis, Greece	38°49′46.06″N 22°22′53.37″E	KC006990	KC007069
				KC006991	KC007066
19		Mexiates, Phthiotis, Greece	38°53′4.09″N 22°18′53.16″E	KC006988	KC007068
				KC006989	KC007067
20	*Dugesia* sp.	Polidrosos, Phocis, Greece	38°38′4.43″N 22°30′49.69″E	KC007022	KC007115
				KC007023	KC007094
21	*Dugesia* sp.	Tripi, Peloponnese, Greece	37°5′38.47″N 22°20′46.29″E	KC007025	KC007100
				KC007021	KC007106
22	*Dugesia* sp.	Agios Floros, Peloponnese, Greece	37°10′8.94″N 22°1′33.92″E	KC007029	KC007086
				KC007008	KC007087
23	*Dugesia* sp.	Dorio – Psari, Peloponnese, Greece	37°18′29.61″N 21°51′55.96″E	KC007024	KC007111
				KC007019	KC007099
24	*Dugesia* sp.	Theisoa – Andritsaina, Peloponnese, Greece	37°29′13.97″N 21°55′4.88″E	KC007031	KC007096
				KC007015	KC007098
25	*D. arcadia*	Chalandritsa, Peloponnese, Greece	38°6′31.85″N 21°47′13.73″E	KC006969	KC007044
				KC006971	KC007047

Table 1 (*continued*)

Locality code	Species	Sampling site	Coordinates	GenBank accession no.	
				COI	ITS-1
26		Sella, Peloponnese, Greece	38°17′3.02″N 21°52′45.80″E	JN376140 KC006970	KC007045 KC007046
27	*D. sagitta*	Roda, Corfu, Greece	39°47′23.94″N 19°47′29.46″E	KC007006 KC007003	KC007077 KC007074
28		Sfakera, Corfu, Greece	39°46′54.55″N 19°47′16.86″E	KC007002 KC006997	KC007081 KC007082
29		Kato vrisi spring, Klimatia, Corfu, Greece	39°44′30.48″N 19°46′49.20″E	KC007004 KC006996	KC007080 KC007075
30		Ano vrisi spring, Klimatia, Corfu, Greece	39°44′12.16″N 19°47′6.33″E	KC006999 KC007007	KC007083 KC007085
31	*D. parasagitta*	Ermones, Corfu, Greece	39°36′37.98″N 19°46′41.64″E	KC006995 KC006994	KC007072 KC007070
32		Ermones, slightly higher than 31, Corfu, Greece	39°36′41.93″N 19°47′1.40″E	KC006993 KC006992	KC007073 KC007071
33	*D. sagitta*	North of Vouniatades, Corfu, Greece	39°31′16.33″N 19°52′38.12″E	KC007000 KC007001	KC007076 KC007079
34		Benitses, Corfu, Greece	39°32′44.39″N 19°54′35.35″E	KC007005 KC006998	KC007078 KC007084
35	*D. aenigma*	Near Agia Eirini, Cephalonia, Greece	38°7′34.92″N 20°44′31.62″E	KC006968 KC006963	KC007040 KC007038
36		Digaleto, Cephalonia, Greece	38°10′46.99″N 20°40′46.80″E	KC006966 KC006967	KC007039 KC007042
37		Near Agia Eirini, Cephalonia, Greece	38°7′35.58″N 20°44′34.80″E	KC006965 KC006964	KC007043 KC007041
38	*Dugesia* sp.	Pastra, Cephalonia, Greece	38°6′4.38″N 20°45′4.14″E	KC007016 KC007027	KC007097 KC007095

Table 2
Forward (F) and reverse (R) primers used in amplification and sequencing.

Name	Sequence 5′–3′	Annealing temperature (°C)	Source
ITS-1			
9F (F)	GTAGGTGAACCTGCGGAAGG	45	Baguñà et al. (1999)
ITSR (R)	TGCGTTCAAATTGTCAATGATC	45	Baguñà et al. (1999)
COI			
BarT (F)	ATGACDGCSCATGGTTTAATAATGAT	43	Álvarez-Presas et al. (2011)
COIEF3 (F)	CCWCGTGCWAATAATTTRAG	48	This study
COIR (R)	CCWGTYARMCCHCCWAYAGTAAA	43	Lázaro et al. (2009)

the separation of Cretan species. Therefore, we also considered a second possible dating, i.e. that the final separation between the Aegean islands lineages and the rest of planarian lineages was a result of the refilling of the Mediterranean after the MSC. In this way, we generated three different temporal scenarios: (1) isolation of Crete at 10 Mya and east–west splitting during MAT opening between 12 and 9 Mya (scenario D1); (2) isolation of Crete at 10 Mya and east–west splitting at the end of MSC 5.3 Mya (scenario D2); (3) isolation of Crete and east–west splitting both occurring at 5.3 Mya (scenario D3). Once the best dating for the Crete separation was evaluated, we inferred a new dating tree using only that calibration point (D4), thus avoiding to fix a date for the separation between eastern and western lineages and to allow that dating to be deduced from the data itself.

We ran BEAST 1.6.1 software package (Drummond and Rambaut, 2007) in order to estimate clade divergence times for the fragment of COI, using relaxed molecular clock settings, following the uncorrelated relaxed lognormal clock. We applied the Yule or 'pure birth' prior process to the speciation model. The model of sequence evolution used was GTR + I + Γ, with runs of 12 million steps, sampling a tree every 1200 steps. Tracer vers. 1.5 (Rambaut and Drummond, 2007) was used to check convergence of parameters and to obtain mean and standard deviation (SD) of the substitution rates. We discarded 10% of the steps as burn-in. We assumed an age of 10 ± 1 Mya (mean of the normal prior distribution ± SD after relative 95% confidence intervals) for the first isolation of Crete, an age of 5.3 ± 0.3 Mya for the end of the MSC (both as the second isolation of Crete and for the splitting between east and west) and an age of 12–9 Mya for the opening of the MAT. Once we had the three calibrated trees, we applied a Bayesian model selection approach to decide which of the three temporal scenarios best fitted the data by running BF with Tracer and evaluating the results following Kass and Raftery (1995) criteria.

2.7. Biogeographic analyses

We used S-DIVA (Statistical Dispersal-Vicariance Analysis) implemented in RASP (Yu et al., 2010) in order to infer the biogeographic history of the Greek *Dugesia* lineages. This method facilitates statistical reconstruction of the ancestral distribution

of species, taking into account phylogenetic uncertainty. We pruned the species tree, leaving one specimen per sampling locality, and excluding all species not present in the northeastern Mediterranean, with the exception of *Dugesia hepta*, which we used as outgroup in this analysis. S-DIVA was run using the trees sampled in a BEAST 1.6.1 analysis for COI and ITS-1. This analysis was performed with 50 million steps, sampling a tree every 10,000 steps. The condensed tree was obtained from the BEAST analysis using the TreeAnnotator 1.6.1 program with a 10% burn-in. We defined eight areas for biogeographic analysis: (A) Sardinia (outgroup); (B) Crete; (C) eastern Aegean islands; (D) Naxos; (E) Peloponnese; (F) Euboea; (G) Mainland; (H) Corfu; (I) Cephalonia. Although some of these areas could potentially be further divided, as mainland for example, we did not do it in order to avoid an excess of divisions.

3. Results

3.1. Taxonomic status

The taxonomic status of the animals from the various populations was determined through morphological analysis of histological sections. In this way we were able to assign the populations to eight out of the nine species known for Greece, viz. *Dugesia aenigma* de Vries, 1984, *D. arcadia* de Vries, 1988, *D. ariadnae* de Vries, 1984, *D. cretica* (Meixner, 1928), *D. damoae* de Vries, 1984, *D. elegans* de Vries, 1984, *D. malickyi* de Vries, 1984, and *D. sagitta* (Schmidt, 1861). Further, three new species were identified through both morphological and molecular markers; the new species names (*D. effusa*, *D. improvisa* and *D. parasagitta*) currently should be treated as *nomina nuda*. In addition, the molecular analysis suggests the presence of a few other genetic lineages, potentially new species (4, 11–17, 20–24 and 38). Unfortunately, we have been unable to ascertain the taxonomic status of these populations due to lack of (1) fixed material, (2) sexual specimens or (3) adequate histological sections. In a companion paper (Sluys et al., in preparation) we will examine and discuss more at length the species status of all populations examined in this study, in particular the new species, based on an integrative approach to taxonomy.

3.2. Sequence characteristics and divergence values

COI (706 bp) and ITS-1 (646 bp) sequences were analyzed for 74 new individuals from Greece and four from Slovakia. The saturation process plot shows that third codon position of the coding COI is not saturated. Therefore, final analyses included third codon positions. ITS-1 is also not saturated (Supplementary data Fig. 1).

Distance data between known species are given in Supplementary data Tables 1–4.

3.3. Phylogenetic and dating analyses

The concatenated tree (Fig. 2) and the tree derived from the COI gene (Supplementary data Fig. 2) have very similar topologies, although the first generally provides more resolved groups. In contrast, ITS-1 data only supports the split of the Cretan clade from the rest of the populations and also a few internal clades (Supplementary data Fig. 3). In particular, the Peloponnese clade is monophyletic and separated by a long branch from the rest, indicating the presence of a number of fixed substitutions in this group. Nevertheless, addition of ITS-1 data to those for the COI gene increases the resolution of the phylogenetic tree obtained from the latter gene. Although ML and BI trees inferred from the concatenated data set show some differences, the basic topology is the same for both.

Summarizing, all analyses reveal a clear correlation between the genetic lineages and their geographic distribution, albeit that ITS-1 provided a less resolved tree. The general picture emerging from these analyses shows a first divergence of the Cretan species, separated in all cases by a long branch from the remaining species, thus suggesting a relatively old event, as compared to the rest of splits. The next node corresponds to the separation of eastern and western MAT lineages. The eastern group, formed by populations from the eastern islands, includes two species (6 and 7) from Naxos (a priori a western island) and the only population studied from Euboea (11). The resolution within this group is poor, with the nodes receiving low support, likely indicating a radiation event that did not leave a clear signal in the molecules studied. In the western clade, the Peloponnese populations constitute a monophyletic group that is highly differentiated from the rest, presumably reflecting a relatively recent dispersal event. The mainland clade is only monophyletic in the dating analysis (Fig. 3); however, all analyses (Figs. 2 and 3, Supplementary data Fig. 2) show populations 15 (Ori Lake in Albania) and 16 (northwestern Greece) in a basal situation as well as two monophyletic clades constituted by the populations 18–20 and 12–14. These two monophyletic clades show a pattern of isolation by distance in the trees (geographically closer populations are more closely related in the trees; Figs. 2 and 3), a pattern to be expected in a case of dispersal, i.e. migration followed by genetic drift. Finally, mainland population 17 and all Ionian populations constitute a monophyletic group in all analyses, while within this group the populations from Corfu and from Cephalonia form two monophyletic clades.

Exceptions to these congruent results among methods concern the Euboea (11), Albanian (15), and Filiates (16) populations. In the COI tree (Supplementary data Fig. 2), the Euboea (11) population is positioned at the base of the western clade (including mainland and Peloponnese clades), whereas in the concatenated tree (Fig. 2) it is at the root of a monophyletic eastern clade (including eastern islands and Naxos species). However, BEAST Bayesian based tree (Fig. 3) positioned the Euboea (11) population within the eastern clade. Additional COI analyses without this Euboea population resulted in a COI tree (not shown) with a monophyletic eastern clade, but with similar node supports. Although the Albanian (15) and Filiates (16) populations always have a basal position with respect to the mainland clade, their relationships vary slightly and never receive high support, whereas the rest of mainland populations form two well-defined clusters.

The datings and substitution rates obtained with the three calibration scenarios (D1, D2, D3) are compared in Table 3, and the results of the BF comparison are presented in Table 4. The D2 model was best supported by the data, with the Bayes Factors providing substantial and strong support (Kass and Raftery, 1995 interpretation) for this scenario, as compared to scenarios D1 and D3, respectively. Scenario D1 received substantial support as compared to D3. Given the strong support for the calibration based on the separation of Crete at 10 Mya, and taking into account the low support for the two clusters defining our second calibration point, we performed a dating analysis using only the first calibration point at 10 Mya (Fig. 3; Table 3), in order to obtain an objective dating for the separation of the eastern and western lineages. This resulted in a value of 7.5 Mya, lying between the two values used in our previous calibration analyses (Table 3).

3.4. Biogeographic analyses

The topology obtained after pruning the tree was very similar to that obtained with the complete data set (Fig. 4, Supplementary data Table 5), with only some differences in nodes with low support

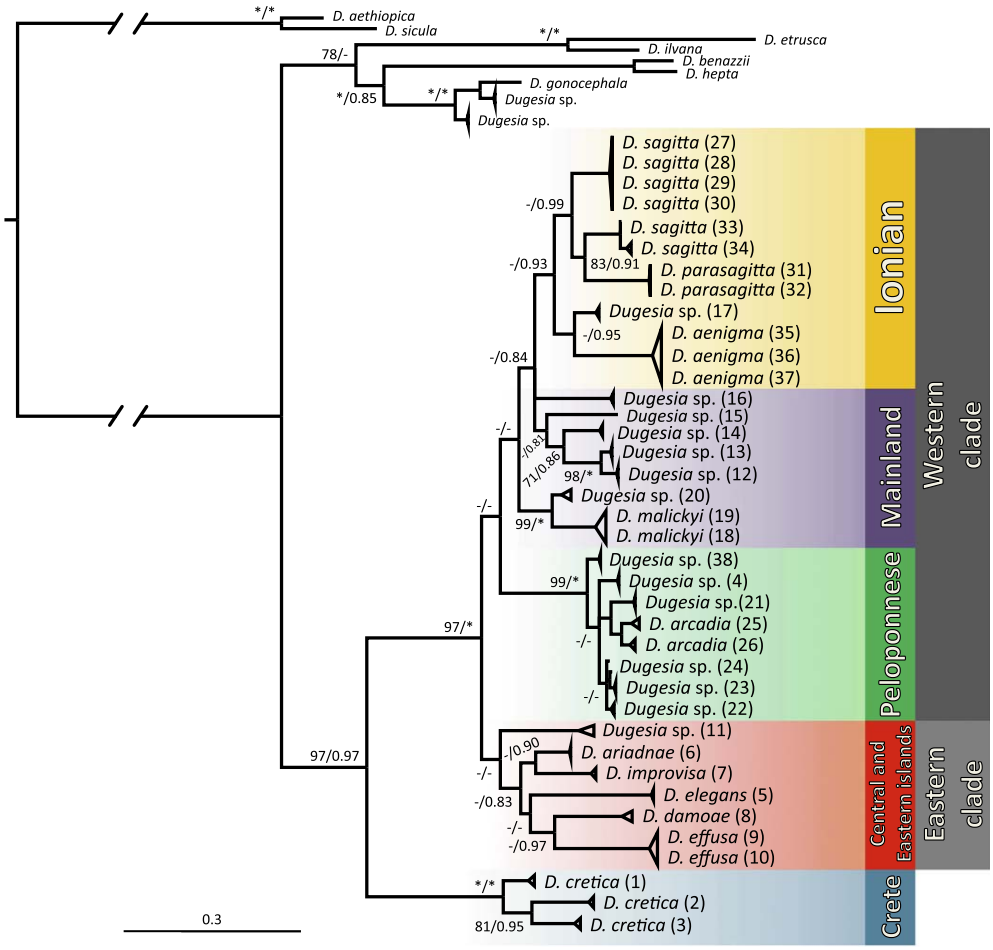


Fig. 2. Bayesian tree inferred from the concatenated data set (COI + ITS-1). Labels correspond to the species name (when known), and the numbers in parentheses refer to the locality codes reported in Table 1. Node numbers correspond to bootstrap (ML)/posterior probability (BI); values are only indicated when >50/>0.80, "*" indicates maximum support. The scale bar indicates substitutions per site.

in our phylogenetic and dating analyses. The specimens from Crete (4) and Cephalonia (38) that clustered together with the Peloponnese individuals were not included in the analysis because there is strong evidence that their distribution has been influenced by human transport. The results suggest seven vicariant events and two dispersals (indicated in green and blue, respectively, in Fig. 4). When we compare the dating studies and the biogeographic analyses (Figs. 3 and 4) it is clear that most of the results are in accordance with the geological history of the region. However, the hypothesized vicariant processes do not always precisely coincide with the presumed geological events. For example, in node 5 a dispersal event is inferred going from the mainland (G) to the western coast (HI, including present Corfu and Cephalonia), followed by a vicariant event (node 6) that splits region H (Corfu) from GI (Cephalonia plus a mainland population) and, subsequently, another vicariant event (node 61) that splits G (the mainland population) from I (Cephalonia). However, those hypothesized vicariant events occurred at a period long before the two islands were actually separated from the continent, implying that they were not the result of that geological event but most probably resulted from the isolation of several drainage basins in the Ionian region. Another complex situation is found in the eastern region, with S-DIVA deducing two vicariant events and one dispersal event. The first vicariant event separated the Euboea population (11) from the rest (node 21), after which dispersal took place from the east (C) to Naxos (D), followed by a vicariant event (nodes 22 and 23). Finally, there are three vicariances coinciding with geological breakages: the separation of Crete from the mainland (node 1), the separation of eastern and western Aegean lineages (node 2), and the separation of Peloponnese populations from the mainland (node 3).

4. Discussion

4.1. Differentiation among genetic lineages, speciation and systematic implications

The phylogenetic trees and also the genetic distances (Supplementary data Table 1), show that the species described previously on morphological grounds coincide with well-defined genetic lineages. The COI distances between species vary between 2.8% and 9.6%. These values are slightly lower than those found between *Dugesia* species in the western Mediterranean (Lázaro et al., 2009), excepting *D. benazzii* and *D. hepta* (the latter two species presumably representing a case of recent speciation). This situation suggests a younger diversification process in the eastern region (as also seen in the dating analyses, Fig. 3), which corresponds well with the fact that when Crete became isolated for the first time (c. 10 Mya) and the MAT began to form (c. 12 Mya), the western part of the Mediterranean had practically reached its present configuration (Rosenbaum and Lister, 2004a,b; Schettino and Turco, 2006).

In general, levels of intraspecific divergence (Table 1) fall within the range found for other planarian species (Lázaro et al., 2009,

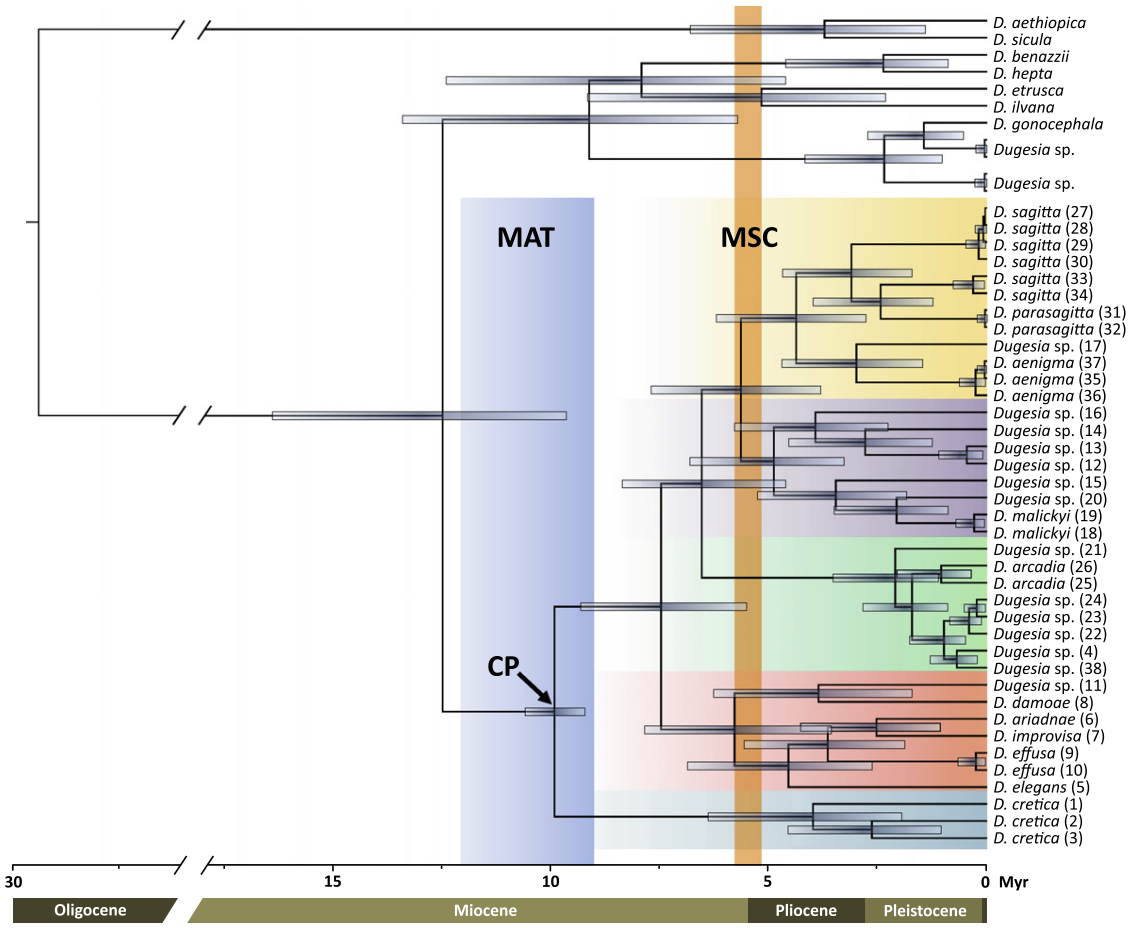


Fig. 3. Divergence times between Greek lineages of *Dugesia* inferred from COI by Bayesian analysis using a relaxed molecular clock and fixing the calibration point (CP) at 11–9 Mya (scenario D4). Bars at nodes represent the 95% highest posterior density (HPD) credibility interval. Vertical color bars indicate the periods of opening of the mid-Aegean trench (MAT; blue) and Messinian salinity crisis (MSC; orange).

Table 3
Inferred mean dates and highest posterior density (HPD) confidence interval for three scenarios using two calibration points: (1) the isolation of Crete from the Greek mainland (CP1: CR-GR) and (2) the split of the east and west Aegean (CP2: WMAT-EMAT); bottom row (D4) presents the calibration inferred from the data, using only the early isolation of Crete (c. 11–9 Mya) (for further explanation, see Material and Methods). Abbreviations: CR (Crete), EMAT (central and eastern islands), GR (All Greek populations without Crete), ION (Ionian), MNL (mainland), PEL (Peloponnese), and WMAT (Peloponnese, Mainland, and Ionian).

	Calibration point (CR-GR)	Node dating (MYA) [95% HPD]				Mean rate[b]
		WMAT-EMAT	PEL-(MNL + ION)	MNL-ION	Naxos-EMAT[a]	
D1	CP1: 11–9 MYA	Fixed	8.1	6.7	4.4	0.015
	CP2: 12–9 MYA		[9.5–6.2]	[8.5–4.8]	[6.5–2.3]	[0.001–0.021]
D2	CP1: 11–9 MYA	Fixed	4.9	4.3	2.7	0.022
	CP2: 5.5–5 MYA		[5.4–4.1]	[4.3–2.3]	[4.3–1.6]	[0.015–0.029]
D3	CP1: 5.5–5 MYA	Fixed	4.4	3.6	2.3	0.028
	CP2: 5.5–5 MYA		[5.2–3.3]	[4.7–2.6]	[3.5–1.2]	[0.002–0.041]
D4	11–9 MYA	7.5	6.6	5.7	3.7	0.017
	[Mean: 10; SD: 0.3]	[9.3–5.5]	[8.4–4.6]	[7.7–3.8]	[5.6–1.9]	[0.011–0.024]

[a] eastern islands excluding Naxos.
[b] Number of substitutions per site divided by tree length.

2011). The populations from Corfu (i.e. *D. sagitta* and *D. parasagitta*) are structured in three differentiated clades, corresponding with their geographical distribution (northern, central and southern part of the island; Figs. 1–4), with distances between the clades reaching 4.7% (Supplementary data Table 2), a value slightly higher than the maximum found between populations of the same species in the western Mediterranean (Lázaro et al., 2009). Despite the similar genetic distances between the three clades, only the central group presents morphological differences with respect to the other two, thus allowing the delimitation of a new species (*D. parasagitta*). For *D. cretica* the distances between populations also reach high values (5.7%), whereas study of their internal anatomy shows all of these animals to be identical.

On the other hand, there are some genetic lineages that do not correspond to any known species. Morphological analysis of some of these populations revealed cases in which defining characters exist for some genetic clades, whereas in other cases the opposite was the case, i.e. that morphological differences appear in genetically closely related populations. This complex situation calls for a deeper analysis, both from a morphological and a molecular point

Table 4
Bayes Factors results for the comparison of the three temporal scenarios. Probability of the three models with standard error and log10 Bayes factors.

Scenario	lnP (model|data)	S.E.	CP1: 11–9 MYA CP2: 12–9 MYA D1	CP1: 11–9 MYA CP2: 5.5–5 MYA D2	CP1: 5.5–5 MYA CP2: 5.5–5 MYA D3
CP1: 11–9 MYA CP2: 12–9 MYA	−5336.426	±0.224	–	−0.74	0.811
CP1: 11–9 MYA CP2: 5.5–5 MYA	−5334.721	±0.218	0.74	–	1.551
CP1: 5.5–5 MYA CP2: 5.5–5 MYA	−5338.292	±0.221	−0.811	−1.551	–

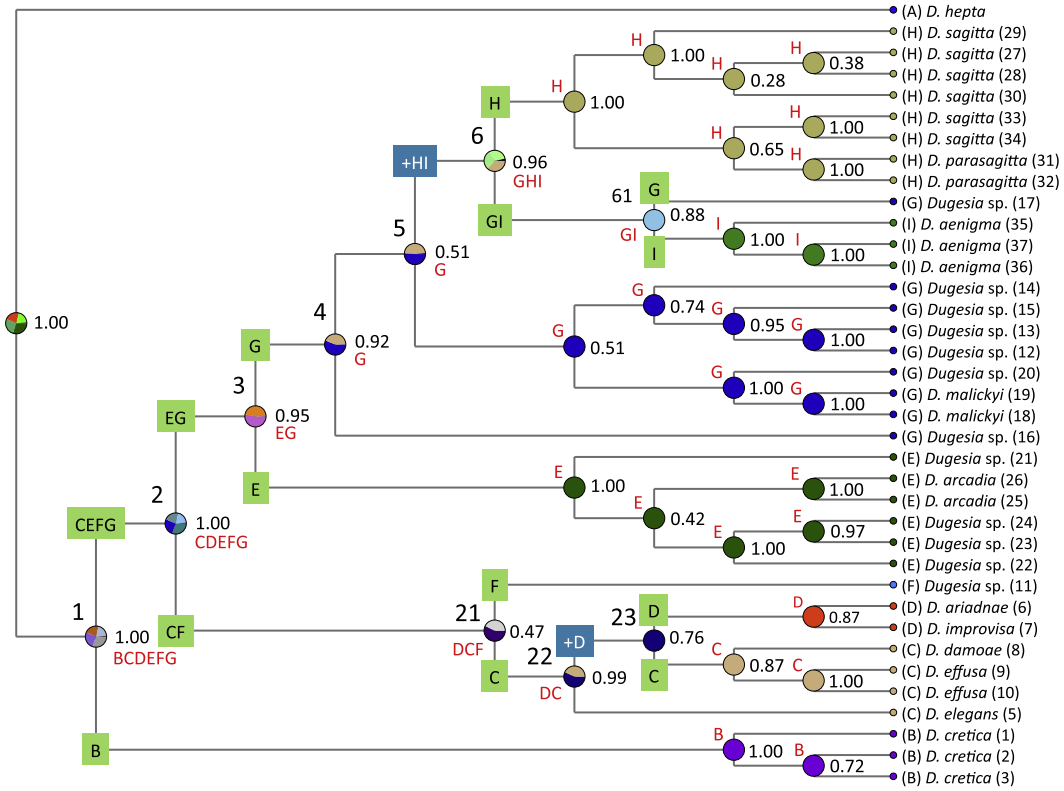


Fig. 4. Cladogram showing the results of the S-DIVA analysis. The node charts show the relative probabilities of alternative ancestral distribution ranges (see Supplementary data Table 5 for the exact values), red letters at the nodes indicate the area with highest probability. Vicariant and dispersal events inferred by the program are highlighted in green and blue, respectively. Numbers shown over some nodes are used to identify them in the text. Posterior probabilities of nodes are shown at their right side. The areas used in this analysis are: (A) Sardinia (outgroup); (B) Crete; (C) eastern Aegean islands; (D) Naxos; (E) Peloponnese; (F) Euboea; (G) Mainland; (H) Corfu; (I) Cephalonia.

of view. In a companion paper we will provide a more in-depth analysis of these cases by taxonomically integrating morphological and molecular data.

4.2. Phylogenetic congruence with geological and climatic history

Although the phylogenetic pattern is congruent with the geographical distribution of the lineages, it remains to be examined whether the timings for the splittings found in the dating analyses coincide with the geological and climatic history of the region. For this we have used the paleogeographic isolation of Crete from the continent as a calibration point for our divergence time analyses. This island became isolated twice, and there has been some contention on which dating is the most adequate to do this calibration. For many terrestrial animals it has been demonstrated that they used the exposed land surface to migrate from the continent to the island during the MSC (Lymberakis and Poulakakis, 2010 and references therein). Even seawater sensitive animals, such as amphibians and freshwater crabs, migrated during the Lago-Mare phase of the MSC (Akın et al., 2010; Jesse et al., 2011). However, unlike freshwater planarians these organisms are able to survive outside of freshwater and therefore their presumed dispersal still provides no firm evidence that the land bridge between Crete and the Peloponnese contained contiguous freshwater bodies facilitating dispersal of the planarians. Nevertheless, we considered the possibility that the Lago-Mare may have offered *Dugesia* an opportunity for dispersal, and hence we calibrated the splitting of the Cretan lineage at the two known moments that this island became isolated in order to compare both scenarios. The results show that our data give stronger support to the 11–9 Mya calibration (Table 4), pointing to a situation where planarians probably did not disperse between Crete and the mainland during the MSC. Either, there was no contact among freshwater bodies between Crete and the continent or planarians did not take the opportunity to disperse.

In the calibration based only on the earlier isolation of Crete at 11–9 Mya (Fig. 3; Table 3), the evolutionary rate (1.7% per site per

lineage per million years) is in agreement with that found in other groups of organisms (e.g. Papadopoulou et al., 2010; Allegrucci et al., 2011) and also with what is considered a universal rate for mitochondrial DNA (Brown et al., 1979). In this scenario, the divergence between western (including the Peloponnese, mainland, and Ionian clades) and eastern (central and western islands clade) Aegean populations occurred between the end of the MAT opening and the beginning of the MSC, thus leaving open the possibility that it was a late consequence of the opening of the MAT. The situation that after the end of the MAT, other climatic and geographic events probably resulted in renewed contact of eastern and western lineages at some point in time (see below) may explain the fact that the support for this splitting is not high, while its dating does not exactly fit the 12–9 Mya period. Furthermore, the divergence times in this scenario, together with the biogeographic analysis, suggest that the common ancestor of the two Naxos species (*D. ariadnae* and *D. improvisa*), which are closely related to the eastern lineages, colonized this island from the eastern Aegean region during the Lago-Mare phase of the MSC (5.5–5.33 Mya). In that case, the freshwater systems on Naxos must have been in contact at one moment with the western Aegean systems during the MSC, perhaps flowing into common freshwater or brackish lakes. Recent human introduction seems a less parsimonious alternative hypothesis, since the two sister species then must have speciated in the east and, subsequently, have been transported on two occasions to the island, given that their speciation is much older than human activity. Additionally, the radiation suggested by the lack of resolution found in this region may have resulted from the cessation of contact between landmasses, due to rapid reflooding of the Mediterranean after the MSC, resulting in vicariant speciation on islands during the latest Messinian (5.33 Mya), although this could not be evaluated in S-DIVA since it was defined as a single distribution area. In a similar way, the Messinian has been postulated as the time of diversification of Mediterranean cyprinids (Bianco, 1990). These freshwater fishes would have dispersed across the basin during the Lago-Mare stage and underwent a fast speciation as a consequence of the return of the basin to marine conditions. This may be reflected in the deep polytomies found in some molecular analyses (Durand et al., 2003; Ketmaier et al., 2004; Tsigenopoulos et al., 2003). Although not all authors agree that the Messinian would have resulted in diversification around the whole Mediterranean basin (Perea et al., 2010), it seems clear that the Lago-Mare stage has acted at a local scale, especially in the eastern Mediterranean (Durand et al., 1999; Ketmaier et al., 2004). The individuals from Euboea (11) are also part of the eastern group in most analyses, if this situation is confirmed, this population and the eastern islands species share a common ancestor. Furthermore, the S-DIVA analysis infers with a higher probability an ancestral area for this clade comprising the eastern islands (C), Naxos (D) and Euboea (F) or CF (node 21 in Fig. 4, Supplementary data Table 5). Taking into account the geological history of the Aegean, this implies that the ancestors of this clade shared the same area in the east, and that the occurrence of species of this clade in the middle Aegean and in the west is due to dispersal. In that case, the Euboea individuals may result from dispersal by some eastern populations, followed by a vicariant event (as estimated by S-DIVA). This situation is also congruent with the findings of Durand et al. (1999) for the cyprinids. In that study a population from Euboea was found to be closely related to species from the rivers in eastern Greece. Our finding reinforces their hypothesis that freshwater habitants in Euboea would have evolved from eastern populations, which arrived as a result of contact between freshwater bodies in the northern Aegean Sea during a decrease in salinity in interglacial seven at about 200,000 years ago (Bianco, 1990).

The Peloponnese clade splits in a vicariant event (node 3 in Fig. 4) from the rest of the continental species at an earlier period (8.5–4.6 Mya) than the geological isolation of the Peloponnese (4–3 Mya), the latter event proposed as an explanation for the evolution of endemic species and lineages on this peninsula (Ursenbacher et al., 2008; Jesse et al., 2011). Presumably, the split between the Peloponnese clade and the other mainland and Ionian lineages was due to the severance of freshwater drainages before the peninsula was actually formed. In fact, the long branch separating this clade from all other groups, both in the COI and ITS-1 trees, and the low variability within it, point to the occurrence of a bottleneck event within the Peloponnese lineage (genetic drift during the bottleneck would have fixed mutations in the DNA that otherwise could have become lost). Although our data do not allow us to statistically test this demographic event, the dating tree shows that *Dugesia* populations on the Peloponnese did not diversify in a period between c. 7 and c. 2 Mya. This last point in time coincides with the beginning of the Pleistocene, a period characterized by an increase in humidity in the Mediterranean area (Haywood et al., 2000), which may have promoted the diversification of *Dugesia* on the Peloponnese peninsula through colonization of newly established freshwater environments.

On the mainland, the dispersal of freshwater planarians seems to have followed a north to south direction along both sides of the Pindus mountain range (east and west, Fig. 3 and Supplementary data Fig. 4). Again, although our data does not allow statistical testing of this dispersal hypothesis, while it could neither be seen in the biogeographic analyses (since we defined all mainland as a single distribution area to avoid an excess of regions), this result is congruent with that for freshwater fishes (Durand et al., 1999). The planarian dispersals are dated at the end of the Pliocene or during the Pleistocene, a little before the datings proposed for the fishes (middle and end of the Pleistocene for western and eastern Pindus lineages, respectively). However, the fishes' datings were based on a rate calculated for other organisms and were used very cautiously by Durand et al. (1999).

These two latter hypotheses, i.e. (1) loss of diversity on the Peloponnese with a recent recovery and (2) a possible north–south recolonization pattern on the mainland, need more detailed population studies in order to test the occurrence of bottleneck or dispersal events. However, the currently available dated tree induced us to erect a bold hypothesis, viz. a freshwater crisis before or during the MSC on the southern part of the Balkan Peninsula that resulted in the disappearance of most planarian populations in that region. When the climate became more suitable, surviving *Dugesia* populations on the Peloponnese would have dispersed through the entire Peloponnese peninsula, resulting in the radiation that is apparent from the phylogenetic trees. Moreover, populations situated in the north of the Balkan Peninsula could move southwards, colonizing Greece along both sides of the Pindus mountain range (Supplementary data Fig. 4).

Finally, the Ionian clade comprises *D. aenigma* from Cephalonia, *D. sagitta* and *D. parasagitta* from Corfu, and population 17 from the mainland in a basal position. This group reflects the presence of a third mainland lineage that, after the biogeographic analysis, presumably dispersed from the mainland to the west coast (node 5 in Fig. 4). Later, it experienced at least two vicariant events: (1) the splitting of Corfu lineages from the rest and (2) the splitting of population 17 from Cephalonia lineages (Fig. 3, Fig. 4). These two vicariant events (at about 4 and 1.7 Mya, respectively; Fig. 3) as well as the diversification of the three Corfu lineages (at about 1.7 Mya, Fig. 3) predate the isolation of Corfu and Cephalonia from the mainland at c. 9000 cal. yr BP. This suggests that during the last glaciation the river drainage basins inhabited by these lineages were not in contact, although Corfu was joined with the mainland through a large coastal plain traversed by many rivers (Van Andel and Shackleton, 1982). Hence, the vicariant events estimated are most probably due to the severance of those drainages.

Interestingly, the rate of substitution obtained for *Dugesia* in this study differs considerably from the only other molecular calibration for triclads available from the literature: *Schmidtea mediterranea* (Dugesiidae), with a 0.27% substitution per lineage and Mya for COI (Lázaro et al., 2011). The different rates of diversification observed for *Dugesia* and *Schmidtea* (the latter genus with only four species and a restricted area of distribution) may also explain the observed differences in molecular substitution rates.

4.3. Impact of human activities on planarian distribution

An unexpected result is that population four from Crete and population 38 from Cephalonia fall within the Peloponnese clade. According to the ingroup genetic distance between the Crete and Cephalonia populations and the Peloponnese specimens (COI: 0.8–2.5%; ITS-1: 0–1.1%) and the rather recent divergence times (0.96 Mya), postdating the last contact between these landmasses, it does not seem likely that they spread by their own means from the Peloponnese to these two islands. It has been suggested that biochore dispersal is of no importance in the dispersal of freshwater triclads (Reynoldson, 1966). However, all evidence in this case points to humans as a vector of their dispersal, a possibility already mentioned by de Vries (1985). The genetic similarity between Cretan and Cephalonian populations suggests that they originated from the same source population or from two genetically and geographically close populations. Despite this case, the total effect of anthropochorous transport on the current distribution of the planarians seems to be limited in this section of the Mediterranean region.

4.4. Fluvial basin history underlies planarian dispersal and speciation patterns

Fluvial basins may act as ''ecological islands'' for exclusively freshwater organisms, even on islands, in the same way as mountain peaks, landslides, or puddles (Heads, 2011). This is also evident in Mediterranean *Dugesia*. For example, the three genetic lineages found in Corfu diverged on the continent in the absence of marine barriers. This illustrates how the lack of contact between freshwater drainages for these organisms is as important in their diversification processes as island formation is for many other organisms. The Peloponnese lineage also represents a case of divergence before a well-known geographical barrier appears, viz. the opening of the Gulf of Corinth. This extreme dependence of planarians on contiguous freshwater bodies for their dispersal makes them an ideal group of organisms (1) to examine the effect of the geological history of freshwater drainages on their evolutionary diversification, and (2) to elucidate geological events, such as presence of land bridges and fluvial basins, which at times may be difficult to ascertain from geological data. In some respects this also applies to freshwater fishes, but planarians have the added advantage that they are able to live in smaller watercourses or even in temporary ones, thus enabling the extension of such historical inferences to a more fine-grained geographic scale. For example, the present study suggests that planarians were able to disperse between eastern and central islands in the Aegean Sea during the MSC, whereas there was likely no full contact between freshwater courses on the continent and Crete during that period.

Acknowledgments

This research was supported by BES-2009-022530 grant from the Ministerio de Ciencia e innovación (to E.S.), and by CGL 2008-00378 Grant to M.R. Completion of the study was made possible by a Grant from the Naturalis Biodiversity Center to R. Sluys. We are indebted to the following persons, who provided material and helped us with the samplings: Eduardo Mateos (also author of the *Dugesia* picture in the Graphical Abstract), Núria Bonada, Joan Solà, Dora Vázquez, Caterina Rodríguez, Miquel Vila-Farré, Margarita Metallinou and Enric Planas. We also thank two anonymous reviewers for their helpful comments and suggestions on the manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2012.11.010.

References

Akın, Ç., Can Bilgin, C., Beerli, P., Westaway, R., Ohst, T., Litvinchuk, S.N., Uzzell, T., Bilgin, M., Hotz, H., Guex, G., et al., 2010. Phylogeographic patterns of genetic diversity in eastern Mediterranean water frogs were determined by geological processes and climate change in the late cenozoic. J. Biogeogr. 37, 2111–2124.

Allegrucci, G., Trucchi, E., Sbordoni, V., 2011. Tempo and mode of species diversification in *Dolichopoda* cave crickets (Orthoptera, Rhaphidophoridae). Mol. Phylogenet. Evol. 60, 108–121.

Álvarez-Presas, M., Baguñà, J., Riutort, M., 2008. Molecular phylogeny of land and freshwater planarians (Tricladida, Platyhelminthes): from freshwater to land and back. Mol. Phylogenet. Evol. 47, 555–568.

Álvarez-Presas, M., Carbayo, F., Rozas, J., Riutort, M., 2011. Land planarians (Platyhelminthes) as a model organism for fine-scale phylogeographic studies: understanding patterns of biodiversity in the Brazilian Atlantic Forest hotspot. J. Evol. Biol. 24, 887–896.

Baguñà, J., Carranza, S., Pala, M., Ribera, C., Giribet, G., Arnedo, M., Ribas, M., Riutort, M., 1999. From morphology and karyology to molecules. New methods for taxonomical identification of asexual populations of freshwater planarians. Ital. J. Zool. 66, 207–214.

Ball, I.R., Fernando, C.H., 1969. Freshwater triclads (Platyhelminthes, Turbellaria) and continental drift. Nature 221, 1143–1144.

Ball, I.R., 1975. Nature and formulation of biogeographical hypotheses. Syst. Biol. 24, 407–430.

Bianco, P., 1990. Potential role of the paleohistory of the Mediterranean and Paratethys basins on the early dispersal of Euro-Mediterranean freshwater fishes. Ichthyol. Explor. Freshwaters 1, 167–184.

Bittkau, C., Comes, H.P., 2005. Evolutionary processes in a continental island system: molecular phylogeography of the Aegean *Nigella arvensis* alliance (Ranunculaceae) inferred from chloroplast DNA. Mol. Ecol. 14, 4065–4083.

Brown, W.M., George, M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA 76, 1967–1971.

Cosentino, D., Gliozzi, E., Pipponzi, G., 2007. The late Messinian lago-mare episode in the Mediterranean basin: preliminary report on the occurrence of Paratethyan ostracod fauna from central Crete (Greece). Geobios 40, 339–349.

de Vries, E.J., 1985. The biogeography of the genus *Dugesia* (Turbellaria, Tricladida, Paludicola) in the Mediterranean region. J. Biogeogr. 12, 509–518.

de Vries, E.J., 1984. On the species of the *Dugesia gonocephala* group (Platyhelminthes, Turbellaria, Tricladida) from Greece. Bijdr. Dierk. 54, 101–126.

de Vries, E.J., 1988. Further contributions to the taxonomy and biogeography of the subgenus *Dugesia* (Platyhelminthes: Tricladida: Paludicola) in the Mediterranean region and the Middle East. Isr. J. Zool. 35, 109–136.

Dermitzakis, M., 1990. Paleogeography, geodynamic processes and event stratigraphy during the late cenozoic of the Aegean area. International symposium on: biogeographical aspects of insularity, Roma 1987. Accad. Naz. Lincei 85, 263–288.

Dermitzakis, M., Papanikolaou, D.J., 1981. Paleogeography and geodynamics of the Aegean region during the neogene. Ann. Geol. Pays Hellen. 20, 245–289.

Douris, V., Giokas, S., Thomaz, D., Lecanidou, R., Rodakis, G.C., 2007. Inference of evolutionary patterns of the land snail *Albinaria* in the Aegean archipelago: is vicariance enough? Mol. Phylogenet. Evol. 44, 1224–1236.

Drummond, A., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Durand, J.D., Bianco, P.G., Laroche, J., Gilles, A., 2003. Insight into the origin of endemic Mediterranean Ichthyofauna: phylogeography of Chondrostoma genus (Teleostei, Cyprinidae). J. Hered. 94, 315–328.

Durand, J.D., Templeton, A.R., Guinand, B., Imsiridou, A., Bouvet, Y., 1999. Nested clade and phylogeographic analyses of the chub, *Leuciscus cephalus* (Teleostei, Cyprinidae), in Greece: implications for Balkan Peninsula biogeography. Mol. Phylogenet. Evol. 13, 566–580.

Hall, T.A., 1999. Bioedit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Haywood, A.M., Sellwood, B.W., Valdes, P.J., 2000. Regional warming: pliocene (3 Mya) paleoclimate of Europe and the Mediterranean. Geology 28, 1063–1066.

Heads, M., 2011. Old taxa on young islands: a critique of the use of island age to date island-endemic clades and calibrate phylogenies. Syst. Biol. 60, 204–218.

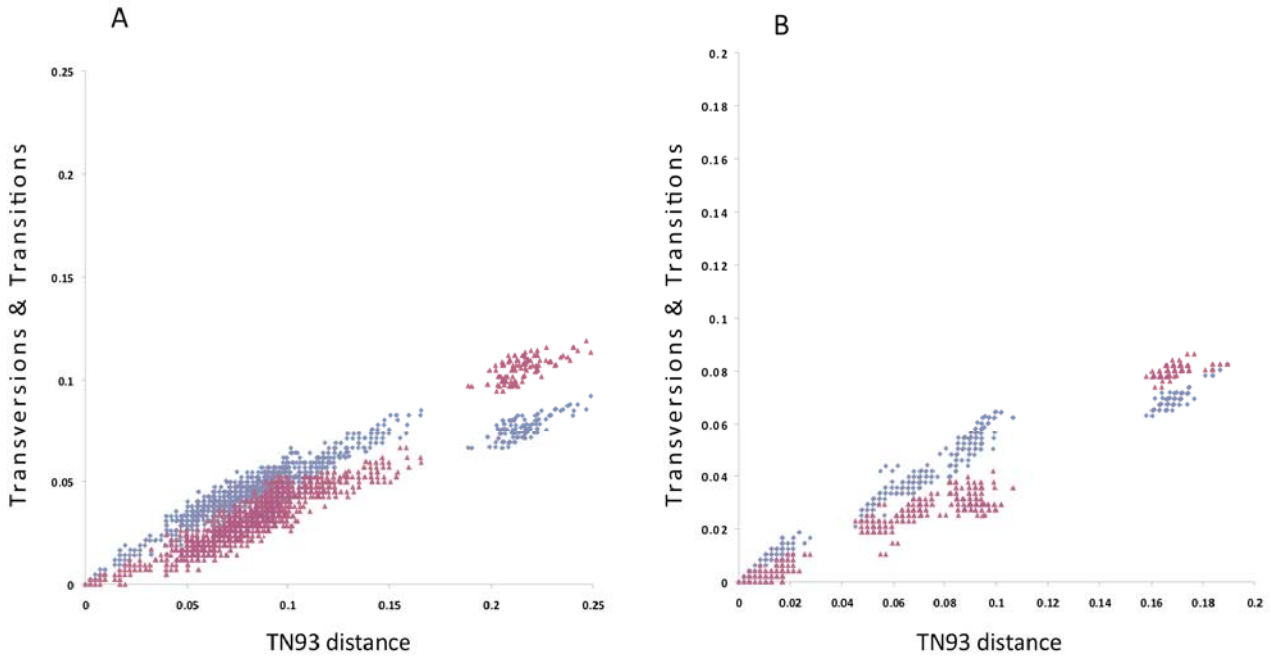
Hsü, K.J., 1972. Origin of saline giants: a critical review after the discovery of the Mediterranean evaporite. Earth-Sci. Rev. 8, 371–396.
Humason, G., 1967. Animal Tissue Techniques. W. H. Freeman and Company, San Francisco and London.
Jesse, R., Grudinski, M., Klaus, S., Streit, B., Pfenninger, M., 2011. Evolution of freshwater crab diversity in the Aegean region (Crustacea: Brachyura: Potamidae). Mol. Phylogenet. Evol. 59, 23–33.
de Jong, H., 1998. In search of historical biogeographic patterns in the western Mediterranean terrestrial fauna. Biol. J. Linn. Soc. 65, 99–164.
Kasapidis, P., Magoulas, A., Mylonas, M., Zouros, E., 2005. The phylogeography of the gecko *Cyrtopodion kotschyi* (Reptilia: Gekkonidae) in the Aegean archipelago. Mol. Phylogenet. Evol. 35, 612–623.
Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 773–795.
Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. 9, 286–298.
Ketmaier, V., Biancob, P.G., Cobollia, M., Krivokapicc, M., Canigliaa, R., De Matthaeisa, E., 2004. Molecular phylogeny of two lineages of Leuciscinae cyprinids (*Telestes* and *Scardinius*) from the peri-Mediterranean area based on cytochrome *b* data. Mol. Phylogenet. Evol. 32, 1061–1071.
Kornilios, P., Poulakakis, N., Mylonas, M., Vardinoyannis, K., 2009. The phylogeny and biogeography of the genus *Zonites* Montfort, 1810 (Gastropoda: Pulmonata): preliminary evidence from mitochondrial data. J. Mollusc. Stud. 75, 109–117.
Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the Messinian salinity crisis. Nature 400, 652–655.
Krijgsman, W., 2002. The Mediterranean: mare nostrum of earth sciences. Earth Planet. Sci. Lett. 205, 1–12.
Lázaro, E.M., Harrath, A.H., Stocchino, G.A., Pala, M., Baguñà, J., Riutort, M., 2011. *Schmidtea mediterranea* phylogeography: an old species surviving on a few Mediterranean islands? BMC Evol. Biol. 11, 274.
Lázaro, E.M., Sluys, R., Pala, M., Stocchino, G.A., Baguñà, J., Riutort, M., 2009. Molecular barcoding and phylogeography of sexual and asexual freshwater planarians of the genus *Dugesia* in the western Mediterranean (Platyhelminthes, Tricladida, Dugesiidae). Mol. Phylogenet. Evol. 52, 835–845.
Lymberakis, P., Poulakakis, N., 2010. Three continents claiming an archipelago: the evolution of Aegean's herpetofaunal diversity. Diversity 2, 233–255.
Maurakis, E.G., Pritchard, M.K., Economidis, P.S., 2001. Historical relationships of mainland river drainages in Greece. BIOS 6, 109–124.
Meixner, J., 1928. Der Genitalapparat der Tricladen und seine Beziehungen zu ihrer allgemeinen Morphologie, Phylogenie, Ökologie und Verbreitung. Z. Morph. Ökol. Tiere. 11, 570–612.
Papadopoulou, A., Anastasiou, I., Keskin, B., Vogler, A.P., 2009. Comparative phylogeography of tenebrionid beetles in the Aegean archipelago: the effect of dispersal ability and habitat preference. Mol. Ecol. 18, 2503–2517.
Papadopoulou, A., Anastasiou, I., Vogler, A.P., 2010. Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Mol. Biol. Evol. 27, 1659–1672.
Parmarkelis, A., Pfenninger, M., Spanos, L., Papagiannakis, G., Louis, C., Mylonas, M., 2005. Inference of a radiation in *Mastus* (Gastropoda, Pulmonata, Pulmonata, Enidae) on the island of Crete. Evolution 59, 991–1005.
Parmakelis, A., Stathi, I., Chatzaki, M., Simaiakis, S., Spanos, L., Louis, C., Mylonas, M., 2006. Evolution of *Mesobuthus gibbosus* (Brullé, 1832) (Scorpiones: Buthidae) in the northeastern Mediterranean region. Mol. Ecol. 15, 2883–2894.
Perea, S., Bohme, M., Zupancic, P., Freyhof, J., Sanda, R., Ozulug, M., Abdoli, A., Doadrio, I., 2010. Phylogenetic relationships and biogeographical patterns in circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. BMC Evol. Biol. 10, 265.
Perissoratis, C., Conispoliatis, N., 2003. The impacts of sea-level changes during latest pleistocene and holocene times on the morphology of the Ionian and Aegean seas (SE alpine Europe). Mar. Geol. 196, 145–156.
Poulakakis, N., Lymberakis, P., Antoniou, A., Chalkia, D., Zouros, E., Mylonas, M., Valakos, E., 2003. Molecular phylogeny and biogeography of the wall-lizard *Podarcis erhardii* (Squamata: Lacertidae). Mol. Phylogenet. Evol. 28, 38–46.
Poulakakis, N., Lymberakis, P., Valakos, E., Zouros, E., Mylonas, M., 2005. Phylogenetic relationships and biogeography of Podarcis species from the Balkan Peninsula, by bayesian and maximum likelihood analyses of mitochondrial DNA sequences. Mol. Phylogenet. Evol. 37, 845–857.
Poulakakis, N., Sfenthourakis, S., 2008. Molecular phylogeny and phylogeography of the Greek populations of the genus *Orthometopon* (Isopoda, Oniscidea) based on mitochondrial DNA sequences. Zool. J. Linn. Soc. 152, 707–715.
Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. <beast.bio.ed.ac.uk/Tracer>.
Reynoldson, T.B., 1966. The distribution and abundance of lake-dwelling triclads — towards a hypothesis. In: Anonymous Advances in Ecological Research. Academic Press, pp. 1–71.
Romeis, B., 1989. Romeis Mikroskopische Technik, 17th ed. Urban and Schwarzenberg, München, Wien, Baltimore, pp. 697.
Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.
Rosenbaum, G., Lister, G.S., 2004a. Formation of arcuate orogenic belts in the western Mediterranean region. In: Sussman, A.J., Weil, A.B. (Eds.), Orogenic Curvature: Integrating Paleomagnetic and Structural Analyses. Geological Society of America, Boulder, Colorado, pp. 41–56.
Rosenbaum, G., Lister, G.S., 2004b. Neogene and quaternary rollback evolution of the Tyrrhenian sea, the Apennines, and the Sicilian maghrebides. Tectonics 23, TC1013.
Schettino, A., Turco, E., 2006. Plate kinematics of the western Mediterranean region during the oligocene and early miocene. Geophys. J. Int. 166, 1398–1423.
Schmidt, O., 1861. Untersuchungen über Turbellarien von Corfu und Cephalonia, nebst Nachträgen zu früheren Arbeiten. Z. wiss. Zool. 11, 1–30.
Sfenthourakis, S., Legakis, A., 2001. Hotspots of endemic terrestrial invertebrates in southern Greece. Biodivers. Conserv. 10, 1387–1417.
Sluys, R., 1989. A monograph of the marine triclads. A. A. Balkema, Rotterdam and Brookfield.
Sluys, R., Solà, E., Vila-Farré, M., Mateos, E., Riutort, M., in preparation. Integrative delimitation of species of Mediterranean freshwater planarians (Platyhelminthes, Tricladida, Dugesidae).
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56, 564–577.
Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Phylogenet. Evol. 28, 2731–2739.
Tsigenopoulos, C.S., Durand, J.D., Ünlü, E., Berrebi, P., 2003. Rapid radiation of the Mediterranean luciobarbus species (Cyprinidae) after the Messinian salinity crisis of the Mediterranean sea, inferred from mitochondrial phylogenetic analysis. Biol. J. Linn. Soc. 80, 207–222.
Ursenbacher, S., Schweiger, S., Tomović, L., Crnobrnja-Isailović, J., Fumagalli, L., Mayer, W., 2008. Molecular phylogeography of the nose-horned viper (*Vipera ammodytes*, Linnaeus (1758)): evidence for high genetic diversity and multiple refugia in the Balkan Peninsula. Mol. Phylogenet. Evol. 46, 1116–1128.
van Andel, T.H., Shackleton, J.C., 1982. Late paleolithic and mesolithic coastlines of Greece and the Aegean. J. Field Archaeol. 9, 445–454.
Veith, M., Mayer, C., Samraoui, B., Barroso, D.D., Bogaerts, S., 2004. From Europe to Africa and vice versa: evidence for multiple intercontinental dispersal in ribbed salamanders (Genus *Pleurodeles*). J. Biogeogr. 31, 159–171.
Xia, X., Xie, Z., 2001. DAMBE: software package for data analysis in molecular biology and evolution. J. Hered. 92, 371–373.
Yu, Y., Harris, A.J., He, X.J., 2010. S-DIVA (Statistical Dispersal-Vicariance Analysis): a tool for inferring biogeographic histories. Mol. Phylogenet. Evol. 56, 848–850.

Supplementary Information

Supplementary Figures & Supplementary Tables

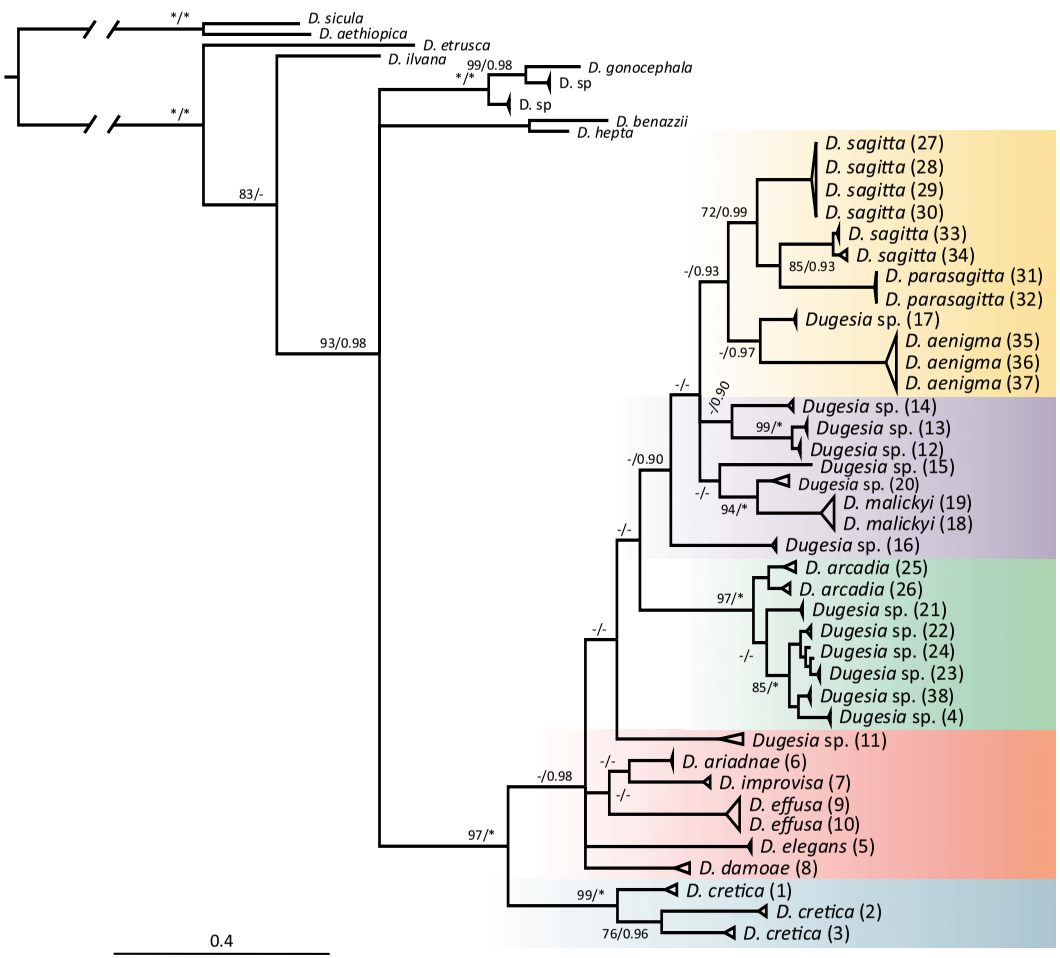
Supplementary figure 1 – Saturation process

Plot of observed transitions and transversions versus corrected distances for COI (A) and ITS-1 (B). The blue rhombs represent transversions, the red triangles transitions.



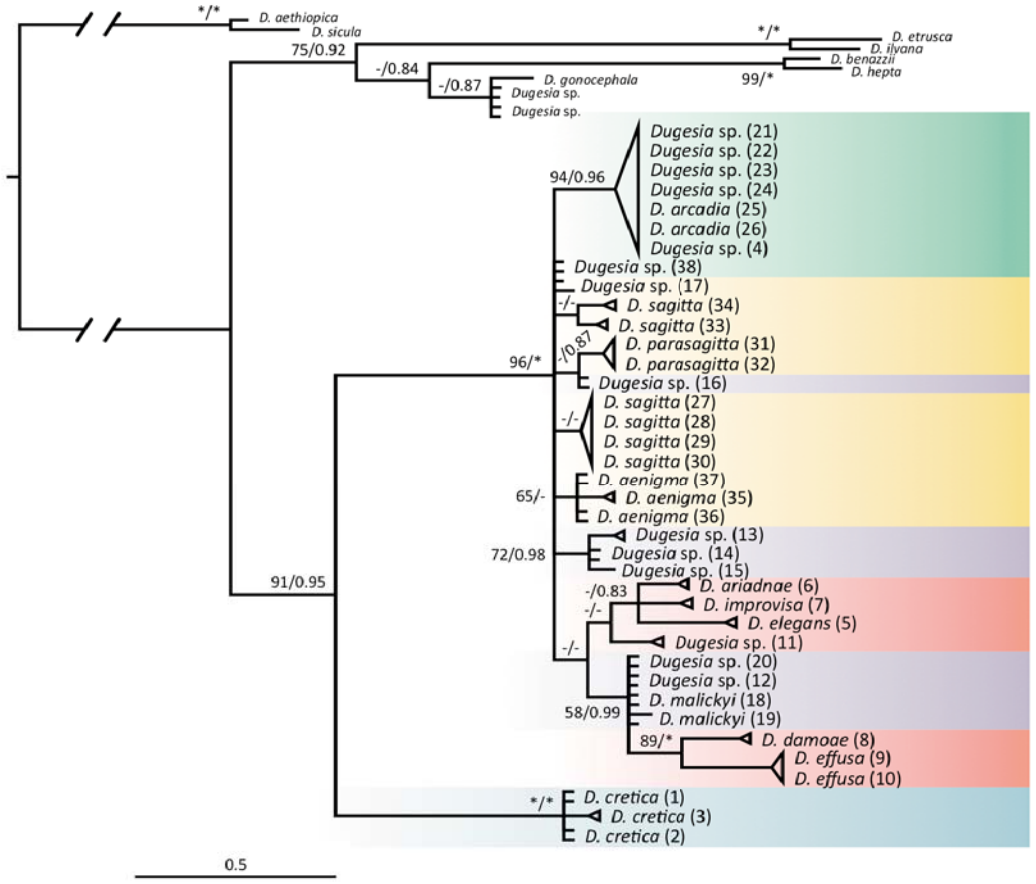
Supplementary figure 2 – Bayesian COI tree

BI phylogenetic tree obtained from the mitochondrial COI alignment. The labels show the species name (when known) and the locality code. Node numbers correspond to bootstrap (ML)/posterior probability (BI). "–" indicates <50/<0.80 and "*" maximum support. Background colour code as in figure 2. The unit of the scale bar is the number of substitution per site.



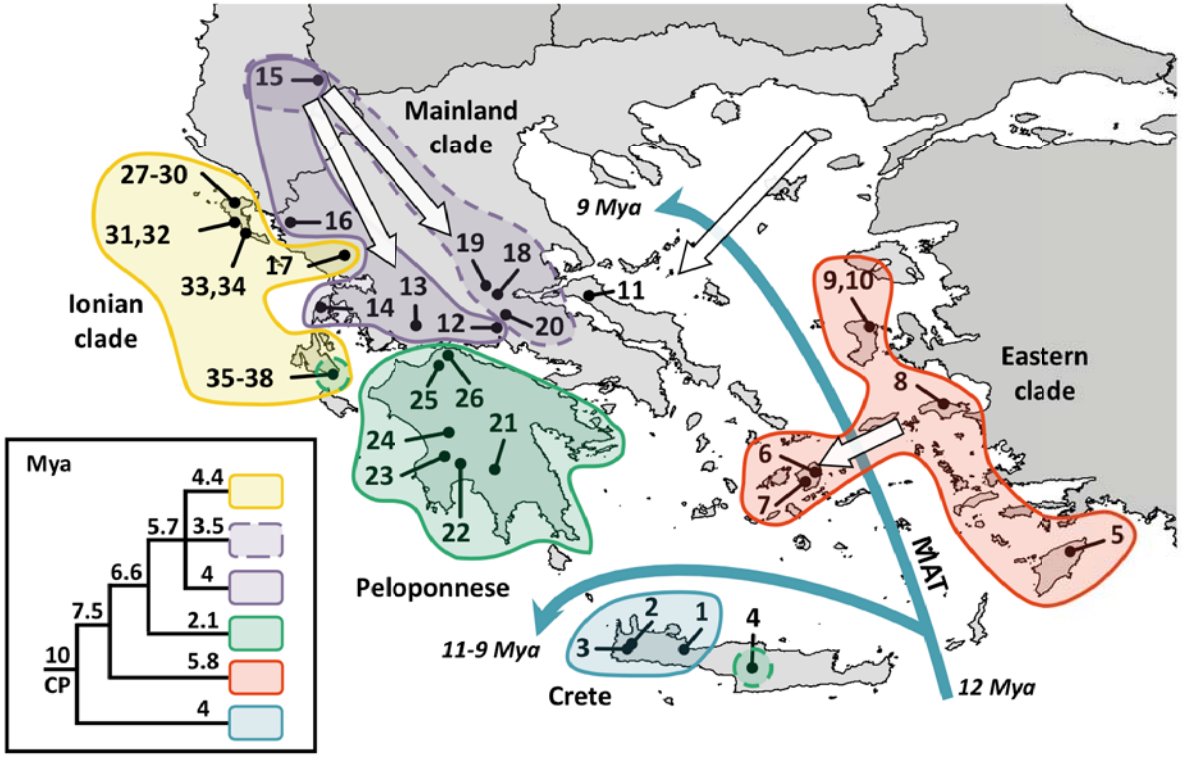
Supplementary figure 3 – Bayesian ITS-1 tree

BI phylogenetic tree inferred from the ITS-1 alignment. The labels show the species name (when known), and the locality code. Node numbers correspond to bootstrap (ML)/posterior probability (BI). "–" indicates <50/<0.80 and "*" maximum support. The unit of the scale bar is the number of substitutions per site.



Supplementary figure 4 – Summary picture

Pictorial scenario showing the distribution of the main clades in the phylogenetic tree (inset shows a condensed tree with the clades and their divergence times) and hypothesized historical biogeographic events. Two main geological events (isolation of Crete and mid-Aegean trench (MAT) opening) are shown as blue arrows. White arrows indicate our main dispersal hypotheses for planarians.



Supplementary table 1. COI genetic distances (in %) between known species used in this study

	Creta *D. cretica*	Samos *D. damoae*	Naxos *D. improvisa*[a]	Naxos *D. ariadnae*	Chios *D. effusa*[a]	Rhodes *D. elegans*	Peloponnisos *D. arcadia*	Mainland *D. malickyi*	Corfu *D. sagitta*	Corfu *D. parasagitta*[a]	Cephalonia *D. aenigma*
D. cretica											
D. damoae	5.4-7.7										
D. improvisa[a]	5.4-7.3	3.8-4.1									
D. ariadnae	6.2-8.2	3.3-3.8	2.8								
D. effusa[a]	6.8-8.4	5.4-6.5	4.6-4.9	5.4-5.7							
D. elegans	8.7-9.3	6-6.5	7.1	6.2	8.4-8.7						
D. arcadia	5.7-7.9	3.6-4.7	4.9-5.4	5.7-6.2	6.8-7.9	7.1-7.9					
D. malickyi	7.3-7.9	5.4	5.2	5.5	7.6-7.9	6.8	6.3-6.8				
D. sagitta	6-9	3.3-6	4.9	4.9-6	5.4-6.5	6.2-8.4	5.4-7.1	5.2-6			
D. parasagitta[a]	7.9-8.5	5.7-6	5.2	6	6.8-7.1	7.6	5.4-6.5	6.8	4.1-4.7		
D. aenigma	8.2-9.6	4.9-6	6-6.5	4.9-6	6.8-7.3	6.2-6.5	6-7.3	4.4-4.9	5.2-7.9	6.8-7.4	

[a] *nomina nuda*

Supplementary table 2. COI genetic distances (in %) between the three main clades from Corfu (*D. sagitta*: 27-30, 33-34; *D. parasagitta*: 31-32).

Corfu	North Corfu (27-30)	Central Corfu (31-32)	Central-South Corfu (33-34)
North Corfu (27-30)			
Central Corfu (31-32)	4.4-4.7		
Central-South Corfu (33-34)	3.8-4.4	4.1-4.4	

Supplementary table 3. COI genetic distances (in %) between three, morphologically identified, *D. cretica* populations from Crete.

Crete	Georgioupoli (1)	Kakopetros (2)	Sasalos (3)
Georgioupoli (1)			
Kakopetros (2)	4.9-5.7		
Sasalos (3)	3.8-4.4	4.6-5.7	

Supplementary table 4. ITS-1 genetic distances (in %) between known species used in this study

	Creta *D. cretica*	Samos *D. damoae*	Naxos *D. improvisa*[a]	Naxos *D. ariadnae*	Chios *D. effusa*[a]	Rhodes *D. elegans*	Peloponnese *D. arcadia*	Mainland *D. malickyi*	Corfu *D. sagitta*	Corfu *D. parasagitta*[a]	Cephalonia *D. aenigma*
D. cretica											
D. damoae	6.2										
D. improvisa[a]	5	1.8									
D. ariadnae	4.7	1.8	0.9								
D. effusa[a]	5.9	1.6	2	1.6							
D. elegans	5	1.6	1.1	1.1	1.8						
D. arcadia	4.8	2	1.6	1.5	2.2	1.8					
D. malickyi	5	1.1	1.1	1.1	1.3	1.3	0.9				
D. sagitta	4.8-5.2	1.3-1.6	1.3-1.6	1.3-1.5	1.5-1.8	1.1-1.3	1.1-1.3	0.7-0.9			
D. parasagitta[a]	5	1.6	1.6	1.6	1.8	0.4	1.3	0.9	0.7-0.9		
D. aenigma	5-5.2	0.9-1.1	1.1-1.3	1.1-1.3	1.3-1.5	0.9-1.1	0.9-1.1	0.4-0.7	0.2-0.7	0.4-0.7	

[a] *nomina nuda*

Supplementary table 5. Ancestral distribution areas probability for the nodes labelled in the supplementary figure 4.

Node	Ancestral distribution probabilities				
1	BCDEFG	BCDEFGHI	BCEFGHI	BCEFG	*
	0.3038	0.2617	0.225	0.1953	0.0142
2	CDEFG	CDEFGHI	CEFG	CEFGHI	*
	0.3038	0.2617	0.225	0.1954	0.0141
21	CDF	CF	*		
	0.5736	0.426	0.0004		
22	CD	C			
	0.5697	0.4303			
23	CD				
	1.0				
3	EG	EGHI	*		
	0.5272	0.4674	0.0054		
4	G	GHI	*		
	0.5642	0.4309	0.0049		
5	G	GHI	*		
	0.5114	0.4728	0.0158		
6	GHI	HI	GH	*	
	0.3717	0.2973	0.2971	0.0339	
61	GI				
	1.0				

* Other ancestral ranges

Chapter 2

Dugesia (Platyhelminthes, Continenticola), a widespread and morphologically homogeneous living genus from the Mesozoic

Reference

Solà E, Stocchino GA, Manconi R, Leria L, Harrath H, Riutort M*. ***Dugesia* (Platyhelminthes, Continenticola), a widespread and morphologically homogeneous living genus from the Mesozoic.** *In preparation*.

* More collaborators will be included in the final manuscript.

Summary

The aim of this study is to seek an answer to an intriguing question surrounding the distribution range of the genus *Dugesia*. These animals are supposed to be poor dispersers as they cannot glide out of freshwater bodies. They are not able to survive in salt water or under desiccation conditions. However, its wide distribution range includes Africa, Europe, Middle East, South Asia, Far East and Australasia. Both assumptions together have led some researchers to wonder about the origin and dispersal routes of *Dugesia* along the geological time. The most recent hypothesis pointed a Gondwanan origin of the genus followed by a posterior dispersal in Eurasia through the Arabian Peninsula and/or India after their collision with the continent, both former pieces of Gondwana. However, the only approach with the aim of answering this question was carried out on the basis of the morphology with no satisfactory answer.

Here we collected with the kind help of many collaborators samples from all across *Dugesia* distribution range and we obtained sequences of 4 gene fragments: 1 mitochondrial and 3 nuclear. The topology of the concatenated phylogenetic tree obtained (four genes) strongly suggests a Pangean origin. Thus the origin of the genus is temptatively fated back to 220 Million years ago (Ma) or to an older time, instead of on Gondwana. An estimation of the lineage divergence times also suggests that the Madagascar *Dugesia*, a group both externally and genetically diverse, was already on this island when it split from Africa about 130 Ma. The European and Middle East *Dugesia* probably colonized the area in a relative recent time from Asian ancestors. Our results make very unlikely a colonization of Eurasia from either ancestors that arribed with India or ancestors crossing through the Arabian plate.

Interestingly, all *Dugesia* species look pretty much the same externally and the inner morphology of the copulatory apparatus in particular is pretty homogeneous. This fact along the putative antiquity of the genus suggests an old morphological stasis.

The present work has only been possible thanks to the collaboration of many researchers around the globe. In the present manuscript we did not include them as authors, but as agreed with them, they will be asked to be among them. The contributors can be found in the Acknowledgements and in the Supplementary Table 1.

Dugesia (Platyhelminthes, Continenticola), a widespread and morphologically homogeneous living genus from the Mesozoic

Eduard Solà[1], Giacinta Angela Stocchino[2], Renata Manconi[2], Laia Leria[1], Halim Harrath[3], Marta Riutort[1*]

[1]Institut de Recerca de la Biodiversitat and Dept. Genètica, Facultat de Biologia, Universitat de Barcelona, Av Diagonal, 643, Barcelona 08028, Spain
[2]Dipartimento di Scienze della Natura e del Territorio, Via Muroni 25, I−07100 Sassari, Italy
[3] Zoology Department, College of Science, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia.

* Correspondence: mriutort@ub.edu

Abstract

Aim

To find out the area of origin and putative historical dispersal routes of the free−living flatworm genus *Dugesia*, a poor and fragile disperser present in Africa, Madagascar, Eurasia, and Australasia. Thus, we aim to explain such wide distribution range and to discover which processes are more likely to have shaped its diversification.

Location

Africa, Madagascar, Arabian Peninsula, Europe, Middle East, India, Far East, Australasia.

Methods

Multilocus molecular Bayesian and Maximum Likelihood−based phylogenetic analyses, divergence time estimations, uncorrelated relaxed clock, and likelihood−based ancestral area reconstructions (AAR).

Results

The phylogenetic analyses clearly split the trees in two equivalent main groups: one mainly containing the genus *Dugesia* from Eurasia and Australasia and the second specimens from Africa, Madagascar and the Arabian Peninsula. The origin of the living *Dugesia* lineages is dated back to the Middle Triassic, before the formation of the European epicontinental seaway (EES) about 220 million years ago (Ma) that severed Europe and Asia from the rest of the supercontinent and before the breakup of Pangaea.

Main conclusions

Dugesia is an old genus that most likely originated on Pangaea during the Late Triassic before the formation of the EES. Therefore, *Dugesia* probably was already widely distributed on Pangaea. More recent alternative calibrations on the split of the two main groups would imply events of wide dispersion overseas that are extremely unlikely due to the incapability of the genus to survive in salt water. Thus, the genus diversified mainly because of ancient dispersal events followed by vicariant processes for a long time. We also found evidences of human−mediated transportation. If correct, *Dugesia* is

a very old genus which actual representatives present both external and internal homogeneous morphology, thus indicating a very long−term morphological stasis.

Keywords

Dugesia, Eurasia, flatworm, Gondwana, historical biogeography, Madagascar, Mesozoic, molecular dating, morphological stasis, Platyhelminthes.

Introduction

The wide distribution of the genus *Dugesia* has been an intriguing issue for over fifty years (Kawakatsu, 1968; Ball, 1974; 1975; Sluys *et al.*, 1998). The species of this freshwater flatworm inhabit lakes, rivers and streams of Africa, Europe, Middle East, South Asia, Far East and Australasia. Among different hypotheses to explain its wide distribution the most recent place the origin of the genus in a Gondwanan scenario (Sluys *et al.*, 1998). The arrival of *Dugesia* in Eurasia from Gondwanan ancestors would be explained either by (i) rafting on the Indian subcontinent between the time it split from Madagascar (*c.* 88 million years ago − Ma) and it collided with Asia (*c.* 55−20 Ma) or (ii) through the impact of the Arabian plate (*c.* 20 Ma).

The distribution of the genus across the continents and islands must be mainly explained by vicariant events posterior to the genus natural dispersion due to the low vagility of freshwater flatworms (Ball, 1974). These free−living platyhelminths are fragile organisms with direct development and without any known resistant stage. Although some species can be found in brackish waters, they are not able to survive in salt water. Moreover, overland or aerial dispersal (e.g. birds) have been considered very unlikely (Reynoldson, 1966). Therefore, it is expected that both landmasses and freshwater bodies splits will be mirrored in the planarians phylogenetic patterns (Ball, 1974).

According to the present distribution of the Dugesiidae representatives it has been agreed that the family probably originated on Pangaea (Ball, 1974; 1975; Sluys, 1998). This supercontinent assembled about 340−320 Ma (Scotese *et al.*, 1979) and broke up in two superterranes (Laurasia and Gondwana) from about 200 to about 160 Ma (Allègre, 1988; Hallam, 1994; Smith *et al.*, 2004). Prior to the Pangaea rifting, at the Late Triassic (*c*. 220−200 Ma), the Tethys Ocean expanded westwards covering present−day Europe with the European epicontinental seaway (ESS), leaving some small islands on the region. This Tethys extension was also contiguous with the boreal Ocean through a narrow connection in the north. Therefore, the EES disconnected Eurasia from the rest of Gondwana (Ziegler, 1988; Newton & Bottrell, 2007). Shortly after the end of the Pangaea breakup, the southern superterrane Gondwana began its own rifting (Jokat *et al.*, 2003, 2005; Scotese, 2004; Schettino & Scotese, 2005). During the Jurassic period (*c.* 175−140 Ma) the Eastern part of Gondwana initiated its breakup,

starting in the Somalia coast of present−day Africa. The sea−floor spreading led to two landblocks separated by the Somalia and Mozambique basins: West Gondwana (South America and Africa) and East Gondwana (Antarctica, India, Seychelles, Madagascar and Australia) (Coffin & Rabinowitz, 1987; Jokat *et al.*, 2003; Rabinowitz & Woods, 2006). Between 160−130 Ma, Madagascar and India together rifted off Eastern Africa and drifted southwards to its present position (Schettino & Scottese, 2005; Ali & Aitchison, 2008 Rabinowitz & Woods, 2006). No later connexion existed between the Indian−Madagascar block and Africa. In the south, Australia and Antarctica migrated away from India−Madagascar about 132 Ma (Powell *et al.*, 1988; Müller *et al.*, 2000; Brown *et al.*, 2003). Madagascar became completely isolated when India and the Seychelles split about 88 Ma and rapidly migrated northwards (e.g., Besse & Courtillot, 1988, 2002; Acton, 1999) until India finally collided with the Tibetan part of Asia about 20−35 Ma (Aitchison *et al.*, 2007; Ali & Aitchison, 2008; Van Hinsbergen *et al.*, 2012). The second contact of a Gondwanan former land and Eurasia took place around 23−16 Ma, when the Arabian plate collided with the Eurasian plate (Robertson, 2000).

The presence of three formaly described *Dugesia* species on Madagascar (*Dugesia milloti* De Beauchamp, 1952; *D. debeauchampi* De Vries, 1988; *D. myopa* De Vries, 1988) may be interpreted as an indirect indicator of the genus antiquity, suggesting an origin anterior to Madagascar isolation in the Cretaceous period. The Malagasy fauna and flora is very rich due to its diversification after a long isolation time. However, the fossil record and molecular divergence time estimation analyses have shown the ancestors of most of the vertebrates on Madagascar to arrive in the island much after its isolation, mainly during the Cenozoic (e.g. Crottini *et al.*, 2012; Tolley *et al.*, 2013). The dispersal way in Madagascar considered as the most likely is rafting overseas, guided by oceanic paleocurrents along the Mozambique channel (Vences *et al.*, 2003; Raxworthy *et al.*, 2002; Poux *et al.*, 2005; Yoder & Nowak, 2006; Samonds *et al.*, 2012; Ali & Huber, 2010). These rafting evidences, along with others clues, have emphasized the importance of long−distance dispersal apart from vicariance in the diversification processes for many fauna and flora across landmasses on Southern continents (Sanmartín & Ronquist, 2004; Yoder & Nowak, 2006).

We have performed a wide sampling effort to obtain specimens of *Dugesia* through all its distribution with the aim to (i) find out the relationships among *Dugesia* populations distributed across Africa, Madagascar, Eurasia and Australasia, (ii) test different biogeographical scenarios to infer the antiquity of the genus and putative

dispersal routes to its present distribution range, (iii) find evidences to support or reject the hypothetical presence of the genus *Dugesia* on Madagascar before its isolation 130 Ma, and (iv) find out which historical paleogeographical events are more likely to have shaped the diversification of the *Dugesia*.

If Gondwanan origin is real, a priori we expected the African species to be older, and hence present a higher genetic diversity among them than the Eurasian lineages. Moreover, either the African or the Malagasy lineages should be more closely related to the present−day Eurasian groups, depending on which of both competing hypothesis is real (dispersal through the Arabian Peninsula or rafting on India). For this reason, in our sampling effort we have specially tried to include specimens from Madagascar, India, Africa and the Arabian Peninsula, because they are vital to test the two existing hypotheses. However, the phylogenetic analyses gave unexpected results showing evidences of a third scenario implying an older origin and dispersion for the genus.

Material and methods

Dugesia sampling and laboratory techniques

Samples of *Dugesia* specimens were obtained from different localities across the genus distribution range (Fig. 1; Supplementary Figs 1, 2 & 8; Supplementary Table 1) by a number of collaborators. Collected individuals were fixed and preserved in absolute ethanol for molecular analyses. Generally, two specimens per locality were sequenced, but only one per locality was included in the analyses, excepting few cases where high molecular divergence was detected.

The collected specimens were referred to their species when they had been previously identified for other works by analyzing their inner morphology. Those specimens for which such identification was not possible to be carried out were checked with a binocular loupe and considered as *Dugesia* sp. after recognizing their typical external characteristics.

Total genomic DNA was extracted from ethanol−preserved specimens using the commercial reagent DNAzol (Molecular Research Center Inc. Cincinnati, OH) by following the manufacturer's instructions.

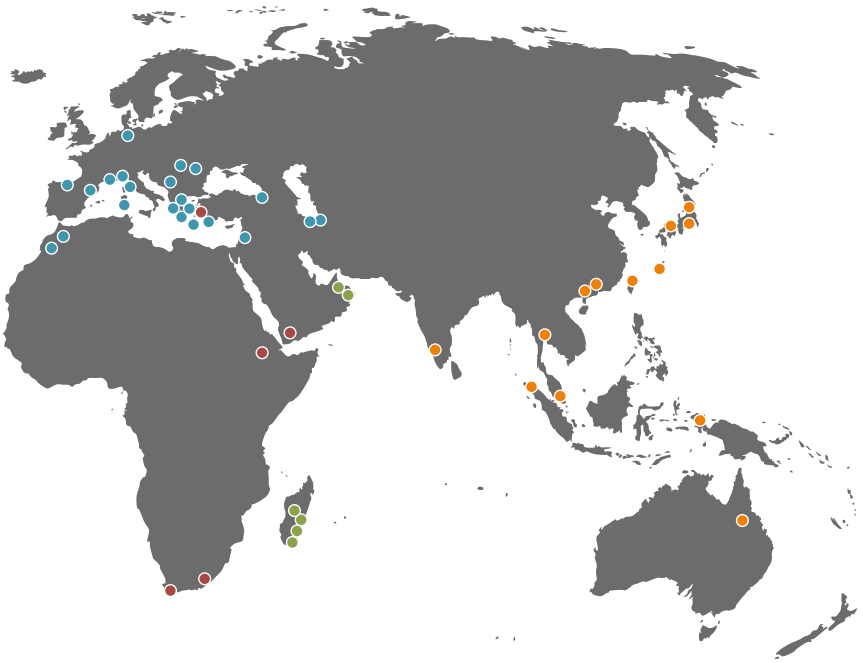


Figure 1 Approximate areas from which *Dugesia* samples used in the present study were obtained. One circle may include different localities. The different colours indicate the major phylogenetic clades to which the specimens belong. More detailed maps of Europe and Madagascar are presented in the Supplementary Material (Supplementary Figs 1, 2 & 8). For further details on specimens localities see Supplementary Table 1.

We obtained four gene sequence fragments by polymerase chain reaction (PCR): the mitochondrial gene *cytochrome c oxidase subunit I* (Cox1), and the nuclear genes *18S ribosomal gene*, *28S ribosomal gene*, and the *ribosomal internal transcribed spacer−1* (ITS−1). The 18S and 28S sequences were obtained by the amplification of two overlapping fragments. The final volume of the PCR reaction was 25 µl. To 1 µl of DNA we added: (1) 5 µl of Promega 5x buffer, (2) 1 µl of dNTPs (10 mM), (3) 0.5 µl of each primer (forward and reverse) (25 µM), (4) 0.15 µl of *Taq* polymerase (GoTaq® Flexi DNA of Promega). To complete the final 25 µl PCR volume we added double distilled and autoclavated water. We used specific primers to amplify the different genome region (Supplementary Table 2). In order to obtain the amplification for some genes it was necessary to vary the annealing temperatures or the amount of $MgCl_2$ or DNA.

Before sequencing, PCR products were purified by the use of a vacuum system (MultiScreen$^{TM}_{HTS}$ Vacuum Manifold, Millipore). The sequencing reactions were carried out and run either in an automated sequencer ABI Prism 3730 by the Unitat de Genòmica of Centres Científics i Tecnològics of the Universitat de Barcelona

(CCiTUB) or by Macrogen Corporation in Europe (Amsterdam, the Netherlands). The primers used for sequencing were the same than those used to amplify the fragments, excepting the Cox1 forward, using a more internal primer (COIEF3, JAPO or COIEFM) when it was not possible to sequence with primer BarT. The sequences were checked by eye in the software GENEIOUS 6.1.7 (Biomatters, 2014).

Sequence alignment and phylogenetic analyses

In order to carry out the phylogenetic analyses we have obtained sequences of the mitochondrial gene Cox1, ITS−1, 18S and 28S. A part from the sequences obtained in this study, we retrieved additional sequences of other *Dugesia* species available in GenBank (Supplementary Table 1).

Sequences of the three nuclear ribosomal genes or regions were aligned using the online software MAFFT version 7 using the G−INS−i algorithm (Katoh & Standley, 2013). The alignments were checked by eye and manually edited with the software GENEIOUS 6.1.7. The mitochondrial coding gene Cox1 sequences were first translated into amino acids (genetic code 9 in NCBI) in order to check the presence of stop codons, then were aligned using the *Translation align* function implemented in GENEIOUS 6.1.7. Regions of 'doubtful' homology in 18S, 28S and ITS−1 alignments were removed using the software Gblocks (Talavera & Castresana, 2007), allowing half gap positions in the alignment and the minimum number of sequences for a flank position to the minimum value allowed.

In order to calculate which evolutionary model fits the best the molecular data JMODELTEST 2.1.1 (Posada, 2008) was run using the Akaike Information Criterion (AIC) calculations for each gene sepparately.

We used two phylogenetic inference approaches (i) maximum likelihood (ML) using the software RAXML 7.4.2 (Stamatakis, 2006) and (ii) Bayesian inference (BI) either with MRBAYES 3.2.2 (Ronquist *et al.*, 2012) or BEAST v.1.7.3 (Drummond & Rambaut, 2007).

MRBAYES analyses were run for each gene independently and for a concatenated dataset including all the genes with two simultaneous runs of 1 cold and 5 hot chains given the high number of terminals. Each run was performed with 10 million generations with sampling parameters every 10^3 and a 25% default burn−in value for the final trees. The convergence of the topologies and model parameters of the two runs was surveyed by ensuring the average standard deviation of split frequencies fell below

0.01. It was also checked that the likelihood had reached stationarity. The maximum likelihood analyses were performed under the GTRGAMMAI and 1000 bootstrap pseudoreplicates.

In order to find out which *Dugesia* lineages must be used to root the genus tree we first carried out a phylogenetic analysis using the methods mentioned above including representatives of its sister genera *Schmidtea* and *Recurva* (Sluys *et al.*, 2013) using the nuclear 18S and 28S genes in a concatenated dataset.

Additionally, we carried out a phylogenetic analysis (both BI and ML) including all populations from the Oriental region from the present study and from previous papers for which there are Cox1 sequences available (Bessho *et al.*, 1992; Álvarez−Presas *et al.*, 2008; Lázaro *et al.*, 2009; Zhang *et al.*, 2010; Sakai & Sakaizumi, 2012). Our purpouse was to draw a focused phylogenetic picture of *Dugesia* distributed in this area in which many different studies dealing with this genus are done.

Estimation of lineage divergence times

We applied different calibration scenarios to estimate the *Dugesia* divergence times using the software BEAST 1.7.3, setting up the input file with the software BEAUTI. We used the 18S, 28S and Cox1 first and second position gene datasets including all *Dugesia* specimens used in this study and the outgroup species (*Recurva postrema*, *Schmidtea mediterranea* and *S. polychroa*). We forced the monophyly of different groups on the basis of the rooted and the concatenated tree results (Fig. 2; Supplementary Fig. 3): (i) all *Dugesia* specimens (ingroup); (ii) the populations affected by the formation of the mid−Aegean trench (MAT) (Solà *et al.*, 2013) and its (iii) Western and (iv) Eastern Aegean subgroups. The substitution models were set as GTR + I + Γ with empirical base frequencies for all the genes. The model clock was assessed as uncorrelated relaxed clock for all the genes and the tree prior was set under the Yule Process of speciation. We used as a fix calibration point for all the analyses the formation of the MAT, for which we applied a normal distribution for the MRCA of species impacted by this event (Mean = 10.5; Stdev = 0.6). The second calibration point was applied on the divergence of the two main *Dugesia* groups (referred here as Gondwanan and Eurasian, see the Results section). We used different times between the formation of the European epicontinental sea (EES) about 220−200 Ma and the aftermath of the Permo−Carboniferous glaciations about 270 Ma, proposed by Ball (1974) as the diversification time for the Dugesiidae family. Thus, we set five normal

distributions along this period every 10 Ma (Means = 220, 230, 240, 250, 260; Stdev = 4). Alternatively we also calibrated at 150, 100 and 50 Ma under the same distribution parameters to check for more recent divergences although no biogeographical hypotheses were assessed for these times in order to discard that we found an optimal but random calibration scenario within our 'local' time set. Depending on the effective sample size (ESS) scores we run 2−3 MCMC Bayesian analyses for 20 million generations, resulting in 2−3 files of 20,000 trees each. Log files were inspected in TRACER v.1.5 (Rambaut & Drummond, 2009) to assess that the ESS of the combined log files reached values over 200 for all parameters (Drummond *et al.*, 2006). The burn−in at 10% and the tree combination were conducted in LOGCOMBINER and TREEANNOTATOR (Drummond & Rambaut, 2007). The different scenarios were posteriorly tested using Bayes Factors (BFs) (1000 replicates for likelihood), to assess which of the alternative datings best explained our data.

Ancestral Area Reconstruction (AAR)

In order to test the putative ancestral distribution range of the different *Dugesia* lineages we carried out the likelihood dispersal−extinction−cladogenesis (DEC) approach implemented in the software LAGRANGE (Ree *et al.*, 2005; Ree & Smith, 2008). We used a calibrated pruned tree without some phylogenetically and geographically close terminals in order to make the analysis more tractable (Supplementary Table 1). The areas were defined according to the tectonic plates excepting the Eurasian and African plates that were split in two; Europe and Asia and Africa and Madagascar respectively. The areas were: i) Europe; ii) Asia; iii) Africa; iv) Arabian Peninsula; v) India; vii) Australasia (Eastern to Wallace line); viii) Madagascar. We also applied stratified dispersal constraint matrices for different spans of time on the basis of their geological isolation (see Supplementary Table 3). The maximum number of areas in ancestral ranges was held at two. Dispersal constraints were set to 1.0 when landmasses were connected and 0.1 when landmasses were disjunct.

Results

Phylogenetic relationships

The final datasets for the alignment including only *Dugesia* specimens, after Gblocks pruning, had a length of: 744 bp for the mitochondrial gene Cox1; 568 bp for the ITS−1; 1,545 bp for the 18S; 1,666 bp for 28S. The concatenated dataset has a total length of 4,523 bp with a 12.1% of missing data. The final dataset including the outgroup (*Schmidtea* and *Recurva*) had a length of: 1,688 bp for the 18S; 1,532 bp for 28S. The concatenated dataset has a total length of 3,220 with a 12.8% of missing data (Supplementary Table 1). The evolutionary model used for all the genes and all the analyses was GTR + I + Γ as this was the result of JMODELTEST 2.1.1. Cox1 was partitioned in all the analyses considering all three positions separately (all positions = GTR + I + Γ). Therefore, the concatenated analyses contained 6 partitions: the 3 nuclear genes and the 3 Cox1 partitions by position. The Cox1 alignment dealing with *Dugesia* specimens from the Oriental region contained 834 bp and was also partitioned by position (1st = KHY + I + Γ; 2 and 3 = GTR + I + Γ).

The preliminary trees including the genera *Recurva* and *Schmidtea* as outgroup showed the *Dugesia* specimens to be split in two symmetric sister clades: one containing animals from Africa, Madagascar and Oman, and the second containing animals from Europe, Middle East and Asia (Supplementary Fig. 3). The first group that we will call 'Gondwanan' contains two 'outliers' from Greece. The second group, the 'Eurasian', also includes 'outlier' specimens from Morocco and Australasia (New Guinea and Australia). The Gondwanan and Eurasian clades were also distinctly separated and well supported when different concatenated and single gene unrooted trees with only *Dugesia* species were checked (Supplementary Figs 4, 5, 6 & 7). Also the relationships within the two main groups were basically the same in the rooted and unrooted trees although with better support for the nodes in the later. Therefore, we used one of the two main lineages to root the following phylogenetic trees.

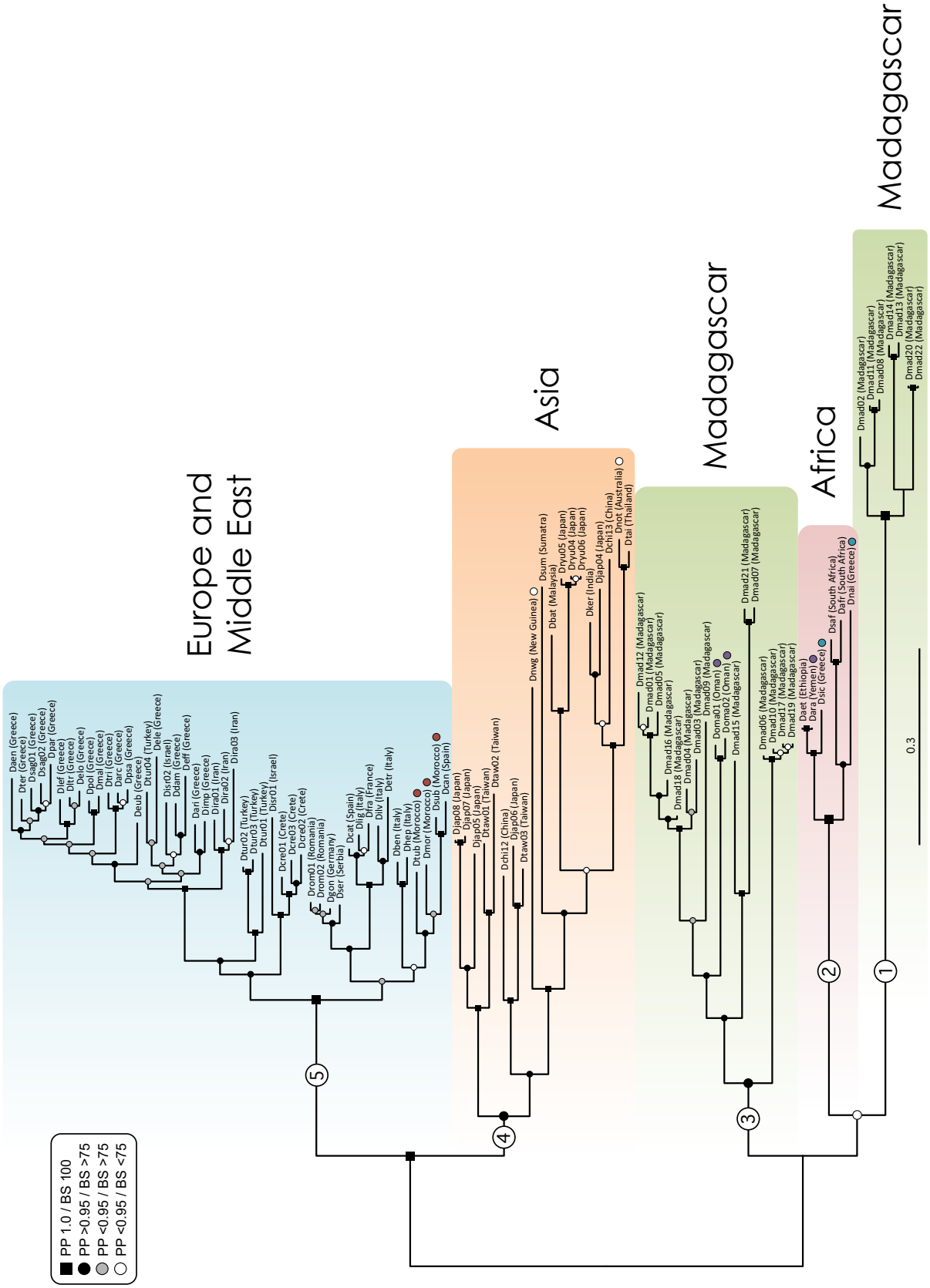
The concatenated tree (Fig. 2) shows three well−differentiated clades within the Gondwanan lineage although the relationships between them could not be resolved as they conform a trichotomy. Two out of the three clades contain *Dugesia* from Madagascar (1 and 3). One of these two subgroups is comprised by different populations of a high triangular−shaped head such as those of *D. milloti*, previously described from the same island (1; see De Vries, 1988). The other Malagasy clade (3)

includes other populations from Madagascar with different external colour patterns and appearances. Interestingly, within the later group there are *Dugesia* from two localities from Oman (Arabian Peninsula) well−nested. The tree branches of the Malagasy specimens of the group 1 are relatively long in the concatenated tree and especially in the ITS−1 and Cox1 gene trees (Supplementary Figs 6 & 7). The third Gondwanan clade (2) contains species from South Africa, Ethiopia, Yemen and *D. sicula* and *D. naiadis* from Chios. *D. sicula* is also known from many localities across the African and European Mediterranean coast, and it is considered to be of African origin having expanded mainly thanks to human activity (Lázaro *et al.*, 2013).

The Eurasian lineage essentially comprises *Dugesia* from Eurasia, but it also includes *Dugesia* from Morocco and Australasia (New Guinea and Australia). The Eurasian clade is split in two main groups, one contains the populations from Asia and Australasia (4), and the second includes *Dugesia* from Europe, Middle East and Morocco (5). Both clades are well supported in the concatenated tree. However, the Asian group is paraphyletic in the 28S and ITS−1 gene trees (Supplementary Figs 5 & 6). The Asian sublineage presents populations from different geographical regions mixed−up. Thus, we observe specimens from Taiwan, Japan, and China in different clades. Interestingly, the Australian species *D. notogaea* (*Dnot*) is molecularly the sister taxon of the Thailandese *Dugesia* (*Dtai*), while the specimen from New Guinea (*Dnwg*) is placed among *Dugesia* from China, Japan, Sumatra and Taiwan.

The European and Middle East lineage (5) is also showing an internal structure, separating the Aegean and Middle East species from the Western and Central European *Dugesia*. The specimens from Morocco are related to the *Dugesia* sp. (*Dcan*) from Spain (Iberian Peninsula). It is interesting that the European and Middle East clade presents relatively shorter branches in comparison with the Asian and Gondwanan lineages.

Figure 2. Bayesian tree inferred from the concatenated dataset (18S, 28S, ITS−1, Cox1). Rooting based on a previous analysis (Supplementary Fig. 3). Taxa labels correspond to codes in Supplementary Table 1 and in brackets the region where they were sampled. Numbers in white circles correspond to the main lineages. Specimens that are considered geographical 'outliers' are accompanied with a colored circle: Red, Africa; White, Australasia; Purple, Arabian Peninsula; Blue, Europe. Supports descriptions are presented within the upper-left box; PP corresponds to the Posterior Probability (BI) and BS to the Bootstrap (ML). The scale bar indicates substitutions per site.


Europe and Middle East
Asia
Madagascar
Africa
Madagascar
PP 1.0 / BS 100
PP >0.95 / BS >75
PP <0.95 / BS >75
PP <0.95 / BS <75
0.3
1
2
3
4
5
Daen (Greece)
Dter (Greece)
Dsag01 (Greece)
Dsag02 (Greece)
Dief (Greece)
Dlef (Greece)
Dltr (Greece)
Delo (Greece)
Dpol (Greece)
Dmal (Greece)
Dtri (Greece)
Darc (Greece)
Dpsa (Greece)
Deub (Greece)
DturO4 (Turkey)
Dele (Greece)
Disr02 (Israel)
Ddam (Greece)
Deff (Greece)
Dari (Greece)
Dimp (Greece)
Dira01 (Iran)
Dira02 (Iran)
Dira03 (Iran)
Dtur02 (Turkey)
Dtur03 (Turkey)
Dtur01 (Turkey)
Disr01 (Israel)
Dcre01 (Crete)
Dcre03 (Crete)
Dcre02 (Crete)
Drom01 (Romania)
Drom02 (Romania)
Dgon (Germany)
Dser (Serbia)
Dcat (Spain)
Dlig (Italy)
Dfra (France)
Dliv (Italy)
Detr (Italy)
Dben (Italy)
Dhep (Italy)
Dtub (Morocco)
Dmor (Morocco)
Dsub (Morocco)
Dcan (Spain)
Djap08 (Japan)
Djap07 (Japan)
Djap05 (Japan)
Dtaw01 (Taiwan)
Dtaw02 (Taiwan)
Dchi12 (China)
Djap06 (Japan)
Dtaw03 (Taiwan)
Dnwg (New Guinea)
Dsum (Sumatra)
Dbat (Malaysia)
Dryu05 (Japan)
Dryu04 (Japan)
Dryu06 (Japan)
Dker (India)
Djap04 (Japan)
Dchi13 (China)
Dnot (Australia)
Dtai (Thailand)
Dmad12 (Madagascar)
Dmad01 (Madagascar)
Dmad05 (Madagascar)
Dmad16 (Madagascar)
Dmad18 (Madagascar)
Dmad04 (Madagascar)
Dmad03 (Madagascar)
Dmad09 (Madagascar)
Doma01 (Oman)
Doma02 (Oman)
Dmad15 (Madagascar)
Dmad21 (Madagascar)
Dmad07 (Madagascar)
Dmad06 (Madagascar)
Dmad10 (Madagascar)
Dmad17 (Madagascar)
Dmad19 (Madagascar)
Daet (Ethiopia)
Dara (Yemen)
Dsic (Greece)
Dsaf (South Africa)
Dafr (South Africa)
Dnal (Greece)
Dmad02 (Madagascar)
Dmad11 (Madagascar)
Dmad08 (Madagascar)
Dmad14 (Madagascar)
Dmad13 (Madagascar)
Dmad20 (Madagascar)
Dmad22 (Madagascar)

None of the obtained topologies give support to any of the two original hypotheses we wanted to test, since the divergence and levels of genetic differentiation of the Eurasian group are equivalent to those of the African−Malagasy lineages, and there is neither a closer relationship of the Eurasian clade to any of the African nor Madagascar lineages as we expected.

The phylogenetic tree obtained from the Cox1 fragment that includes all *Dugesia* of the Oriental region for which sequences of this gene is available (Supplementary Figs 8 & 9) supported the geographical mixed condition of the group already seen in the general phylogenetic tree. Although the topology of the tree is not generally well−supported some structure can be guessed. The clade with more representatives (1) includes animals from China, Japan and Taiwan, but it is not supported. On the other hand, *D. ryukyuensis* (*Druy04−06*) species from the small island of Okinawa is well delimited as well as its sister relationship with *D. batuensis* (*Dbat*) from Malaysia, fact also seen in the general tree (Fig. 2). These two species are clustered with low support within a clade including *D. notogaea* from Australia (*Dnot*), the Thailandese species (*Dtai*) and a specimen from Japan (*Djap04*) that may be an undescribed species (2). The third distinctive lineage also encompasses specimens from these three areas (3).

Estimation of lineage divergence times

The divergence time estimation was carried out without including the ITS−1 due to the difficulty to align this sequence between genera. We neither used the third position of Cox1 because its saturation led to low ESS values in previous analyses (DAMBE analysis, data not shown; Xia & Xie, 2001).

The divergence time estimation in Platyhelminthes is a rather complicated issue as far as its fossil record is almost non−existant (Dentzien−Dias *et al*., 2013 and references therein) providing no clue about the relative antiquity of the different groups. Therefore, the only way to carry out dating analyses when dealing with organisms of this phylum implies the use of paleogeographical events to calibrate phylogenetic trees (Lázaro *et al.*, 2011; Scarpa *et al*., 2013; Solà *et al.*, 2013) or alternatively indirect estimations from other groups by the use of subtitution rates or secondary calibration points. Despite the associated limitations and risks of calibrating trees with paleogeographical events (Magallon, 2004; Heads, 2005; Kodandaramaiah, 2011), the poor dispersal capability of freshwater flatworms (Ball, 1974) may be an advantage to

cautiously trust estimations based on them. *Dugesia* dispersion is conditioned by the continuity of freshwater bodies on landmasses. Therefore, vicariant events may be treated as minimum times of divergence for this group.

According to the topology of the phylogenetic tree obtained using the concatenated dataset (Fig. 2), the mixed geographical distribution of different specimens across Asia and their general wide divergence within the Asian lineage prevent the use of paleogeographical−based calibration points using nodes within this clade. On the other hand, within the European lineage (5) we detected one subgroup that may be useful for calibration and already reported in a previous work (Solà *et al.*, 2013). It corresponds to the formation of the mid−Aegean trench (MAT) about 11−9 Ma (Dermitzakis & Papanikolau, 1981), which split the Aegean area in a Western and an Eastern part. It had an impact on the diversification of different fauna on the region and also on that of *Dugesia* species present in this area. Therefore, we used this event in all the divergence time estimation analyses as an inner calibration point. As we also aimed to find out if the Malagasy *Dugesia* lineages diversification would be explained by the split of Madagascar from Africa (*c.* 130 Ma) we did not use any calibration point within the lineage containing the Malagasy and African populations. As an external calibration point for our phylogenetic tree we propose different scenarios. First, the symmetrical and very well−supported configuration of two main *Dugesia* lineages in the rooted and the concatenated phylogenetic tree presented in this work consisting of a Gondwanan and an Eurasian clade (Fig. 2), and the comparatively similar lengths of the basal branches (Fig. 3) suggest that the most recent common ancestor of the present representatives diversified approximately at the same time. To explain such topology it is necessary to hypothesize a diversification of the genus previous to the split of Eurasia from the rest of Pangaea, which occurred when it was isolated by the formation of an epicontinental seaway about 220 Ma, before the supercontinent breakup starting at 200 Ma (Ziegler, 1998; Newton & Bottrell, 2007). However, it is possible that these two lineages already diverged before this landmass severing. Supporting this idea is the fact that at that time *Dugesia* have had to be already widely distributed across the present Eurasia and Gondwana former landmasses. Therefore, we tested different scenarios between 220 Ma and the aftermath of the Permo−Carboniferous glaciation (Ball, 1974). On the other hand we tested three more recent scenarios based on no biogeographical hypotheses (50, 100, 150 Ma) that would have implied long overseas dispersal of these fragile salt−sensitive animals.

Between the eight different scenarios used to calibrate the *Dugesia* phylogenetic tree the analysis by Bayes Factors supported the most likely to be the one establishing the split of the two main *Dugesia* lineages at about 240 Ma (Table 1). This scenario was substantially or strongly better than the 220, 230 and 250 Ma but not different than the 260 Ma scenario. Nonetheless, its likelihood was slightly higher than the last one. Interestingly, the majority of the comparisons of the splits calibrated at more recent times (50, 100, 150 Ma) were strongly or decisively worst than any of the times between 220 and 260 Ma. They also received much lower likelihood values, thus pointing to lower supports for these divergence times to fit our data.

The best scenario (240 Ma) showed a divergence of the two main groups at 237 Ma (229−245 Ma) (Fig. 3). Within the Gondwanan group, the first Malagasy group to diverge (3) split at 144.6 Ma but showing a wide 95% high posterior density credibility interval (95% HPD; 105−191). The divergence of the second Malagasy group (2) from the African clade (1) occurs at 132.8 Ma also showing a wide HPD 95% (95.5−174.3 Ma). The topology of the relationships between the African and Malagasy lineages is different from that obtained in MrBayes (Fig. 2). This is explained by the fact that such relationships are not well−supported and differently recovered across the different analyses. The Oman specimens within the second group of Madagascar split at 46.7 Ma (30.3−66.5 Ma), thus postdating the isolation of the island. The common ancestor of the Eurasian lineages diverged at 129 Ma (97.7−166.8 Ma), at an equivalent time in comparison with the Gondwanan group (144.6 Ma). The Asian clade began to diversify at 112.6 Ma (86.5−146.4 Ma). Within this group, the 'outliers' from New Guinea (*Dnwg*) and Australia (*Dnot*) split from its sister lineages at 88.1 Ma (52−95.9 Ma) and 7.8 Ma (3−14.8 Ma) respectively. The diversification of the European and Middle East clade started at 36.5 Ma (26.8−47.9 Ma).

It is noteworthy to highlight the fact that these calibrations are based on paleogeographical processes temptatively related to vicariant events. Thus, many risks are associated such as the underestimation of the divergence times (considering that *Dugesia* is not able to disperse overseas) and a circular reasoning (Kodandaramaiah, 2011).

The obtained evolutionary rate were $3.03 \cdot 10^{-4}$ ($\pm$ $1.265 \cdot 10^{-6}$) per site per lineage per million years for 18S; $2 \cdot 10^{-4}$ ($\pm$ $1.55 \cdot 10^{-6}$) for 28S; and $2.15 \cdot 10^{-3}$ ($\pm$ $1.57 \cdot 10^{-5}$) for the first and second position of Cox1.

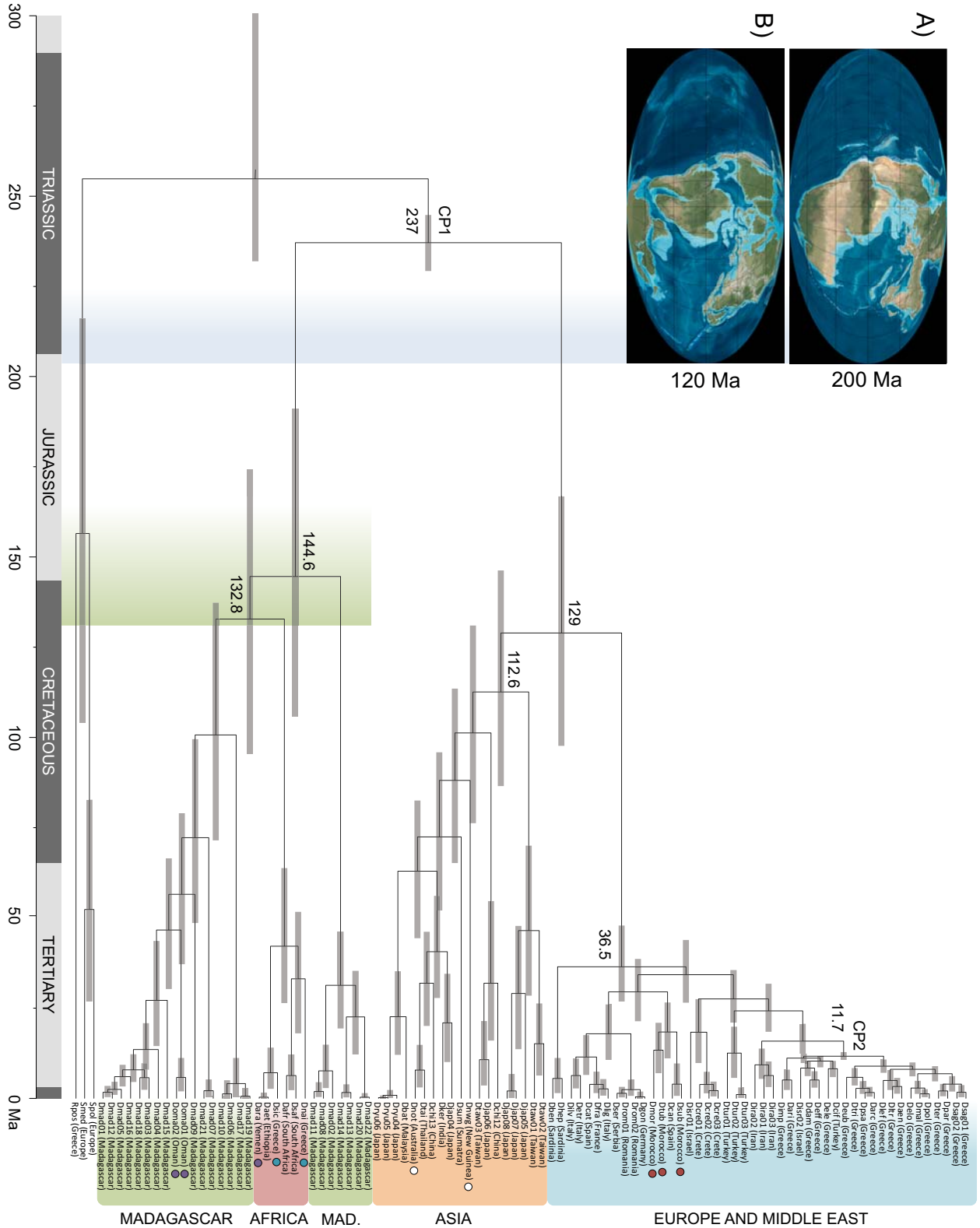
Table 1. Bayes Factors results for the comparison of the temporal scenarios. Probabilty of the three models with standard error and log10 Bayes Factors. * Indicates a substantial evidence against H_0; **, strong evidence; ***, decisive evidence.

	Scenario	lnP (model|data)	S.E	1	2	3	4	5	6	7	8
1	CP1: 12−9 Ma CP2: 210−230 Ma	−20854.832	+/− 0.363	−	−0.491	-1.272^{**}	-0.758^{*}	-0.863^{*}	2.336^{***}	1.708^{**}	0.529^{*}
2	CP1: 12−9 Ma CP2: 220−240 Ma	−20853.701	+/− 0.385	0.491	−	-0.781^{*}	−0.267	−0.372	2.827^{***}	2.199^{***}	1.02^{**}
3	**CP1: 12−9 Ma CP2: 230−250 Ma**	**−20851.904**	**+/− 0.374**	$\mathbf{1.272^{**}}$	$\mathbf{0.781^{*}}$	−	$\mathbf{0.514^{*}}$	**0.409**	$\mathbf{3.608^{***}}$	$\mathbf{2.98^{***}}$	$\mathbf{1.801^{**}}$
4	CP1: 12−9 Ma CP2: 240−260 Ma	−20853.087	+/− 0.373	0.758^{*}	0.267	-0.514^{*}	−	−0.105	3.094^{***}	2.466^{***}	1.287^{**}
5	CP1: 12−9 Ma CP2: 250−270 Ma	−20852.846	+/− 0.328	0.863^{*}	0.372	−0.409	0.105	−	3.199^{***}	2.571^{***}	1.392^{**}
6	CP1: 12−9 Ma CP2: 40−60 Ma	−20860.211	+/− 0.366	-2.336^{**}	-2.827^{***}	-3.608^{***}	-3.094^{***}	-3.199^{***}	−	-0.628^{*}	-1.807^{**}
7	CP1: 12−9 Ma CP2: 90−110 Ma	−20858.765	+/− 0.396	-1.708^{**}	-2.199^{***}	-2.98^{***}	-2.466^{***}	-2.571^{***}	0.628^{*}	−	-1.179^{**}
8	CP1: 12−9 Ma CP2: 140−160 Ma	−20856.05	+/− 0.436	-0.529^{*}	-1.02^{**}	-1.801^{**}	-1.287^{**}	-1.392^{**}	1.807^{**}	1.179^{**}	−

Ancestral area reconstruction

The likelihood AAR were implemented under the dispersal−extinction−cladogenesis (DEC) model in LAGRANGE. Likelihood AAR for living *Dugesia* results suggested that the family was already distributed across Europe or Middle East and Africa when it split from its sistergroup (*Schmidtea* and *Recurva*) (Fig. 4). About 240 Ma, prior to the Pangaea breakup, *Dugesia* diverged by vicariance, most probably one lineage on each hemisphere. Later, the European clade dispersed to Asia. The presence of *Dugesia* on Australia, India and New Guinea from Asian ancestors is here interpreted as dispersal followed by vicariance.

Figure 3. Dated tree using as a calibration point a mean of 240 Ma (Stdev = 4) for the split of the two main *Dugesia* lineages (CP1) according to the Bayes Factors analysis (Table 1). CP2, calibration point using the age of formation of the mid-Aegean trench (MAT). Grey bars at nodes represent the 95% highest posterior density (HPD) credibility interval. The meaning of the circles accompanying some specimens is explained in Fig. 2. The mean diversification age of the main lineages are shown on the tree. Vertical color bars indicate the periods of the formation of the European epicontinental seaway (ESS) (blue) and the split of Madagascar from Africa (Green). Mollewide projections showing the Pangaea configuration at 200 Ma (A), where the EES on Europe can be observed, and at 120 Ma (B) showing the separation of Madagascar and India from Africa. Paleogeographical maps from Ron Blakey at http://jan.ucc.nau.edu/rcb7/.


A)
B)
120 Ma
200 Ma
300
250
200
150
100
50
0 Ma
TRIASSIC
JURASSIC
CRETACEOUS
TERTIARY
CP1
237
144.6
132.8
129
112.6
36.5
CP2
11.7
MADAGASCAR
AFRICA
MAD.
ASIA
EUROPE AND MIDDLE EAST

The *Dugesia* distributed on the present−day Africa dispersed to Madagascar in first place when they still were contiguous landmasses, where it diversified in a new lineage. Later, a second Malagasy and an African lineages diverged by vicariance probably because the split of the island. *D. naiadis* (*Dnai*) on Greece and *D. arabica* (*Dara*) on the Arabian Peninsula are explained by dispersal from African ancestors followed by vicariance. The same situation applies for the Moroccan specimens nested within the European lineage.

Discussion

Which hypothesis fits best the *Dugesia* observed diversification patterns?

Considering the *Dugesia* specimens analyzed in the present study it seems very unlikely that the Eurasian species originated from Gondwanan ancestors as previously suggested by Ball (1974, 1975) and Sluys (1998). Our assumption rises from the fact that the Eurasian lineages are neither nested within the diverse and also deep Malagasy group nor within (or directly related with) the African lineage. In contrast, there is a clear polarity between the Gondwanan and the Eurasian lineages (Fig. 2; Supplementary Fig. 3).

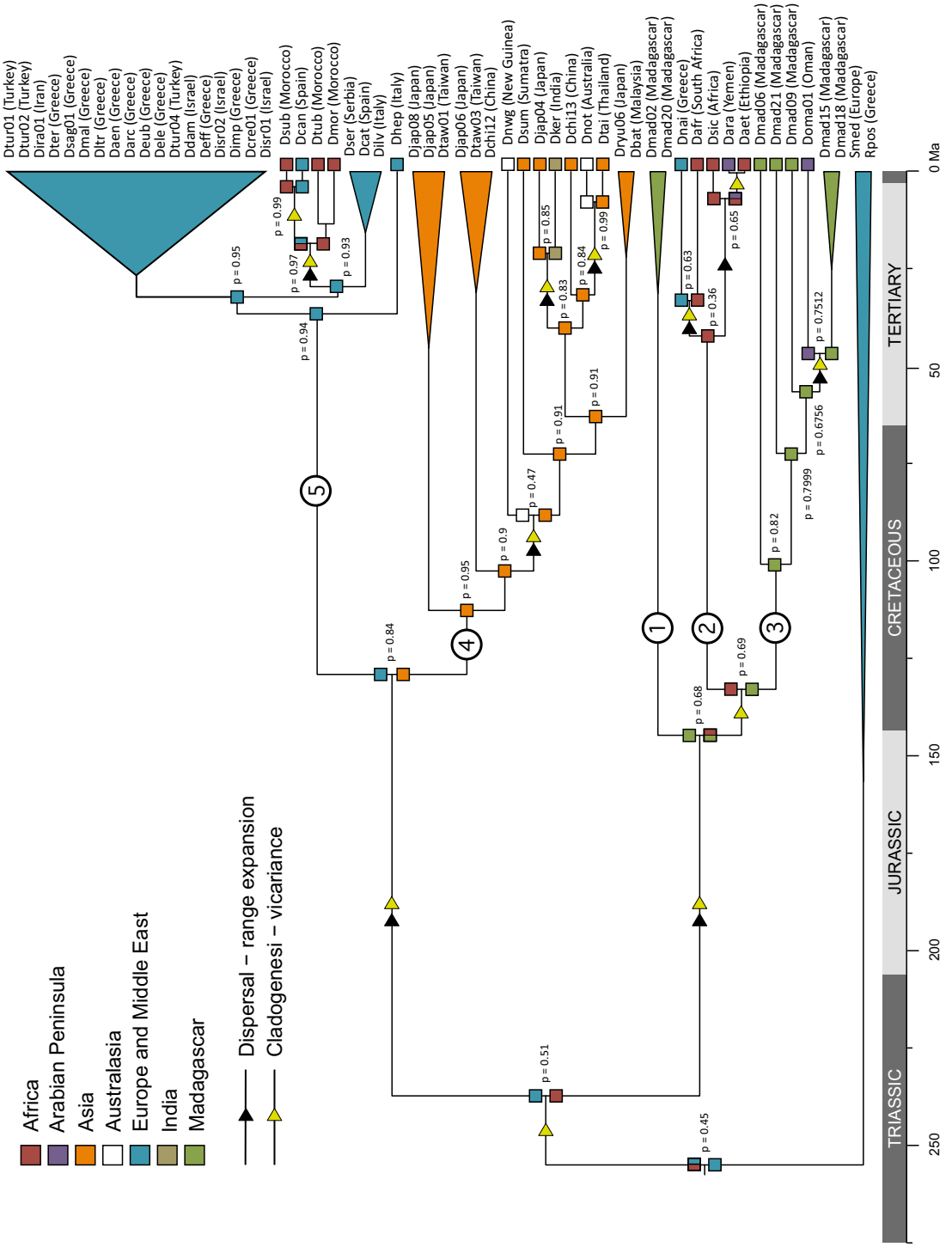
In case the Eurasian lineage originated from Madagascar ancestors, we would expect them to be nested within the Malagasy lineage. This would have suggested the dispersal within Asia through the collision of India with the continent about 60 Ma after its split and drifting from Madagascar about 88 Ma. However, the only specimen of *Dugesia* we obtained from India is also nested within the Asian clades in an inner position. We intentionally pursued to obtain samples from the Western Ghats region on Southwest India due to the fact that its freshwater bodies are known to contain a high number of endemic species, some of them related with species from Madagascar or the Seychelles (Datta−Roy & Karanth, 2009 and references therein). Other studies willing to check if Indian freshwater species are of Gondwanan origin showed that many organisms present in India actually dispersed out of Asia in this region (e.g. freshwater gastropods, Köhler & Glaubrecht, 2007). Although it is still possible that some *Dugesia* species on Asia were brought on India thus according to the ferry model/out−of−India hypothesis (Datta−Roy & Kranth, 2009), our samplings did not yield specimens related with Malagasy animals. However, we cannot discard that they inhabit the Southern

Asian fresh waters or that they went extinct. Interestingly, *D. astrocheta* from Africa (not sampled) is considered as closely related with *D. burmaensis* from India according to their morphology (Sluys *et al.*, 1998).

The second hypothesis considered by both Ball (1974) and Sluys (1998) proposed the dispersal of *Dugesia* through the impact of the Arabian plate with Eurasia at about 20 Ma. This scenario is also very unlikely under the light of our results. First, according to this hypothesis we would have expected a closer relationship between the African and Eurasian lineages. Second, even considering an African sampling or extinction bias, the similar relative diversification times between the Eurasian and the Gondwanan lineages also temptatively rejects this dispersion−through−Arabia hypothesis (Fig. 3).

In consequence the present day distribution of *Dugesia* could only be explained if (1) some species of the genus were already present throughout the Pangaea when Eurasia was isolated due to the EES; or (2) a more recent origin occurred either within Gondwana or Laurasia derived regions and it dispersed posteriorly. The second alternative however would probably imply long dispersals overseas, between Eurasia or Madagascar and Africa. For such fragile animals this possibility seems highly unlikely. In consequence we considered the Pangaean origin the best explanation for our results and tested the putative date for the original splitting by using multiple calibration dates along the period comprised between 220 and 260 Ma (means), that we later compared through Bayes Factors. We also tested multiple more recent dates to make sure that the data could not be better explained by some younger splitting although we did not have a priori any paleogeographical event that could explain such recent splitting. The results show our data to be better explained by a basal split for *Dugesia* occurring around 237 Ma (from 254 to 229 Ma).

Figure 4. Ancestral area reconstruction (ARR) for *Dugesia* and its sister genera. The seven areas used in the analysis are shown on the upper−left corner of the figure. The AAR with the higest likelihood are shown as colored boxed at each node along its relative probability of the global likelihood. Boxes with one color indicate the ancestor to be confined to a single geographic area; combined boxes indicate an ancestor with a distribution across two areas. Two boxes, one on each branch, indicate the ancestral ranges inherited by each of the daughter lineages arising from the node.


Dtur01 (Turkey)
Dtur02 (Turkey)
Dira01 (Iran)
Dter (Greece)
Dsag01 (Greece)
Dmal (Greece)
Dltr (Greece)
Daen (Greece)
Darc (Greece)
Deub (Greece)
Dele (Greece)
Dtur04 (Turkey)
Ddam (Israel)
Deff (Greece)
Disr02 (Israel)
Dimp (Greece)
Dcre01 (Greece)
Disr01 (Israel)
Dsub (Morocco)
Dcan (Spain)
Dtub (Morocco)
Dmor (Morocco)
Dser (Serbia)
Dcat (Spain)
Dliv (Italy)
Dhep (Italy)
Djap08 (Japan)
Djap05 (Japan)
Dtaw01 (Taiwan)
Djap06 (Japan)
Dtaw03 (Taiwan)
Dchi12 (China)
Dnwg (New Guinea)
Dsum (Sumatra)
Djap04 (Japan)
Dker (India)
Dchi13 (China)
Dnot (Australia)
Dtai (Thailand)
Dryu06 (Japan)
Dbat (Malaysia)
Dmad02 (Madagascar)
Dmad20 (Madagascar)
Dnai (Greece)
Dafr (South Africa)
Dsic (Africa)
Dara (Yemen)
Daet (Ethiopia)
Dmad06 (Madagascar)
Dmad21 (Madagascar)
Dmad09 (Madagascar)
Doma01 (Oman)
Dmad15 (Madagascar)
Dmad18 (Madagascar)
Smed (Europe)
Rpos (Greece)
p = 0.99
p = 0.97
p = 0.93
p = 0.95
p = 0.94
p = 0.84
p = 0.95
p = 0.9
p = 0.47
p = 0.85
p = 0.83
p = 0.84
p = 0.99
p = 0.91
p = 0.91
p = 0.63
p = 0.36
p = 0.65
p = 0.82
p = 0.7999
p = 0.6756
p = 0.7512
p = 0.68
p = 0.69
p = 0.51
p = 0.45
1
2
3
4
5
Africa
Arabian Peninsula
Asia
Australasia
Europe and Middle East
India
Madagascar
Dispersal – range expansion
Cladogenesi – vicariance
TRIASSIC
JURASSIC
CRETACEOUS
TERTIARY
250
200
150
100
50
0 Ma

Historical biogeography of the genus *Dugesia*

The well−supported clustering of the African with the Malagasy specimens drawing a trichotomy within the Gondwanan clade strongly suggests that the split of these three lineages happened before or during the Madagascar split from Africa between 160 and 130 Ma. The dating analysis we took (140 Ma) temptatively supports this hypothesis (mean of the two Malagasy lineage split at 144.6 and 132.8 Ma). The Eurasian lineages remaining out of this group may indicate that they diverged before this event and had a former extense distribution along the eastern part of Pangaea.

The *Dugesia* lineages from Madagascar are not yet described formally because all of them were collected as asexual populations. However, many of them show wide genetic and external appearance diversity between their populations. The trichotomy conformed by the two main Malagasy and the African groups may indicate a radiation event that could be due to the split of Madagascar from Africa, thus supporting the presence of the genus on the island prior to this event. Indeed, the AAR analysis (Fig. 4) suggests an African distribution of the Gondwanan *Dugesia*, diversifying first after it dispersed in the present−day Madagascar (at that time a contiguous landmass) and later, a second lineage diversified on the island and a third on Africa when they split. This hypothesis is implicitly supported by the dating analysis that points the clade divergence of the two Malagasy groups at 144.6 and 132.8 Ma respectively (Fig. 3). Within the Gondwanan clade, the group 1 containing those Malagasy populations with high triangular heads may have experienced an accelerated evolution rate on their Cox1 and ITS−1 genes as it can be seen in the gene trees (Supplementary Figs 6 & 7).

The case of the *Dugesia* collected from two different localities in Oman and well nested within a Malagasy group (3) in the concatenated and all gene trees is particularly interesting (Supplementary Figs 3, 4, 5 & 6). Some biogeographical studies have suggested a contact between the migrating Indian plate and the Arabian Peninsula along its way to Asia according to the close relationships found between organisms from India with those of Oman (Rage, 2003; Van Bocxlaer *et al.*, 2006). However, this proposal has been controversial and strongly rejected by some researchers (Ali & Aitchison, 2008). We here do not propose such geological situation to be the cause of the Arabian−Malagasy tight phylogenetic connection. On the other hand, according to our dating analysis the divergence of the Arabian animals from its *Dugesia* sistergroup postdates the Indian drifting from Madagascar, thus excluding a round trip from

Madagascar to India and from there to the Arabian Peninsula after it collided with Asia. Therefore, the best explanation for this Arabian−Malagasy relationship would imply accidental human transportation from populations not sampled for this work, an event already known for a few distributional 'outliers' *Dugesia* species (Solà *et al.*, 2013; Lázaro *et al.*, 2014). In a speculative manner we propose the next situation that may explain such disjunct distribution. About 2,200 B.C., mariners from Southern Asia went in regular voyages from the Indus Valley to Mesopotamia and to the Horn of Oman (Ratnagar, 1981; Cleuziou and Tosi, 1994). It has been speculated that this sailors explored farther south along the coasts of Arabia and Africa (Wright & Rakotoarisoa, 2003), perhaps reaching Madagascar. It is possible that they reprovisioned of fresh water there, bringing accidentally Malagasy *Dugesia* specimens with them to their next stop.

The Asian and Australasian lineage (4) also presents interesting features. Its relative diversification is slightly more recent than the split of the Malagasy lineage from the African group (Fig. 3), thus suggesting a comparable divergence time. Additionally, the dating analysis suggests an old diversification of the genus (112.6 Ma) on the region, flourishing in multiple deep and shallower lineages. According to the AAR results, the Asian *Dugesia* ancestors could have originated from European populations that dispersed eastwards (Fig. 4). The ancient diversification on the area would explain the presence of different lineages on China and on the islands studied (Taiwan, Japan, and Okinawa), which *Dugesia* would have reached before these landmasses isolated from the mainland or taking advantatge of sea level changes. Due to the relative low depth of the submerged continental shelf of this region (Sunda shelf, Sahul shelf or Taiwan shelf) the Far East had become subaerial during eustatism episodes, mainly during the last 30 Ma and especially during the Pleistocene glacial maxima, connecting different islands with the continent and between them. These events have allowed freshwater bodies from the different landmasses to converge during these episodes (Boggs *et al*., 1979; Voris, 2000). The focus on the Oriental lineage carried out in the present work on the basis of a fragment of the mitochondrial Cox1 gene (Supplementary Fig. 9) also showed some interesting results that may have future taxonomic implications. The tree includes the species *Dugesia notogaea*, *D. ryukyuensis*, *D. batuensis*, and a bunch of unidentified species or temptatively referred as *D. japonica* (*Dchi01−12*, *Djap01−03*). All these latter specimens are contained in a not well−supported clade (1) that may include the actual *D. japonica*. However, the

relatively long branches would suggest an old origin of different independent lineages and therefore the presence of different putative species. On the other hand, the weakly−supported clade may point to a not monophyletic condition of this group. Interestingly, it has been suggested that the well−supported clade (3) with relatively long branches sepparating the different taxons may be related with *D. austroasiatica* (M. Kawakatsu personal communiaction) for which the original area is still unknown (Sluys *et al.*, 2010). These results point to *D. japonica* being a conundrum of species.

The Wallace Line has been traditionally considered a barrier between the Southeastern fauna and the Australasian, drawing a limit between Borneo and Sulawesi (Celebes) and through the Lombok Strait between Bali and Lombok (Mayr, 1944). Interestingly, both *Dugesia* specimens from New Guinea (*Dnwg*) and Australia (*Dnot*) are nested within the Asian clade (4). Furthermore, these two 'outliers' are neither monophyletic nor molecularly close. The Australian *D. notogaea* is the sistergroup of the Thailandese (*Dtai*) specimen by a relatively short branch separating them (7.8 Ma in the dating analysis; Fig. 3), while the New Guinean *Dugesia* is not closely related to any sampled specimen (divergence at 88.1 Ma), phylogenetically placed in the middle of the Asian lineage. As far as the two sides of the Wallace Line have never been in contact, it seems that this situations may be explained either by a human−mediated transportation or because they naturally reached the islands by an unkown dispersal way. In freshwater crabs there are also reports of animals that apparently crossed the Wallace Line, probably by rafting (Klaus *et al.*, 2010). However, rafting is a very unlikely way to disperse for freshwater flatworms because of their fragility under diseccation conditions or in salt water.

The European and Middle East lineage (5) is a relatively compact clade in comparison to the Asian or Gondwanan lineages, sharing a younger diversification. In spite of its relative branch shortness it contains a wide diversity of described *Dugesia* species covering a wide distribution range. Thus suggesting a relatively recent dispersion and diversification of the genus on the area. According to the dating analysis, the common ancestor of the present representatives of this clade began to diversify about 36.5 Ma (47.9−26.8). Interestingly, about this time, the Middle East and Europe that used to be constituted by many islands started to be a contiguous landmass with Asia. The long branch separating the European and Middle East clade from its common ancestor with the Asian clade is suggesting a major European extinction when the Gondwanan and Eurasian lineages split followed by a colonization from not sampled

Asian ancestors. Alternatively, some *Dugesia* populations could have been isolated in the remaining European islands after the Tethys Sea covered Europe and expanded once Europe became a continguous landmass again. Indeed, the AAR results would point to the last hypothesis (Fig. 4). However, it could be either an artifact because of undersampling or because of the outgroup (exclusively European genera with very few species).

Gathering all the different evidences from the present study, the most reasonable explanation is that the origin and diversification of *Dugesia* took place on the Pangaea, the first split of *Dugesia* would more probably have been prior to its breakage. A more recent origin in the Gondwana would have implied long dispersal overseas of such a fragile and salt water sensitive animal between, for instance, Eurasia or Madagascar and Africa. Although a more recent origin of *Dugesia* seems very unlikely, we cannot discard an older origin of the genus than we propose in the present work. Despite the caution with which we have to take our dating analysis, it is interesting that the divergence inferred for the outgroup species *Schmidtea polychroa* and *S. mediterranea* (52 Ma) is reasonably similar to that inferred in a previous work (40 Ma; Lázaro *et al.*, 2011). The diversification of fauna groups that matches or predates the fragmentation of Laurasia and Gondwana has also been found in reptiles, amphibians, mammals and invertebrate lineages (e.g. Springer *et al.*, 2003; Roelants & Bossuy, 2005; San Mauro *et al.*, 2005; Wildman *et al.*, 2007; Gamble *et al.*, 2008; Giribet et al., 2012) and in plants (Mao *et al.*, 2012).

Our results would imply a wide distribution of the genus already on Pangaea, probably on the western region. The absence of the genus *Dugesia* on North and South America could be interpreted in three different ways: (i) the genus is actually on these continents but it has never been reported, (ii) *Dugesia* was occupying this territory long ago but it came extinct, or (iii) geological or climatical barriers on Pangaea and Gondwana did not allow the expansion of *Dugesia* from the Central−Eastern Pangaea to the West. Considering the third point, the proposal of barriers in the former Pangaea was already suggested by Scotese (2004) regarding terrestrial organisms. Unfortunatelly, no fossil will probably ever resolve such biogeographical enigmas. Comparative historical biogeography, especially on other freshwater triclads and organisms, will probably lead to a more complete picture of Pangaea and Gondwana, considering both geological and historical biogeographical advances.

Morphological stasis in the genus *Dugesia*

The species of this genus are known to have a very similar external appearance presenting the characteristic triangle−shaped head and the two eyes in free−pigment patches. Moreover, the inner morphology of the copulatory apparatus although presenting different characteristics among the *Dugesia* species is quite homogeneous. The results of this work likely place the origin of the genus *Dugesia* in the Early Mesozoic, back to about 237 Ma. This case suggests an extreme morphological stasis in this Dugesiid genus. It has kept a very constant appearance among its lineages through a long geological period. The morphological stasis is not rare in freshwater flatworms, as it has also been proposed for the genus *Girardia* from South America (Sluys *et al.*, 2005). Morphological stasis has been also reported in many animal groups, for organisms such as mygalomorphs (Hamilton *et al.*, 2014), salamanders (Min *et al.*, 2005), and coelacanths (Holder *et al.*, 1999), among others. One example of a long−time period morphological stasis case includes the *Pantodon* fishes, which is a considered an extreme case of phenotypic stasis because they have barely changed over 57 Ma (Lavoué *et al.*, 2011). Nonetheless, according to our tree topology and dating analysis, the *Dugesia* could even be a more extreme case, being a genus as old as at least two hundred million years. This would make arise questions about how is selective pressure acting on these animals, leading to temptative answers pointing to a 'comfort state', meaning that they do not need more adaptations to survive successfuly in freshwater. On the contrary, *Dugesia* could be under a strong selective pressure or it may just be morphologicaly constrained. Supporting this hypothesis is its wide distribution range, only absent in polar or very cold areas. The small changes detected in the inner morphology of the different species would be driven by stochasticity after vicariant events.

Conclusions

The previous hypotheses on the origin of *Dugesia* have placed its diversification in a Gondwanan scenario and the origin of the Dugesiidae family on Pangaea. In the present work we gathered different evidences from molecular phylogenies that suggest an older origin for the genus, on the supercontinent Pangaea. According to our dating results, *Dugesia* origin would have taken place about 237 Ma. However, due to the limitations

of our calibration approach, the presence of big areas still not sampled and the intrinsic exclusion of extinct lineages we are cautious with our results. Nonetheless, the present work rejects the previous hypotheses, implying an improvement from previous approaches trying to disentangle the striking wide distribution of such a poor disperser.

Acknowledgements

We are indebted to Abdolvahab Masghsoudlou, Benedicta Oshuware, Biju Kumar, Duenguen Krailas, Fengqing Li, Goran Zivanovik, Khang Tsung Fei, Leon Blaunstein, Michael Balke, Mei−Hui Li, Midori Matsumoto, Ori Segev, Reza Bagherzadeh, Savel R. Daniels, Sonih Khisihing, and Xiaoli Tang who kindly contributed with samples. They will be kindly requested to be part of the authorship of the present paper. We are also indebted to Ignacio Ribera, Miklos Balint and Núria Bonada who also provided material. This research was supported by CGL 2011-23466 Grant (to M.R) and by the doctoral fellowship BES−2009−022530 from the Ministerio de Ciencia e Innovacion from Spain (to E.S.).

References

Acton, G.D. (1999) Apparent polar wander of India since the Cretaceous with implications for regional tectonics and true polar wander. *The Indian Subcontinent and Gondwana: a Palaeomagnetic and Rock Magnetic Perspective* (ed. by T. Radhakrishna and J.D.A Piper), pp. 135-159. Geological Society of India Memoirs, Vol. 44.

Aitchison, J.C., Ali, J.R. & Davis, A.M. (2007) When and where did India and Asia collide? *Journal of Geophysical Research: Solid Earth,* **112B**.

Ali, J.R., & Aitchison, J.C. (2008) Gondwana to Asia: Plate tectonics, paleogeography and the biological connectivity of the Indian sub−continent from the Middle Jurassic through latest Eocene (166−35 Ma). *Earth−Science Reviews,* **88**, 145−166.

Ali, J.R. & Huber, M. (2010) Mammalian biodiversity on Madagascar controlled by ocean currents. *Nature,* **463,** 653−657.

Allègre, C.J. (1988) *The Behavior of the Earth*. Harvard University Press, Cambridge.

Álvarez−Presas, M., Baguñà, J. & Riutort, M. (2008). Molecular phylogeny of land and freshwater planarians (Tricladida, Platyhelminthes): from freshwater to land and back. *Molecular Phylogenetics and Evolution,* **47**, 555−568.

Ball, I.R. (1974) A contribution to the phylogeny and biogeography of the freshwater triclads (Platyhelminthes, Turbellaria). *Biology of the Turbellaria* (ed. by N. W. Riser and P. M. Morse), pp. 339−401. McGraw−Hill, New York.

Ball, I.R. (1975) Nature and formulation of biogeographical hypotheses. *Systematic Zoology,* **24**, 407–430.

Besse, J. & Courtillot, V. (1988) Paleogeographic maps of the continents bordering the Indian Ocean since the early Jurassic. *Journal of Geophysical Research: Solid Earth,* **93B,** 1791–1808.

Besse, J. & Courtillot, V. (2002) Apparent and true polar wander and the geometry of the geomagnetic field over the last 200 Myr. *Journal of Geophysical Research: Solid Earth,* **107B**.

Bessho, Y., Ohama, T. & Osawa, S. (1992) Planarian mitochondria. I. Heterogeneity of cytochrome c oxidase subunit I gene sequences in the freshwater planarian, *Dugesia japonica*. *Journal of Molecular Evolution,* **34,** 324–330.

Biomatters (2014) *Geneious version 6.1.7*. Available at http://www.geneious.com/. Biomatters Limited.

Boggs, Jr.S., Wang, W.C., Lewis, F.S. & Chen, J.C. (1979) Sediment properties and water characteristics of the Taiwan shelf and slope. *Acta Oceanographica Taiwanica,* **10,** 10–49.

Brown, B., Müller, R.D., Struckmeyer, H.I.M., Gaina, C., Stagg, H. & Symonds, P. (2003) Formation and evolution of Australian passive margins: implications for locating the boundary between continental and oceanic crust. *Evolution and Dynamics of the Australian Plate* (ed. by R.R. Hillis and R.D. Müller), pp. 223–243. *Geological Society* of *Australia* Special Publication N. 22 and Geological Society of America Special Paper N. 273.

Cleuziou, S. & Tosi, M. (1994) Black boats of Magan, some thoughts on Bronze−Age water transport in Oman and beyond from the impressed bitumen slabs of Ra's al−Junayz. *South Asia Archaeology 1993* (ed. by A. Parpola and P. Koskikallio Soumalaisen), pp. 745−61. Vol. 2. Tiedakatemian Toimituskia, Helsinki.

Coffin, M.F. & Rabinowitz, P.D. (1987) Reconstruction of Madagascar and Africa: Evidence from the Davie Fracture Zone and western Somali Basin. *Journal of Geophysical Research,* **92,** 9385–9406.

Crottini, A., Madsen, O, Poux, C., StrauB, Vieites, D. & Vences, M. (2012) Vertebrate time−tree elucidates the biogeographic pattern of a major biotic change around the K−T boundary in Madagascar. *Proceedings of the National Academy of Sciences USA,* **109,** 5358–5363.

Datta-Roy, A. & Karanth, K.P. (2009) The Out−of−India hypothesis: what do molecules suggest? *Journal of Biosciences,* **34,** 687–697.

De Beauchamp, P. (1952) Sur les *Dugesia* (Turbellariés Triclades) d'Afrique tropicale et de Madagascar. *Bulletin de la Société Zoologique de France,* **77,** 362–370.

De Vries, E. (1988) A synopsis of the nominal species of the subgenus *Dugesia* (Platyhelminthes: Tricladida: Paludicola) from Africa and Madagascar. *Zoological Journal of the Linnean Society,* **92,** 345−382.

Dentzien−Dias, P.C., Poinar, G., de Figueiredo, A.E.Q., Pacheco A.C.L., Horn B.L.D. & Schultz C.L. (2013) Tapeworm eggs in a 270 million−year−old shark coprolite. *PlosOne,* **8,** e55007.

Dermitzakis, M. & Papanikolaou, D.J. (1981) Paleogeography and geodynamics of the Aegean region during the neogene. *Annales Géologiques des Pays Hélléniques,* **20,** 245–289.

Drumond, A. & Rambaut A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology,* **7,** 214.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. & Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology,* **4,** e88.

Gamble, T., Bauer, A.M., Greenbaum, E. & Jackman, T.R. (2008) Evidence for Gondwanan vicariance in an ancient clade of gecko lizards. *Journal of Biogeography,* **35,** 88–104.

Giribet, G., Sharma, P.P., Benavides, L.R., Boyer, S.L., Clouse, R.M., De Bivort, B.L., Dimitrov, D., Kawauchi, G.Y., Murienne, J. & Schwendinger, P.J. (2012) Evolutionary and biogeographical history of an ancient and global group of arachnids (Arachnida: Opiliones: Cyphophthalmi) with a new taxonomic arrangement. *Biological Journal of the Linnean Society,* **105,** 92–130.

Hallam, A. (1994) *An Outline of Phanerozoic Biogeography*. Oxford University Press, New York.

Hamilton, C.A., Hendrixson, B.E., Brewer, M.S. & Bond, J.E. (2014) An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: A case study of the North American tarantula genus Aphonopelma (Araneae, Mygalomorphae, Theraphosidae). *Molecular Phylogenetics and Evolution,* **71,** 79−83.

Heads, M. (2005) Dating nodes on molecular phylogenies: a critique of molecular biogeography. *Cladistics,* **21,** 62−78

Holder, M.T., Erdmann, M.V., Wilcox, T.P., Caldwell, R.L. & Hillis, D.M. (1999) Two living species of coelacanths? *Proceedings of the National Academy of Sciences USA,* **96,** 12616–12620.

Jokat, W., Boebel, T., König, M. & Meyer, U. (2003) Timing and geometry of early Gondwana breakup. *Journal of Geophysisical Research: Solid Earth,* **108,** 2428.

Jokat, W., Ritzmann, O., Reichert, C. & Hinz, K (2005) Deep crustal structure of the continental margin of the Explora Escarpment and in the Lazarev Sea. East Antarctica. *Marine Geophysical Researches,* **25,** 283–304.

Katoh, K. & Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Evolution,* **30,** 772−780.

Kawakatsu, M. (1968) On the origin and phylogeny of Turbellarians: Suborder Paludicola. *Japanese Society* of *Systematic Zoology*, *Circular,* **38−41**, 11−22.

Klaus, S., Schubart, C.D., Streit, B. & Pfenninger, M. (2010) When Indian crabs were not yet Asian—evidence for Eocene proximity of India and Southeast Asia from freshwater crab biogeography. *BMC Evolutionary Biology,* **10,** 287.

Kodandaramaiah, U. (2011) Tectonic calibrations in molecular dating. *Current Zoology,* **57,** 116−124.

Köhler, F. & Glaubrecht, M. (2007) Out of Asia and into India: on the molecular phylogeny and biogeography of the endemic freshwater gastropod *Paracrostoma* Cossmann, 1900 (Caenogastropoda: Pachychilidae). *Zoological Journal of the Linnean Society,* **91,** 621–657.

Lavoué, S., Miya, M., Arnegard, M.E., McIntyre, P.B., Mamonekene, V. & Nishida. M. (2011) Remarkable morphological stasis in an extant vertebrate despite tens of millions of years of divergence. *Proceedings of the Royal Society B: Biological Sciences,* **278**, 1003−1008

Lázaro, E.M. & Riutort, M. (2013) *Dugesia sicula* (Platyhelminthes, Tricladida): the colonizing success of an asexual planarian. *BMC Evolutionary Biology,* **13,** 268.

Lázaro, E.M., Harrath, H., Stocchino, G., Pala, M., Baguñà, J. & Riutort, M. (2011) *Schmidtea mediterranea* phylogeography: an old species surviving on a few Mediterranean islands? *BMC Evolutionary Biology,* **11,** 274.

Lázaro, E.M., Sluys, R., Pala, M., Stocchino, G.A., Baguñà, J. & Riutort, M. (2009) Molecular barcoding and phylogeography of sexual and asexual freshwater planarians of the genus *Dugesia* in the Western Mediterranean (Platyhelminthes, Tricladida, Dugesiidae). *Molecular Phylogenetics and Evolution,* **52,** 835−845.

Magallon, S.A. (2004) Dating lineages: molecular and paleontological approaches to the temporal framework of clades. *International Journal of Plant Sciences,* **165,** S7–21.

Mao, K., Milne, R.I., Zhang, L., Peng, Y., Liu, J., Thomas, P., Mill, R.R. & Renner, S.S. (2012) Distribution of living Cupressaceae reflects the breakup of Pangea. Proceedings of the National Academy of Sciences of USA 109, 7793–7798.

Mayr, E. (1944) Wallace's Line in the light of recent zoogeographic studies. The Quarterly Review of Biology 19,1–14.

Min, M.S., Yang, S.Y., Bonett, R.M., Vieites, D.R., Brandon, R.A. & Wake, D.B. (2005) Discovery of the first Asian plethodontid salamander. *Nature,* **435,** 87–90.

Müller, R.D., Gaina, C. & Clarke, S. (2000) Seafloor spreading around Australia. Billion−year Earth History of Australia and Neighbours in Gondwanaland (ed. by J.J. Veevers), pp. 18–28. Gemoc Press, Sydney.

Newton, R.J., & Bottrell, S.H. (2007) Stable isotopes of carbon and sulphur as indicators of environmental change: past and present. *Geological Society of London Journal,* **164,** 691–708.

Posada, D. (2008) jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution,* **25,** 1253–1256.

Poux, C., Madsen, O., Marquard, E., Vieites, D.R., de Jong, W.W. & Vences, M. (2005) Asynchronous Colonization of Madagascar by the four endemic clades of primates, tenrecs, carnivores, and rodents as inferred from Nuclear genes. *Systematic Biology,* **54,** 719–730.

Powell, C. McA, Roots, S.R. & Veevers, J.J. (1988) Pre−breakup continental extension in East Gondwanaland and the early opening of the eastern Indian Ocean. *Tectonophysics,* **155,** 261–283.

Rabinowitz, P.D. & Woods, S. (2006) The Africa–Madagascar connection and mammalian migrations. *Journal of African Earth Sciences,* **40,** 270–276.

Rage, J.C. (2003) Relationships of the Malagasy fauna during the Late Cretaceous: northern or southern routes? *Acta Palaeontologica Polonica,* **48,** 661–662.

Rambaut, A. & Drummond, A.J. (2009) Tracer v1.5. Available at: http://beast.bio.ed.ac.uk/Tracer.

Ratnagar, S. (1981) *Encounters: The westerly trade of the Harappa civilization.* Oxford University Press, Delhi.

Raxworthy, C.J., Forstner, M.R.J. & Nussbaum, R.A. (2002) Chameleon radiation by oceanic dispersal. *Nature,* **415,** 784–787.

Ree, R.H. & Smith, S.A. (2008) Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology,* **57,** 4–14.

Ree, R.H., Moore, B.R., Webb, C.O. & Donoghue, M.J. (2005) A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution,* **59,** 2299–2311.

Reynoldson, T.B. (1966) The distribution and abundance of lake−dwelling triclads − towards a hypothesis. *Advances in Ecological Research,* **3,** 1−71.

Robertson, A.H.F. (2000) Mesozoic−Tertiary tectonic−sedimentary evolution of a south Tethyan oceanic basin and its margins in southern Turkey. *Tectonics and Magmatism in Turkey and the Surrounding Area* (ed. by E. Bozkurt, J.A. Winchester, and J.D.A. Piper), pp. 97–138. Geological Society London Special Publications, **173**.

Roelants, K. & Bossuyt, F. (2005) Archaeobatrachian paraphyly and Pangaean diversification of crown-group frogs. *Systematic Biology,* **54,** 111–126.

Ronquist, F., Teslenko, M., Mark, P.V.D., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. (2012) MRBAYES 3.2: Efficient Bayesian phylogenetic inference and model selection across a large model space. *Systematic Biology,* **61,** 539−542.

Sakai, M. & Sakaizumi, M. (2012) The complete mitochondrial genome of *Dugesia japonica* (Platyhelminthes; Order Tricladida). *Zoological Science,* **29,** 672−680.

Samonds, K., Godfrey, L.R., Ali, J.R., Goodman, S.W., Vences, M., Sutherland, M.R., Irwin, M.T. & Krause, D.W. (2012) Spatial and temporal arrival patterns of Madagascar's vertebrate fauna explained by distance, ocean currents, and ancestor

type. *Proceedings of the National Academy of Sciences of the United States of America,* **109,** 5352–5357.

San Mauro, D., Vences, M., Alcobendas, M., Zardoya, R. & Meyer, A. (2005) Initial diversification of living amphibians predated the breakup of Pangaea. *The American Naturalist,* **165,** 590–599.

Sanmartín, I. & Ronquist, F. (2004) Southern hemisphere biogeography inferred by event−based models: Plant versus animal patterns. *Systematic Biology,* **53**, 216–243.

Scarpa, F., Cossu, P., Sanna P., Lai, T., Dedola G.L., Delogu V., Curini−Galleti M. & Casu, M. (2013) Preliminary study for calibrating molecular clock in marine Platyhelminthes using germinate species. *Biologia Marina Mediterranea,* **20,** 248–249.

Schettino, A. & Scotese, C.R. (2005) Apparent polar wander paths for the major continents (200 Ma to the present day): a palaeomagnetic reference frame for global plate tectonic reconstructions. *Geophysical Journal International,* **163,** 727–759.

Scotese, C.R., Bambach, R.K., Barton, C., Van der Voo, R., & Ziegler, A.M. (1979) Paleozoic Base Maps. *The Journal of Geology,* **87,** 217−277.

Scotese, C.R. (2004) Cenozoic and Mesozoic paleogeography: changing terrestrial pathways. Frontiers of Biogeography (ed. by M.V. Lomolino and L.R. Heaney), pp. 9−26, Sinauer Associates.

Sluys, R., Kawakatsu, M., & Ponce de León, R. (2005). Morphological stasis in an old widespread group of species: contribution to the taxonomy and biogeography of the genus Girardia (Platyhelminthes, Tricladida, Paludicola). *Studies on Neotropical Fauna and Environment,* **40,** 155–180.

Sluys, R., Kawakatsu, M. & Winsor, L. (1998) The genus *Dugesia* in Australia, with its phylogenetic analysis and historical biogeography (Platyhelminthes, Tricladida, Dugesiidae). *Zoologica Scripta,* **27**, 273–289.

Sluys, R., Kawakatsu, M. & Yamamoto, K. (2010) Exotic freshwater planarians currently known from Japan. *Belgian Journal of Zoology,* **140(Supp),** 103–109.

Sluys, R., Solà, E., Gritzalis, K., Vila−Farré, M., Mateos, E. & Riutort, M. (2013) Integrative delineation of species of Mediterranean freshwater planarians (Platyhelminthes: Tricladida: Dugesiidae). *Zoological Journal of the Linnean Society,* **169,** 523–547.

Smith, A.G., Smith, D.G. & Funnell, B.M. (2004) *Atlas of Mesozoic and Cenozoic coastlines.* Cambridge University Press, Cambridge.

Solà, E., Sluys, R., Gritzalis, E. & Riutort, M. (2013) Fluvial basin history in the northeastern Mediterranean region underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae). *Molecular Phylogenetics and Evolution,* **66**, 877–888.

Springer, M.S., Murphy, W.J., Eizirik, E. & O'Brien, S.J. (2013) Placental mammal diversification and the Cretaceous–Tertiary boundary. *Proceedings of the National Academy of Sciences USA,* **100,** 1056–1061.

Stamatakis A (2006) RAxML−VI−HPC: maximum likelihood−based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics,* **22,** 2688–2690

Talavera, G. & Castresana, J. (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology,* **56,** 564–577.

Tolley, K.A., Townsend, T.M. & Vences, M. (2013) Large−scale phylogeny of chameleons suggests African origins and Eocene diversification. *Proceedings of the Royal Society B,* **280,** 20130184.

Van Bocxlaer, I., Roelants, K., Biju, S.D., Nagaraju, J. & Bossuyt, F. (2006) Late Cretaceous vicariance in Gondwanan amphibians. *Plos One,* **1,** e74.

Van Hinsbergen, D.J.J., Lippert, P.C., Dupont−Nivet, G., McQuarrite, N., Doubrovine, P.V., Spakman, W. & Torsvik, T.H. (2012) Greater India Basin hypothesis and a two−stage Cenozoic collision between India and Asia. *Proceedings of the National Academy of Sciences USA,* **109,** 7659−7664.

Vences, M., Vieites, D.R., Glaw, F., Brinkmann, H., Kosuch, J., Veith, M. & Meyer, A. (2003) Multiple overseas dispersal in amphibians. *Proceedings of the Royal Society of London Series B, Biological Sciences,* **270,** 2435–2442.

Voris, H.K. (2000) Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography,* **27,** 1153–1167.

Vullo, R., Gheerbrant, E., de Muizon, C. & Néraudeau, D. (2009) The oldest modern therian mammal from Europe and its bearing on stem marsupial paleobiogeography. *Proceedings of the National Academy of Sciences USA,* **106,** 19910−19915.

Wildman, D.E., Uddin, M., Opazo, J.C., Liu, G., Lefort, V., Guindon, S., Gascuel, O., Grossman, L.I., Romero, R. & Goodman, M. (2007) Genomics, biogeography, and the diversification of placental mammals. *Proceedings of the National Academy of Sciences USA,* **104,** 14395–14400.

Wright, H.T. & Rakotoarisoa, J.A. (2003) The Rise of Malagasy Societies: New Developments in the Archaeology of Madagascar. *The Natural History of Madagascar* (ed. by Goodman, S.M. and Benstead, J.P.), p. 112. The University of Chicago Press, Chicago and London.

Xia, X. & Xie, Z. (2001) DAMBE: software package for data analysis in molecular biology and evolution. *Journal of Heredity,* **92,** 371–373.

Yoder, A.D. & Nowak, M.D. (2006) Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell. *Annual Review of Ecology, Evolution, and Systematics,* **37,** 405–31.

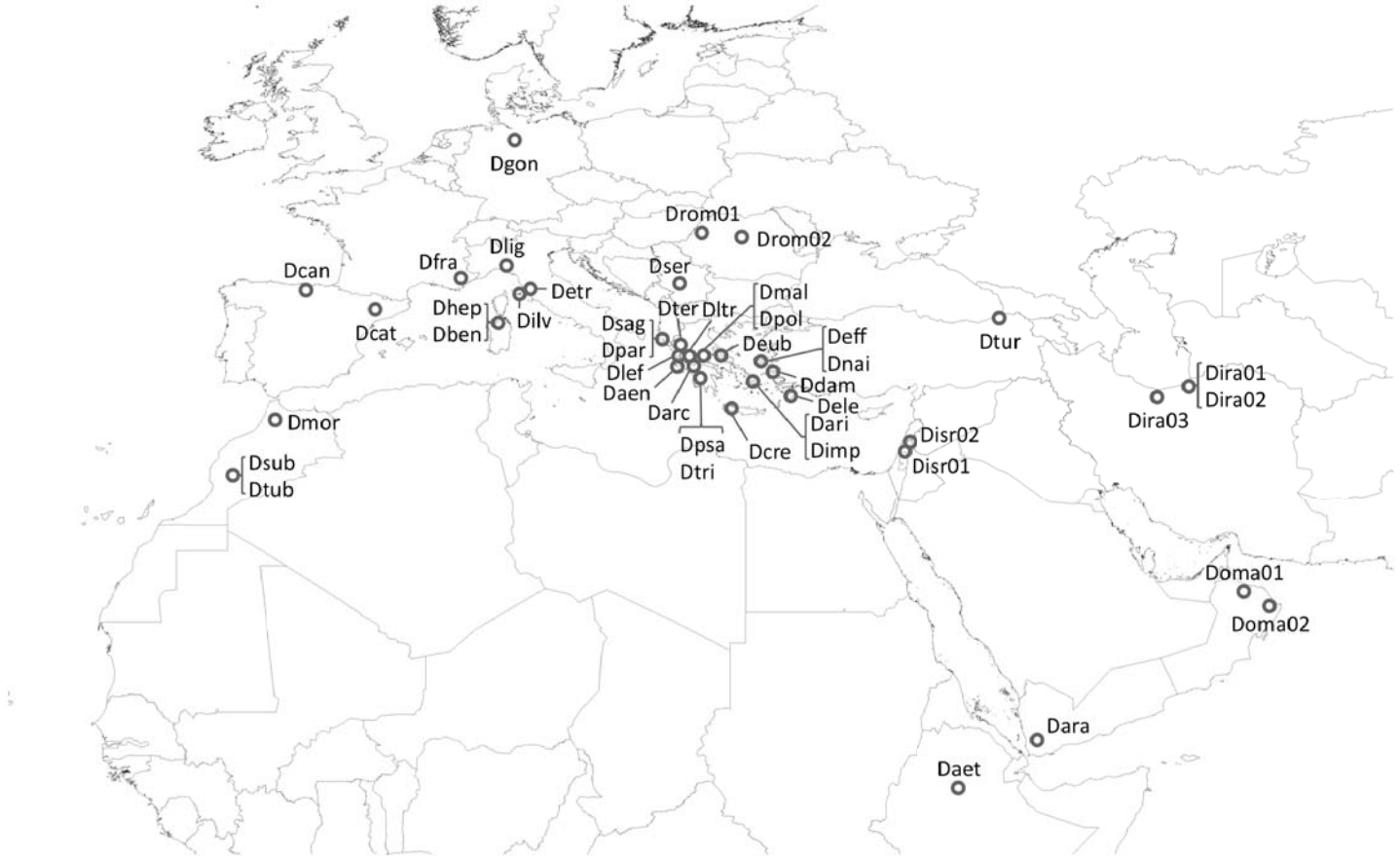
Zhang, H.−C., Sun, J., Chen, G−W. & Liu, D.−Z. (2010) Molecular systematics of eight populations of *Dugesia japonica* in Henan Province based on COI sequences. *Acta Zootaxonomic Sinica,* **35,** 770−775.

Ziegler, P.A. (1988) Evolution of the Arctic–North Atlantic and Western Tethys. *American Association Petroleum Geologists Memoir,* **43**.

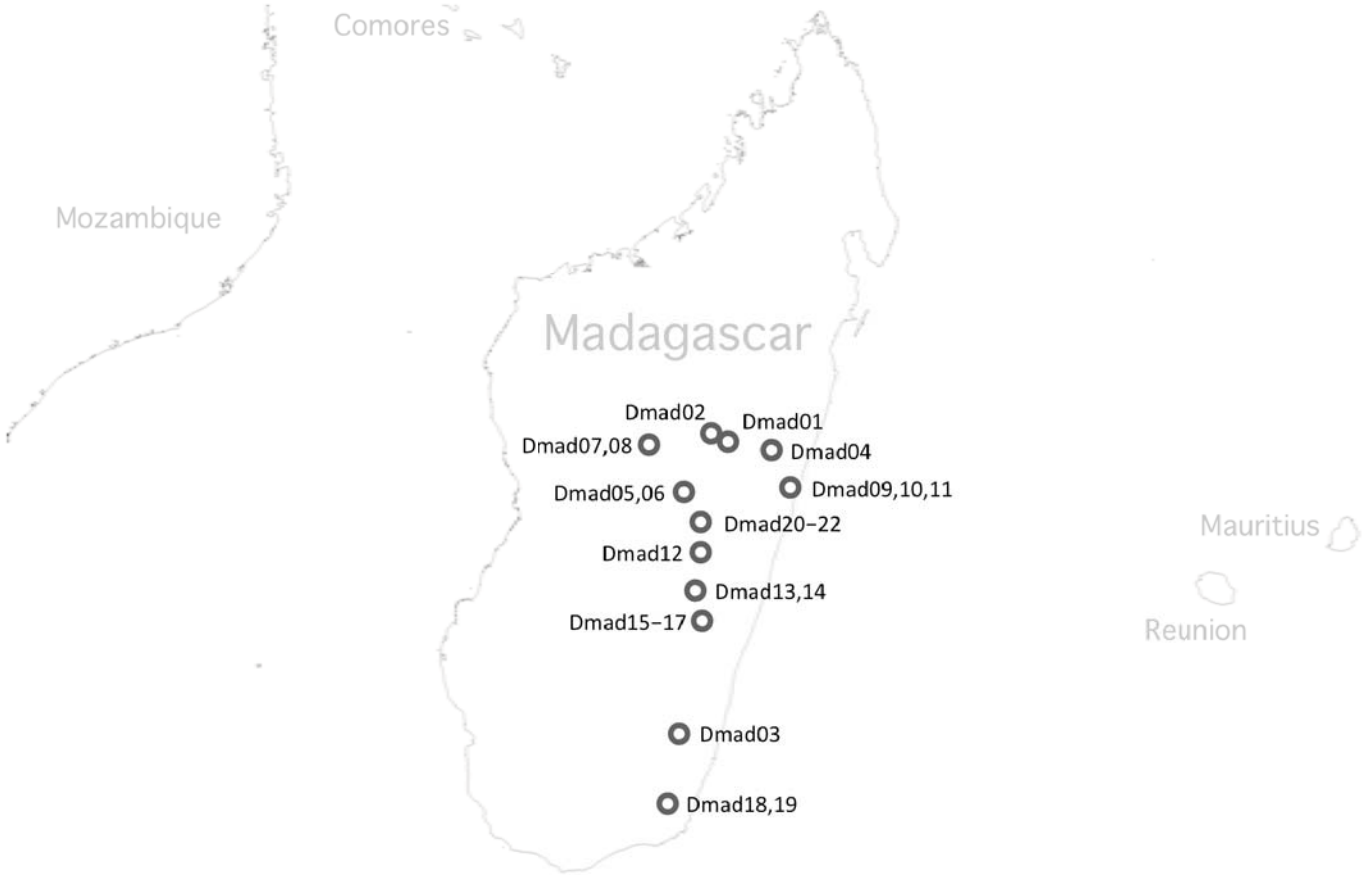
Supplementary Information

Supplementary Figures & Supplementary Tables

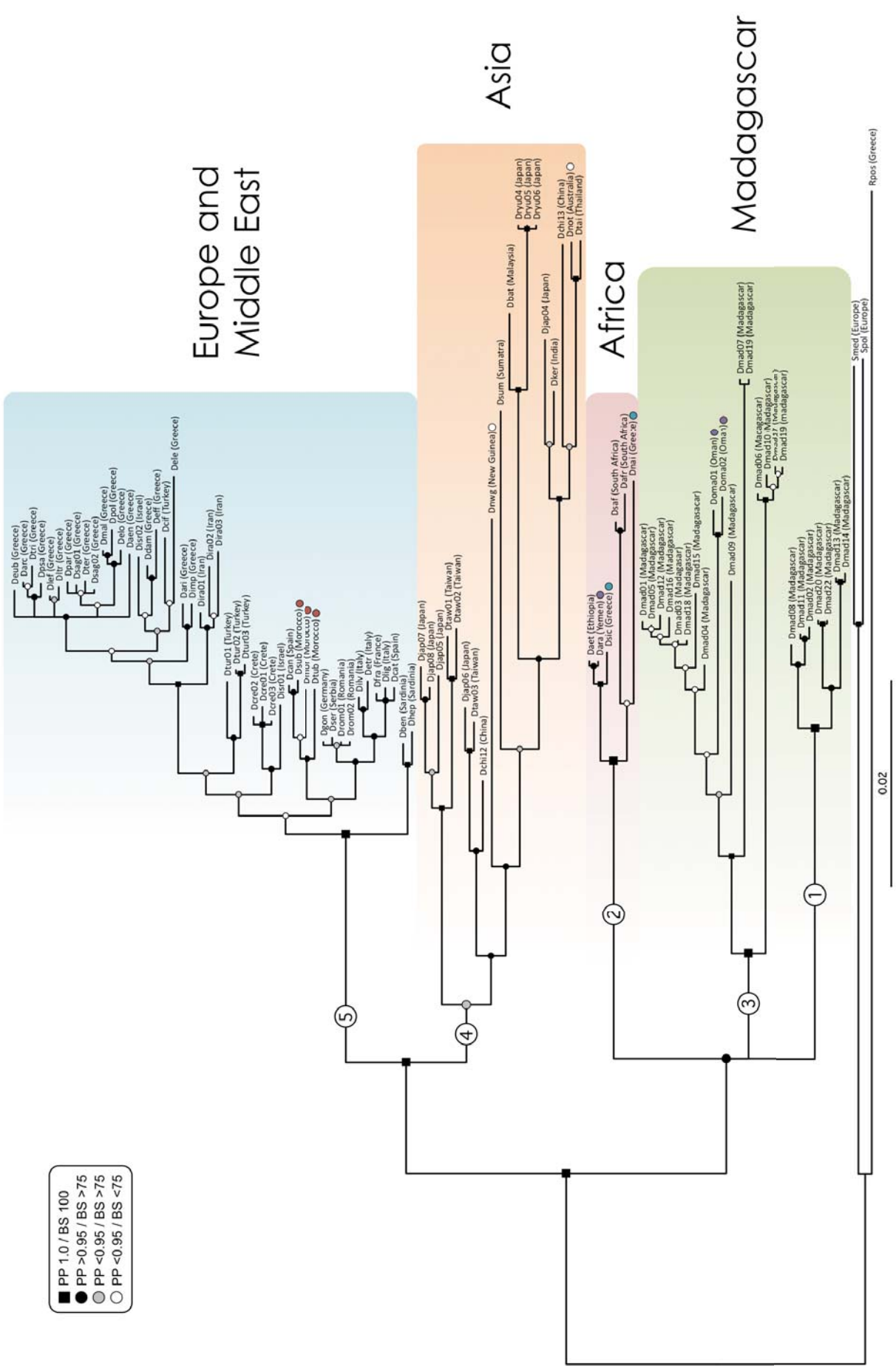
Supplementary Figure 1. Map showing the approximate sampling sites of the *Dugesia* sampled from North Africa, Europe, Middle East and the Arabian Peninsula. Sampling details are included in the Supplementary Table 1.



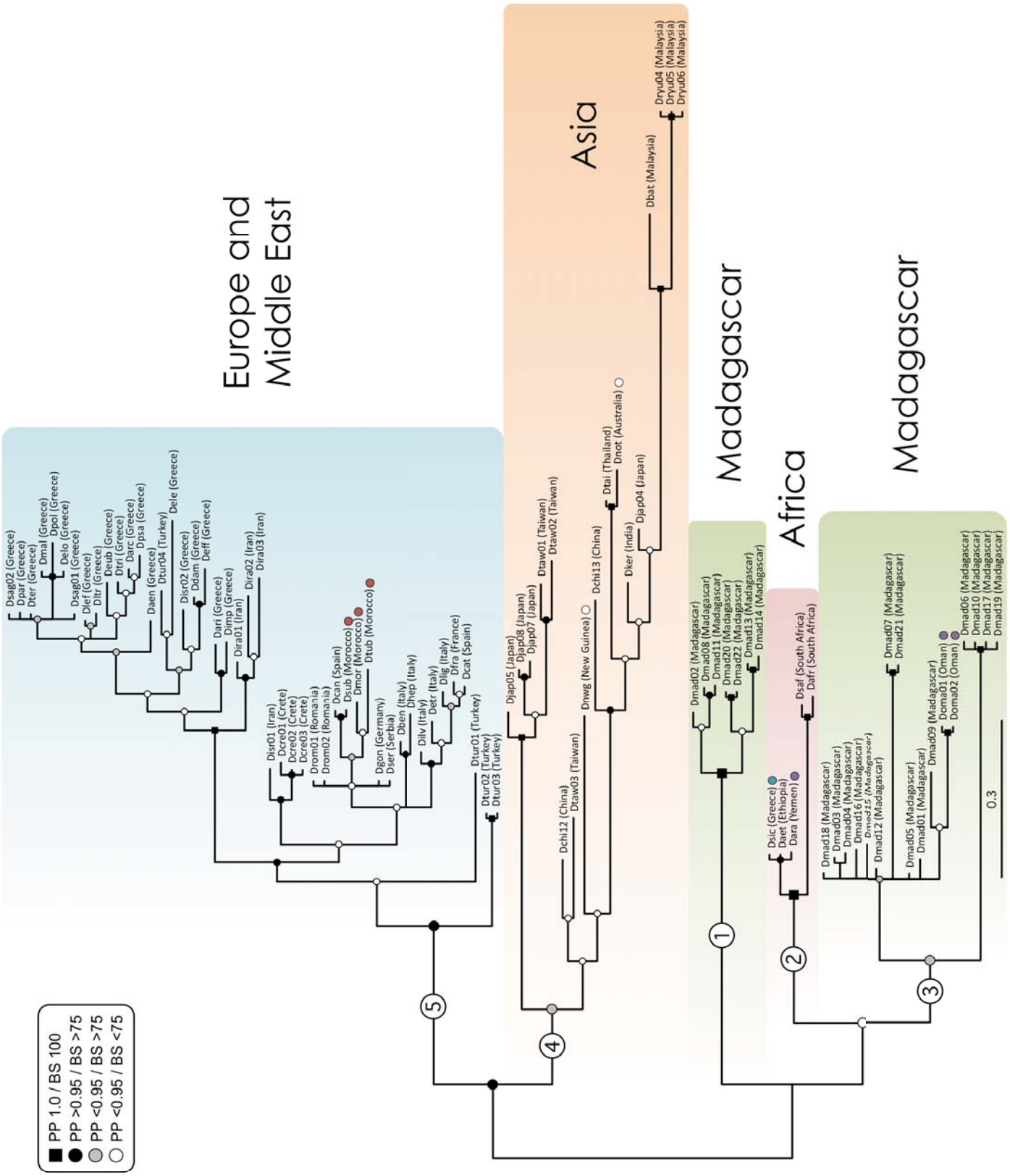
Supplementary Figure 2. Map showing the approximate sampling sites of the *Dugesia* obtained from Madagascar. Sampling details are included in the Supplementary Table 1.



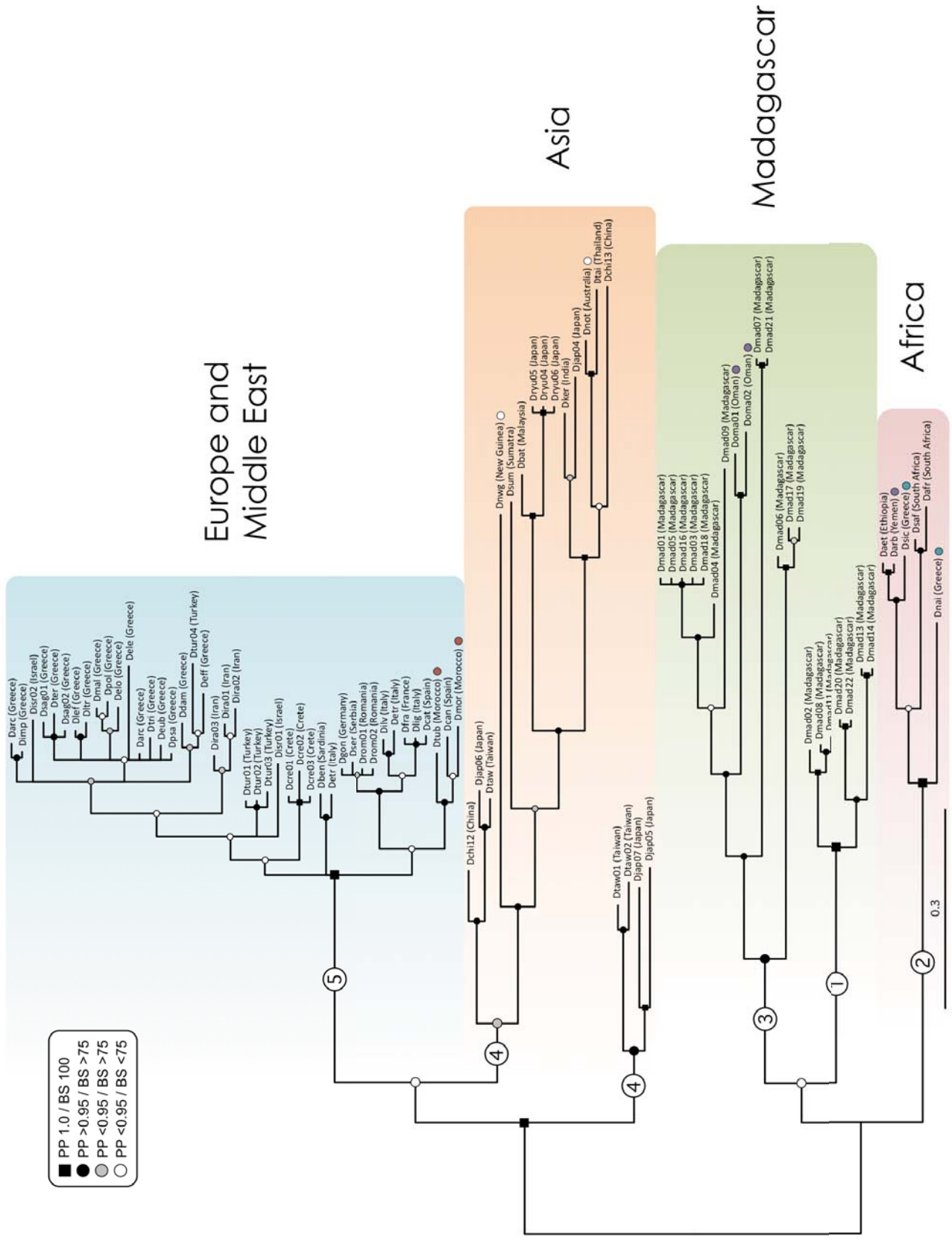
Supplementary Figure 3. Bayesian tree inferred from the 18S and 28S dataset. Rooting using the *Dugesia* sister genera *Recurva* and *Schmidtea*. Taxa labels correspond to codes in Supplementary Table 1. In brackets it is indicated the regions where they were sampled. Numbers in white circles correspond to the main lineages. Specimens that are considered geographical 'outliers' are accompanied with a circle: Red, Africa; White, Australasia; Purple, Arabian Peninsula; Blue, Europe. Supports descriptions are presented within the upper−left box; PP corresponds to the Posterior Probability (BI) and BS to the Bootstrap (ML). The scale bar indicates substitutions per site.


Europe and Middle East
Asia
Africa
Madagascar
PP 1.0 / BS 100
PP >0.95 / BS >75
PP <0.95 / BS >75
PP <0.95 / BS <75
0.02
Deub (Greece)
Darc (Greece)
Dtri (Greece)
Dpsa (Greece)
Dief (Greece)
Dhr (Greece)
Dpar (Greece)
Dsag01 (Greece)
Dter (Greece)
Dsag02 (Greece)
Dmal (Greece)
Dpol (Greece)
Delo (Greece)
Dsta (Greece)
Dsis01 (Israel)
Ddam (Greece)
Deff (Greece)
Dcif (Turkey)
Dele (Greece)
Dari (Greece)
Dimy (Greece)
Dira01 (Iran)
Dira02 (Iran)
Dira03 (Iran)
Dtur01 (Turkey)
Dtur02 (Turkey)
Dtur03 (Turkey)
Dcre02 (Crete)
Dcre01 (Crete)
Dcre03 (Crete)
Disr01 (Israel)
Dcan (Spain)
Dsub (Morocco)
Dmur (Morocco)
Dtub (Morocco)
Dgon (Germany)
Dser (Serbia)
Drom01 (Romania)
Drom02 (Romania)
Dliv (Italy)
Dsar (Italy)
Dfra (France)
Dlig (Italy)
Dcat (Spain)
Dben (Sardinia)
Dhep (Sardinia)
Djap07 (Japan)
Djap08 (Japan)
Djap05 (Japan)
Dtaw01 (Taiwan)
Dtaw02 (Taiwan)
Djap06 (Japan)
Dtaw03 (Taiwan)
Dchi12 (China)
Dnwg (New Guinea)
Dsum (Sumatra)
Dbat (Malaysia)
Dryu04 (Japan)
Dryu05 (Japan)
Dryu06 (Japan)
Djap04 (Japan)
Dker (India)
Dchi13 (China)
Dnot (Australia)
Dtai (Thailand)
Daet (Ethiopia)
Dara (Yemen)
Dsic (Greece)
Dsaf (South Africa)
Dafr (South Africa)
Dnai (Greece)
Dmad01 (Madagascar)
Dmad05 (Madagascar)
Dmad12 (Madagascar)
Dmad16 (Madagascar)
Dmad03 (Madagascar)
Dmad18 (Madagascar)
Dmad04 (Madagascar)
Dmad15 (Madagascar)
Doma01 (Oman)
Doma02 (Oman)
Dmad09 (Madagascar)
Dmad07 (Madagascar)
Dmad19 (Madagascar)
Dmad06 (Madagascar)
Dmad10 (Madagascar)
Dmad17 (Madagascar)
Dmad19 (madagascar)
Dmad08 (Madagascar)
Dmad11 (Madagascar)
Dmad02 (Madagascar)
Dmad20 (Madagascar)
Dmad22 (Madagascar)
Dmad13 (Madagascar)
Dmad14 (Madagascar)
Sned (Europe)
Spol (Europe)
Rpos (Greece)
1
2
3
4
5

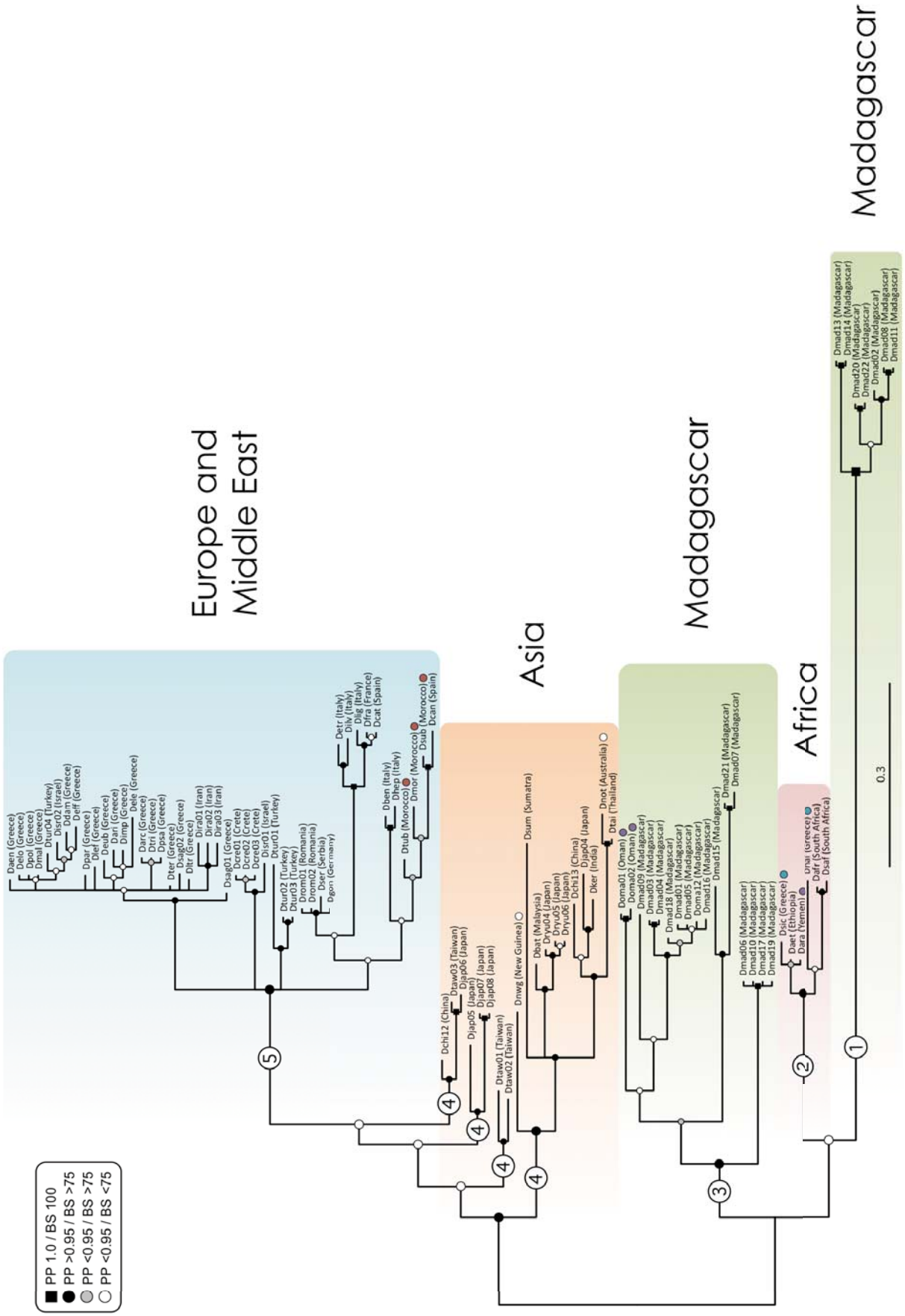
Supplementary Figure 4. Bayesian tree inferred from the 18S dataset. Taxa labels correspond to codes in Supplementary Table 1. In brackets the regions where they were sampled are indicated. Numbers in white circles correspond to the main lineages. Specimens that are considered geographical 'outliers' are accompanied with a circle: Red, Africa; White, Australasia; Purple, Arabian Peninsula; Blue, Europe. Supports descriptions are presented within the upper−left box; PP corresponds to the Posterior Probability (BI) and BS to the Bootstrap (ML). The scale bar indicates substitutions per site.


Europe and Middle East
Asia
Madagascar
Africa
Madagascar
PP 1.0 / BS 100
PP >0.95 / BS >75
PP <0.95 / BS >75
PP <0.95 / BS <75
Dsag02 (Greece)
Dpar (Greece)
Dter (Greece)
Dmal (Greece)
Dpol (Greece)
Delo (Greece)
Dsag01 (Greece)
Dief (Greece)
Deub (Greece)
Dltr (Greece)
Dtri (Greece)
Darc (Greece)
Dpsa (Greece)
Daen (Greece)
Dtur04 (Turkey)
Dele (Greece)
Disr02 (Greece)
Ddam (Greece)
Deff (Greece)
Dari (Greece)
Dimp (Greece)
Dira01 (Iran)
Dira02 (Iran)
Dira03 (Iran)
Disr01 (Iran)
Dcre01 (Crete)
Dcre02 (Crete)
Dcre03 (Crete)
Drom01 (Romania)
Drom02 (Romania)
Dcan (Spain)
Dsub (Morocco)
Dmor (Morocco)
Dtub (Morocco)
Dgon (Germany)
Dser (Serbia)
Dben (Italy)
Dhep (Italy)
Dliv (Italy)
Detr (Italy)
Dlig (Italy)
Dfra (France)
Dcat (Spain)
Dtur01 (Turkey)
Dtur02 (Turkey)
Dtur03 (Turkey)
Djap05 (Japan)
Djap08 (Japan)
Djap07 (Japan)
Dtaw01 (Taiwan)
Dtaw02 (Taiwan)
Dchi12 (China)
Dtaw03 (Taiwan)
Dnwg (New Guinea)
Dchi13 (China)
Dtai (Thailand)
Dnot (Australia)
Dker (India)
Djap04 (Japan)
Dbat (Malaysia)
Dryu04 (Malaysia)
Dryu05 (Malaysia)
Dryu06 (Malaysia)
Dmad02 (Madagascar)
Dmad08 (Madagascar)
Dmad11 (Madagascar)
Dmad20 (Madagascar)
Dmad22 (Madagascar)
Dmad13 (Madagascar)
Dmad14 (Madagascar)
Dsic (Greece)
Dael (Ethiopia)
Dara (Yemen)
Dsaf (South Africa)
Dafr (South Africa)
Dmad18 (Madagascar)
Dmad03 (Madagascar)
Dmad04 (Madagascar)
Dmad16 (Madagascar)
Dmad15 (Madagascar)
Dmad12 (Madagascar)
Dmad07 (Madagascar)
Dmad21 (Madagascar)
Dmad05 (Madagascar)
Dmad01 (Madagascar)
Dmad09 (Madagascar)
Doma01 (Oman)
Doma02 (Oman)
Dmad06 (Madagascar)
Dmad10 (Madagascar)
Dmad17 (Madagascar)
Dmad19 (Madagascar)
0.3
1
2
3
4
5

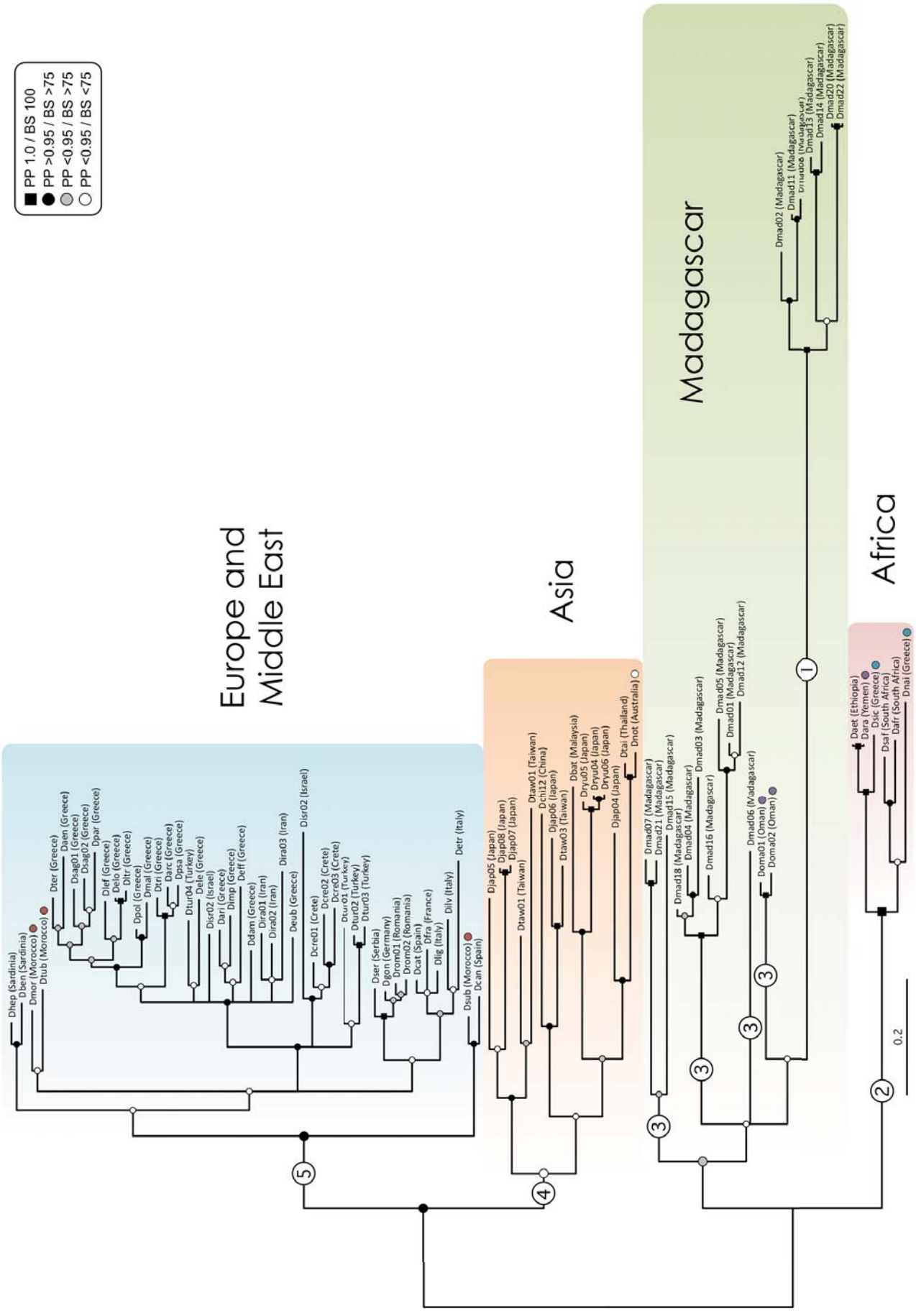
Supplementary Figure 5. Bayesian tree inferred from the 28S dataset. Taxa labels correspond to codes in Supplementary Table 1. In brackets the regions where they were sampled are indicated. Numbers in white circles correspond to the main lineages. Specimens that are considered geographical 'outliers' are accompanied with a circle: Red, Africa; White, Australasia; Purple, Arabian Peninsula; Blue, Europe. Supports descriptions are presented within the upper−left box; PP corresponds to the Posterior Probability (BI) and BS to the Bootstrap (ML). The scale bar indicates substitutions per site.


Europe and Middle East
Asia
Madagascar
Africa
PP 1.0 / BS 100
PP >0.95 / BS >75
PP <0.95 / BS >75
PP <0.95 / BS <75
Darc (Greece)
Dimp (Greece)
Disr02 (Israel)
Dsag01 (Greece)
Dsag02 (Greece)
Dief (Greece)
Ditr (Greece)
Dmal (Greece)
Dpol (Greece)
Delo (Greece)
Dele (Greece)
Darc (Greece)
Dtri (Greece)
Deub (Greece)
Dpsa (Greece)
Ddam (Greece)
Dtur04 (Turkey)
Deff (Greece)
Dira03 (Iran)
Dira01 (Iran)
Dira02 (Iran)
Dtur01 (Turkey)
Dtur02 (Turkey)
Dtur03 (Turkey)
Disr01 (Israel)
Dcre01 (Crete)
Dcre02 (Crete)
Dcre03 (Crete)
Dben (Sardinia)
Detr (Italy)
Dgon (Germany)
Dser (Serbia)
Drom01 (Romania)
Drom02 (Romania)
Dliv (Italy)
Detr (Italy)
Dfra (France)
Dlig (Italy)
Dcat (Spain)
Dtub (Morocco)
Dcan (Spain)
Dmor (Morocco)
Dch12 (China)
Djap06 (Japan)
Dtaw (Taiwan)
Dtaw01 (Taiwan)
Dtaw02 (Taiwan)
Djap07 (Japan)
Djap05 (Japan)
Dnwg (New Guinea)
Dsum (Sumatra)
Dbat (Malaysia)
Dryu05 (Japan)
Dryu04 (Japan)
Dryu06 (Japan)
Dxer (India)
Djap04 (Japan)
Dnot (Australia)
Dtat (Thailand)
Dch13 (China)
Dmad01 (Madagascar)
Dmad05 (Madagascar)
Dmad16 (Madagascar)
Dmad03 (Madagascar)
Dmad18 (Madagascar)
Dmad04 (Madagascar)
Dmad09 (Madagascar)
Doma01 (Oman)
Doma02 (Oman)
Dmad07 (Madagascar)
Dmad21 (Madagascar)
Dmad06 (Madagascar)
Dmad17 (Madagascar)
Dmad19 (Madagascar)
Dmad02 (Madagascar)
Dmad08 (Madagascar)
Dmad11 (Madagascar)
Dmad20 (Madagascar)
Dmad22 (Madagascar)
Dmad13 (Madagascar)
Dmad14 (Madagascar)
Daet (Ethiopia)
Darb (Yemen)
Dsic (Greece)
Dsaf (South Africa)
Dafr (South Africa)
Dnai (Greece)
0.3
1
2
3
4
4
5

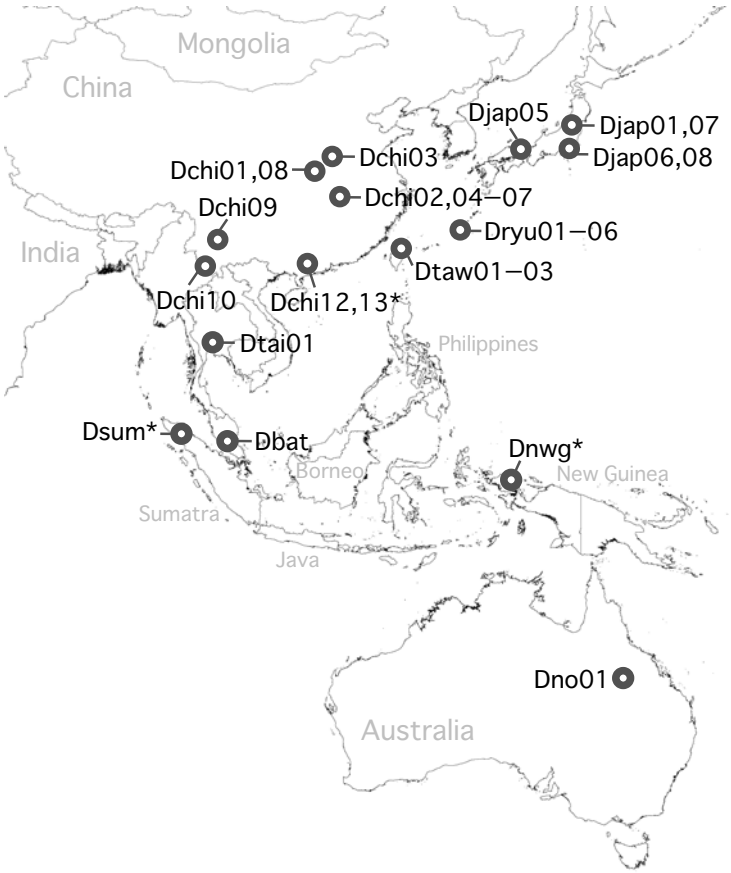
Supplementary Figure 6. Bayesian tree inferred from the ITS−1 dataset. Taxa labels correspond to codes in Supplementary table 1. In brackets the regions where they were sampled are indicated. Numbers in white circles correspond to the main lineages. Specimens that are considered geographical 'outliers' are accompanied with a circle: Red, Africa; White, Australasia; Purple, Arabian Peninsula; Blue, Europe. Supports descriptions are presented within the upper−left box; PP corresponds to the Posterior Probability (BI) and BS to the Bootstrap (ML). The scale bar indicates substitutions per site.


Europe and Middle East
Asia
Madagascar
Africa
Madagascar
PP 1.0 / BS 100
PP >0.95 / BS >75
PP <0.95 / BS >75
PP <0.95 / BS <75
0.3

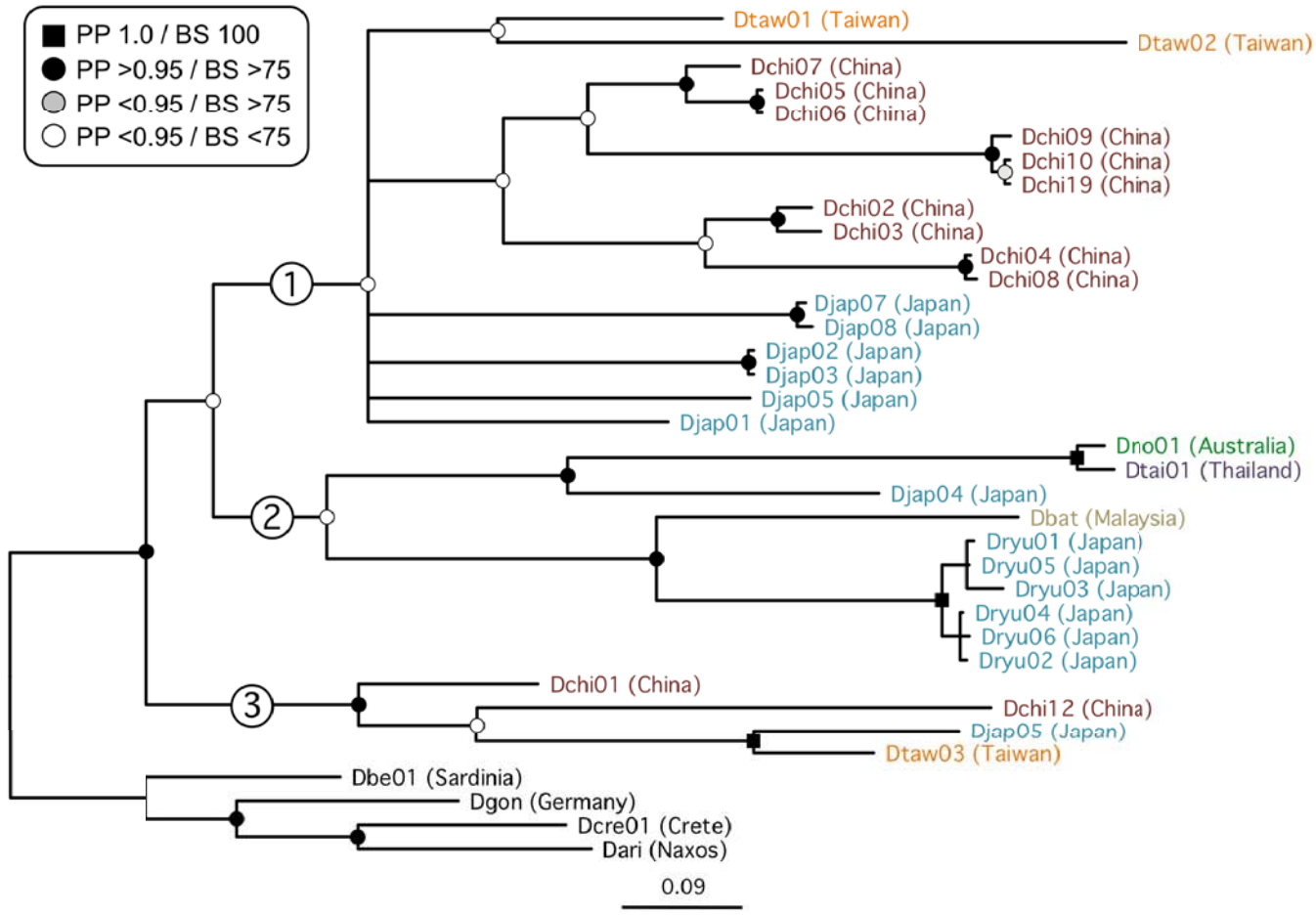
Supplementary Figure 7. Bayesian tree inferred from the Cox1 dataset including the three codon positions. Taxa labels correspond to codes in Supplementary Table 1. In brackets the region where they were sampled are indicated. Numbers in white circles correspond to the main lineages. Specimens that are considered geographical 'outliers' are accompanied with a circle: Red, Africa; White, Australasia; Purple, Arabian Peninsula; Blue, Europe. Supports descriptions are presented within the upper–right box; PP corresponds to the Posterior Probability (BI) and BS to the Bootstrap (ML). The scale bar indicates substitutions per site.


PP 1.0 / BS 100
PP >0.95 / BS >75
PP <0.95 / BS >75
PP <0.95 / BS <75
Europe and Middle East
Asia
Madagascar
Africa
Dhep (Sardinia)
Dben (Sardinia)
Dmor (Morocco)
Dtub (Morocco)
Dter (Greece)
Daen (Greece)
Dsag01 (Greece)
Dsag02 (Greece)
Dpar (Greece)
Dief (Greece)
Delo (Greece)
Ditr (Greece)
Dpol (Greece)
Dmal (Greece)
Dtri (Greece)
Darc (Greece)
Dpsa (Greece)
Dtur04 (Turkey)
Dele (Greece)
Disr02 (Israel)
Dari (Greece)
Dimp (Greece)
Deff (Greece)
Ddam (Greece)
Dira01 (Iran)
Dira02 (Iran)
Deub (Greece)
Dira03 (Iran)
Disr02 (Israel)
Dcre01 (Crete)
Dcre02 (Crete)
Dcre03 (Crete)
Dtur01 (Turkey)
Dtur02 (Turkey)
Dtur03 (Turkey)
Dser (Serbia)
Dgon (Germany)
Drom01 (Romania)
Drom02 (Romania)
Dcat (Spain)
Dfra (France)
Dlig (Italy)
Dliv (Italy)
Detr (Italy)
Dsub (Morocco)
Dcan (Spain)
Djap05 (Japan)
Djap08 (Japan)
Djap07 (Japan)
Dtaw01 (Taiwan)
Dtaw01 (Taiwan)
Dchi12 (China)
Djap06 (Japan)
Dtaw03 (Taiwan)
Dbat (Malaysia)
Dryu05 (Japan)
Dryu04 (Japan)
Dryu06 (Japan)
Djap04 (Japan)
Dtai (Thailand)
Dnot (Australia)
Dmad07 (Madagascar)
Dmad21 (Madagascar)
Dmad15 (Madagascar)
Dmad18 (Madagascar)
Dmad04 (Madagascar)
Dmad03 (Madagascar)
Dmad16 (Madagascar)
Dma05 (Madagascar)
Dmad01 (Madagascar)
Dmad12 (Madagascar)
Dmad06 (Madagascar)
Doma01 (Oman)
Doma02 (Oman)
Dmad02 (Madagascar)
Dmad11 (Madagascar)
Dmad06 (Madagascar)
Dmad13 (Madagascar)
Dmad14 (Madagascar)
Dmad20 (Madagascar)
Dmad22 (Madagascar)
Daet (Ethiopia)
Dara (Yemen)
Dsic (Greece)
Dsaf (South Africa)
Dafr (South Africa)
Dnai (Greece)
0.2
1
2
3
3
3
4
5

Supplementary Figure 8. Map showing the approximate sampling sites of the *Dugesia* shown in the phylogenetic analysis in Supplementary Figure 9. * indicates those localities used in the general work but not included in the Oriental *Dugesia* focused analysis because of the failure in sequencing of the Cox1 gene. We were not able to locate the specimens: Dchi11 (China), Djap02 (Japan), Djap 03 (Japan), and Djap04 (Japan).



Supplementary Figure 9. Bayesian tree inferred from the Cox1 dataset considering all three codon positions. It includes the populations for which there is Cox1 data available from the Far East and Australasia. Taxa labels correspond to codes in Supplementary Table 4 and Supplementary Figure 8. In brackets is shown the region where they were sampled. Numbers in white circles correspond to the main lineages. Supports descriptions are presented within the upper–left box; PP corresponds to the Posterior Probability (BI) and BS to the Bootstrap (ML). The scale bar indicates substitutions per site.



Supplementary Table 1. Specimens, localities, collection information and GenBank Accession Numbers of the genes used in the present study. The specimens in the table are ordered by countries. The two consecutive marks in 18S and 28S columns indicates the availability of the first and second fragment of the gene. ✓, indicates the obtention of the sequences (not uploaded to Genbank yet); ✗, sequence not obtained; −, indicates missing information.

Specimen code	Species	Locality	Collector	Collection date	Accession number			
					cox1	ITS−1	18S	28S
Outgroup								
Rpos	*Recurva postrema*	Greece: Lacrma, Rhodes	EM, ES	04/05/2009	KF308763.1	✓	KF308691.1	✓✓
Smed	*Schmidtea mediterranea*	Spain: Montjuïc, Barcelona	−	−	JF837059.1		U31085.1	✓✓
Spol	*Schmidtea polychroa*	Central Europe	−	−	AF178323.1		AF013154.1	DQ665993.1
Ingroup								
Dnot*	*D. notogaea*	Australia: North Queensland	RC	08/24/2002	FJ646993.1 +FJ646945.1	FJ646908.1	✓✓	✗✓
Dchi12*	*Dugesia* sp.	China: Guandong, Luoding, Silun, Fuchang Village	FL	11/20/2011	✓	✓	✓✓	✓✓
Dchi13*	*Dugesia* sp.	China: Guandong, Luoding, Silun, Oingtong Village	FL	11/21/2010	✗	✓	✓✗	✓✓
Daet*	*D. aethiopica*	Ethiopia: Lake Tana	RM	−	✓	✓	✓✓	✓✓
Dfra	*Dugesia* sp.	France: Alpes Maritimes 3km NW Le Logis−du−Pin, rd N85	IR, JF	2011	✓	✓	✓✓	✓✓
Dgon	*D. gonocephala*	Germany: Niedersachsen, nr Lüder Röhrsen, river Ilmenau	IR, AC	07/10/2012	✓	✓	✓✓	✓✓
Dcre01*	*D. cretica*	Greece: Georgioupoli,	EM, ES	04/07/2009	KC006976.1	KC007050.1	✓✓	✓✓

		Chania, Crete						
Dcre01	*D. cretica*	Greece: Kakopetros, Chania, Crete	EM, ES	04/07/2009	KC006974.1	KC007054.1	✓✓	✓✓
Dcre03	*D. cretica*	Greece: Sasalos, Chania, Crete	EM, ES	04/07/2009	KC006977.1	KC007055.1	✓✓	✓✓
Deff*	*D. effusa*	Greece: Nagos, Chios	MV	04/30/2010	KC006983.1	KC007058.1	✓✓	✓✓
Dari	*D. ariadnae*	Greece: Apollonas, Naxos	EM, ES	04/09/2009	JN376142.1	KC007048.1	✓✓	✓✓
Dimp*	*D. improvisa*	Greece: Melanes, Naxos	EM, ES	04/09/2009	KC006986.1	KC007064.1	✓✓	✓✓
Dele*	*D. elegans*	Greece: Rhodes, Petaloudes Valley	EM, ES	04/04/2009	KC006985.1	KC007062.1	KF308695.1+✓	✓✓
Ddam*	*D. damoae*	Greece: Manolates, Samos	MV	05/05/2010	KC006979.1	KC007057.1	✓✓	✓✓
Deub*	*Dugesia* sp.	Greece: Prokopi, Euboea	CR	05/2010	KC007026.1	KC007112.1	✓✓	✓✓
Dpol	*Dugesia* sp.	Greece: Polidrosos, Phocis	ES, KG	03/14/2010	KC007023.1	KC007094.1	✓✓	✓✓
Dmal*	*D. malickyi*	Greece: Mexiates, Phthiotis	ES, KG	03/29/2010	KC006988.1	KC007068.1	✓✓	✓✗
Dlef	*Dugesia* sp.	Greece: Vafkeri, Lefkada	ES, KG	05/07/2010	KC007034.1	KC007088.1	✓✓	✓✓
Dltr*	*Dugesia* sp.	Greece: Varia, Aetolia–Acarnania	ES, KG	05/07/2010	KC007020.1	KC007092.1	✓✓	✓✓
Delo	*Dugesia* sp.	Greece: Eleonas, Phocis	ES, KG	03/24/2010	KC007014.1	KC007101.1	✓✓	✓✓
Dsag01*	*D. sagitta*	Greece: Roda, Corfu	RS	05/21/2009	KC007006.1	KC007077.1	✓✓	✓✓
Dsag02	*D. sagitta*	Greece: Vouniatades, Corfu	RS	05/24/2009	KC007001.1	KC007079.1	✓✓	✓✓
Dpar	*D. parasagitta*	Greece: Ermones, Corfu	RS	05/24/2009	KC006993.1	KC007073.1	✓✗	✗✗
Dter*	*Dugesia* sp.	Greece: Potamia, Preveza	ES, KG	05/07/2010	KC007036.1	KC007105.1	✓✓	✓✓
Daen*	*D. aenigma*	Greece: Agia Eirini, Cephalonia	RS	05/26/2009	KC006968.1	KC007040.1	KF308698.1+✗	✗✗
Dtri	*Dugesia* sp.	Greece: Tripi, Laconia,	ES, JS, DV	05/13/2010	KC007021.1	KC007106.1	✓✓	✓✓

		Peloponnese						
Darc*	*D. arcadia*	Greece: Sella, Achaea, Peloponnese	ES, KG	03/18/2010	JN376140.1	KC007045.1	KF308694.1+✓	✓✓
Dpsa	*Dugesia* sp.	Greece: Dorio, Messenia, Peloponnese	ES, JS, DV	04/17/2010	KC007019.1	KC007099.1	✓✓	✓✓
Dnai*	*D. naiadis*	Greece: Kimpouries, Chios	MV	05/01/2010	✓	KF308756.1	✗✗	✓✗
Dsic*	*D. sicula*	Greece: Tripes, Chios	MV	05/02/2010	KF308797.1	✓	✓✗	✓✓
Dker*	*Dugesia* sp.	India: South Western Ghats, Kerala	BK	10/2013	✗	✓	✓✓	✓✓
Dsum*	*Dugesia* sp.	Indonesia: Aceh, Sumatra	MB	10/10/2013	✗	✓	✗✗	✓✓
Dnwg*	*Dugesia* sp.	Indonesia: Papua Barat, New Guinea	MB	11/08/2013	✗	✓	✓✓	✓✓
Dira01*	*Dugesia* sp.	Iran: Golestan Province, Qarnabad Forest	AM	06/2012	✓	✓	✓✓	✓✓
Dira01	*Dugesia* sp.	Iran: Golestan Province, Tuskestan Forest	AM	06/2012	✓	✓	✓✓	✓✓
Dira02	*Dugesia* sp.	Iran: Cheshme Ali	RB	2013	✓	✓	✓✓	✓✓
Disr01*	*Dugesia* sp.	Israel: Einot Huga	OS	05/09/2010	✓	✓	✓✗	✓✓
Dirs02*	*Dugesia* sp.	Israel: Dan Springs	OS	08/18/2010	✓	✓	✓✓	✓✓
Dben	*D. benazzii*	Italy: R. Lernu, Sardinia	GS, MP	Lázaro *et al.*, 2009	FJ646977.1 +FJ646933.1	FJ646890.1	✓✗	✓✓
Dilv*	*D. ilvana*	Italy: I. Elba, Tuscany	GS, MP	Lázaro *et al.*, 2009	FJ646989.1 +FJ646944.1	FJ646903.1	✓✓	✓✓
Dlig	*D. liguriensis*	Italy: Grotta Tana di Morbello, Piedmont	GS, MP	Lázaro *et al.*, 2009	FJ646992.1	FJ646907.1	✓✓	✓✓
Dhep*	*D. hepta*	Italy: Sardinia	GS, MP	Lázaro *et al.*, 2009	FJ646988.1 +FJ646943.1	FJ646902.1	✓✓	✓✓
Detr	*D. etrusca*	Italy: Tuscany	GS, MP	Lázaro *et al.*, 2009	FJ646984.1 +FJ646939.1	FJ646898.1	✓✓	✓✓
Dryu04	*D. ryukyuensis*	Japan: Okinawa	MM	1986	✓	✓	✓✓	✓✓
Dryu05	*D. ryukyuensis*	Japan: Okinawa, Oyama	MM	10/08/2011	✓	✓	✓✓	✓✓

Dryu06*	*D. ryukyuensis*	Japan: Okinawa	MM	07/01/2010	✓	✓	✓✓	✓✓
Djap04*	*Dugesia* sp.	Japan: Okinawa, Oyama	MM	01/29/2012	✓	✓	✓✓	✓✓
Djap05*	*Dugesia* sp.	Japan: Hyogo, Iwayadani park	MM	–	✓	✓	✓✓	✓✓
Djap06*	*Dugesia* sp.	Japan: Saitama, Koma river	MM	–	✓	✓	✗✗	✓✗
Djap07	*Dugesia* sp.	Japan: Yamagata, Tateyama	MM	02/28/2013	✓	✓	✓✓	✓✓
Djap08*	*Dugesia* sp.	Japan: Tokyo, Titech	MM	03/07/2013	✓	✓	✓✗	✗✗
Dmad01	*Dugesia* sp.	Madagascar: Lake Mantasoa	RM	10/2011	✓	✓	✓✓	✓✓
Dmad02*	*Dugesia* sp.	Madagascar: Suaranu rivulet	RM	10/2011	✓	✗	✓✗	✓✗
Dmad03	*Dugesia* sp.	Madagascar: Ampasibe	RM	10/2011	✓	✓	✓✓	✓✓
Dmad04	*Dugesia* sp.	Madagascar: Pernit−Tana	RM	10/2011	✓	✓	✓✓	✓✓
Dmad05	*Dugesia* sp.	Madagascar: Tana−Antsirabe	RM	10/2011	✓	✓	✓✓	✓✓
Dmad06*	*Dugesia* sp.	Madagascar: Antsariboti	RM	10/2011	✓	✓	✓✗	✓✓
Dmad07	*Dugesia* sp.	Madagascar: River Landratsay	RM	10/2011	✓	✓	✓✗	✓✓
Dmad08	*Dugesia* sp.	Madagascar: River Landratsay	RM	10/2011	✓	✓	✓✗	✓✓
Dmad09*	*Dugesia* sp.	Madagascar: Ilaka	RM	10/2011	✗	✓	✓✗	✓✓
Dmad10	*Dugesia* sp.	Madagascar: Ilaka	RM	10/2011	✗	✓	✓✗	✗✗
Dmad11	*Dugesia* sp.	Madagascar: Ilaka	RM	10/2011	✓	✓	✓✓	✓✓
Dmad12	*Dugesia* sp.	Madagascar: River Ivato	RM	10/2011	✓	✓	✓✓	✗✗
Dmad13	*Dugesia* sp.	Madagascar: Ambohimahasoa	RM	10/2011	✓	✓	✓✓	✓✓
Dmad14	*Dugesia* sp.	Madagascar: Ambohimahasoa	RM	10/2011	✓	✓	✓✓	✓✓
Dmad15*	*Dugesia* sp.	Madagascar: River Mananatana	RM	10/2011	✓	✓	✓✓	✗✗
Dmad16	*Dugesia* sp.	Madagascar: Lake	RM	10/2011	✓	✓	✓✓	✓✓

		Sahambavy						
Dmad17	*Dugesia* sp.	Madagascar: Lake Sahambavy	RM	10/2011	✗	✓	✓✗	✗✗
Dmad18*	*Dugesia* sp.	Madagascar: River Manaka	RM	10/2011	✓	✓	✓✓	✓✓
Dmad19	*Dugesia* sp.	Madagascar: River Manaka	RM	10/2011	✗	✓	✓✗	✗✗
Dmad20*	*Dugesia* sp.	Madagascar: Ambositra−Fandriana	RM	10/2011	✓	✓	✓✓	✓✓
Dmad21*	*Dugesia* sp.	Madagascar: Ambositra−Fandriana	RM	10/2011	✓	✓	✓✗	✓✓
Dmad22	*Dugesia* sp.	Madagascar: Ambositra−Fandriana	RM	10/2011	✓	✓	✓✓	✓✓
Dbat*	*D. batuensis*	Malaysia: Batu Caves	KT	2013	✓	✓	✓✓	✓✓
Dtub*	*D. tubqalis*	Morocco: Taddert, Toubkal, Marrakech	HH	12/23/2009	✓	✓	✓✓	✓✓
Dsub*	*D. subtentaculata*	Morocco: Imlil, Toubkal, Marrakech	HH	12/23/2009	✓	✓	✓✗	✗✗
Dmor*	*Dugesia* sp.	Morocco	NB	12/24/2009	✓	✓	✓✗	✓✓
Doma01*	*Dugesia* sp.	Oman: Al−Akhdar source of wadi Bani Awf	IR, AC, CH	04/06/2010	✓	✓	✓✓	✓✓
Doma2	*Dugesia* sp.	Oman: Muqal wadi Bani Khalid	IR, AC, CH	04/09/2010	✓	✓	✓✓	✓✓
Drom01	*Dugesia* sp.	Romania: Apuseni Mts., Bratca, Crişul Repede	MKB	07/04/2009	✓	✓	✓✓	✓✓
Drom02	*Dugesia* sp.	Romania: Filia, above Filia Village	MKB	07/18/2009	✓	✓	✓✓	✓✓
Dser01*	*Dugesia* sp.	Serbia: 100km SW from Belgrade	GZ	2012	✓	✓	✓✓	✓✓
Dafr*	*D. afromontana*	South Africa: Amatola Mountains, Eastern Cape Province	RM	–	✓	✓	✓✓	✓✓
Dsaf	*Dugesia* sp.	South Africa: Jonkershoek, Stellenbosch, Western	SD	04/10/2013	✓	✓	✓✓	✓✓

		Cape						
Dcan*	*Dugesia* sp.	Spain: Ruente, Cantabria	MV, LL, ES	10/05/2011	✓	✓	✓✓	✓✓
Dcat*	*Dugesia* sp.	Spain: Font de l'Hedra, Vilanova de Meià, Lleida, Catalonia	MV	05/20/2011	✓	✓	✓✗	✗✓
Dtaw01*	*Dugesia* sp.	Taiwan: Luku	ML	07/24/2012	✓	✓	✓✓	✓✓
Dtaw02	*Dugesia* sp.	Taiwan: Pinglin, Pei−shih streams	ML	07/08/2012	✓	✓	✓✓	✓✓
Dtaw03*	*Dugesia* sp	Taiwan: Caoling Historic Trail, Gongliao	ML	08/20/2012	✓	✓	✓✓	✓✓
Dtai*	*Dugesia* sp	Thailand	RM	12/2012	✓	✓	✓✓	✓✓
Dtur01*	*Dugesia* sp	Turkey: Artvin, Salkimli stream	IR, JF, AF	06/12/2011	✓	✓	✓✓	✓✓
Dtur02*	*Dugesia* sp	Turkey: Artvin, Salkimli stream	IR, JF, AF	06/12/2011	✓	✓	✓✓	✓✓
Dtur03	*Dugesia* sp	Turkey: Artvin, Salkimli stream	IR, JF, AF	06/12/2011	✓	✓	✓✓	✓✓
Dtur04*	*Dugesia* sp	Turkey: Artvin, Çifteköprü stream	IR, JF, AF	06/13/2011	✓	✓	✓✓	✓✓
Dara*	*D. arabica*	Yemen	HH	2011	✓	✓	✓✓	✓✓

Collectors ordered by first name: **AC**, Alexandra Cieslak; **AF**, Arnaud Faille; **AM**, Abdolvahab Maghsoudlou; **BK**, Biju Kumar; **CH**, Carles Hernando; **CR**, Caterina Rodríguez; **DV**, Dora Vázquez; **EM**, Eduardo Mateos; **ES**, Eduard Solà; **FL**, Fengqing Li, **GZ**, Goran Zivanovik; **HH**, Halim Harrath; **IR**, Ignacio Ribera; **JF**, Javier Fresneda; **JS**, Joan Solà; **KG**, Konstantinos Gritzalis; **KT**, Khang Tsung Fei; **LL**, Laia Leria; **MB**, Michael Balke; **MKB**, Miklos Balint; **ML**, Mei−Hui Li; **MM**, Midori Matsumoto; **MP**; Maria Pala; **MV**, Miquel Vila−Farré; **NB**, Núria Bonada; **OS**: Ori Segev; **RB**, Reza Bagherzadeh; **RC**, Ross St. Clair; **RM**, Renata Manconi; **RS**, Ronald Sluys; **SD**, Savel R. Daniels.

* Species used in the DEC analysis implemented in Lagrange (Fig. 4).

Supplementary Table 2.

Forward (F) and reverse (R) primers used in amplification and sequencing. The forward sequences is followed by the corresponding reverse primer. COIEFM, F18SE1, F18SE2, 28SMF1, 28SMR1, 28SMF2 and 28SMR2 were used for Malagasy populations when the other primers failed. JAPO was used for the sequencing (rarely for amplification) of some oriental populations.

Name	Direc.	Sequence 5'−3'	Annealing Temp. (ºC)	Source
ITS−1			45	
9F	F	GTAGGTGAACCTGCGGAAGG		Baguñà et al., 1999
ITSR	R	TGCGTTCAAATTGTCAATGATC		Baguñà et al., 1999
Cox1			43	
BarT	F	ATGACDGCSCATGGTTTAATAATGAT		Álvarez−Presas *et al.*, 2011
COIEF3	F	CCWCGTGCWAATAATTTRAG		Solà et al., 2013
COIEFM	F	GGWGGKTTTGGWAAWTG		This study
JAPO	F	GGWGGYTTTGGTAATTGG		This study
COIR	R	CCWGTYARMCCHCCWAYAGTAAA		Lázaro et al., 2009
18S			45	
1F	F	TACCTGGTTGATCCTGCCAGTAG		Carranza *et al.*, 1996
F18SE1	F	TMTAATCTATTTGCCACAAG		This study
5R	R	CTTGGCAAATGCTTTCGC		Carranza *et al.*, 1996
4F	F	CCAGCAGCCGCGCTAATTC		Carranza *et al.*, 1996
F18SE2	F	GTCGTCGTGTRTATTGTG		This study
9R	R	GATCCTTCCGCAGGTTCACCTAC		Carranza *et al.*, 1996
28S				
28S1F	F	TATCAGTAAGCGGAGGAAAAG	52	Álvarez−Presas *et al.*, 2008
28S4R	R	CCAGCTATCCTGAGGG	49	Álvarez−Presas *et al.*, 2008
28S2F	F	CTGAGTCCGATAGCAAACAAG	49	Álvarez−Presas *et al.*, 2008
28S6R	R	GGAACCCCTTCTCCACTTCAGT	53	Álvarez−Presas *et al.*, 2008
28SMF1	F	GTTGTGTTTTTAATTGAAYAG	43	This study
28SMR1	R	TGCAGACTTTAGATC	43	This study
28SMF2	F	TCTTAATATGYGGTTG	43	This study
28SMR2	R	CTCCACTCTGACTTAC	43	This study

Supplementary Table 3. Dispersal probabilities applied for *Dugesia* in the dispersal−extinction−cladogenesis (DEC) likelihood implemented in LAGRANGE analysis. The migration probabilities among delimited geographic areas (tectonic plates) are shown. The temporal constraints on migration probablities were taken from the paleogeographic reconstructions of this areas position through time available from different sources (See Supplementary Table 5). 0.1 means no contact between landmasses while 1 indicates contiguous landmasses.

Regions	A	B	C	D	E	F	G
20−0 Million years ago (Ma)							
(A) Europe	−	1	1	1	1	0.1	0.1
(B) Asia	1	−	1	1	1	0.1	0.1
(C) Africa	1	1	−	1	1	0.1	0.1
(D) Arabian Peninsula	1	1	1	−	1	0.1	0.1
(E) India	1	1	1	1	−	0.1	0.1
(F) Australia	0.1	0.1	0.1	0.1	0.1	−	
(G) Madagascar	0.1	0.1	0.1	0.1	0.1	0.1	−
25−20 Ma							
(A) Europe	−	1	0.1	0.1	1	0.1	0.1
(B) Asia	1	−	0.1	0.1	1	0.1	0.1
(C) Africa	0.1	0.1	−	1	0.1	0.1	0.1
(D) Arabian Peninsula	0.1	0.1	1	−	0.1	0.1	0.1
(E) India	1	1	0.1	0.1	−	0.1	0.1
(F) Australia	0.1	0.1	0.1	0.1	0.1	−	
(G) Madagascar	0.1	0.1	0.1	0.1	0.1	0.1	−
88−25 Ma							
(A) Europe	−	1	0.1	0.1	0.1	0.1	0.1
(B) Asia	1	−	0.1	0.1	0.1	0.1	0.1
(C) Africa	0.1	0.1	−	1	0.1	0.1	0.1
(D) Arabian Peninsula	0.1	0.1	1	−	0.1	0.1	0.1
(E) India	0.1	0.1	0.1	0.1	−	0.1	0.1
(F) Australia	0.1	0.1	0.1	0.1	0.1	−	
(G) Madagascar	0.1	0.1	0.1	0.1	0.1	0.1	−
130−88 Ma							
(A) Europe	−	1	0.1	0.1	0.1	0.1	0.1
(B) Asia	1	−	0.1	0.1	0.1	0.1	0.1
(C) Africa	0.1	0.1	−	1	0.1	0.1	0.1
(D) Arabian Peninsula	0.1	0.1	1	−	0.1	0.1	0.1
(E) India	0.1	0.1	0.1	0.1	−	0.1	1
(F) Australia	0.1	0.1	0.1	0.1	0.1	−	0.1
(G) Madagascar	0.1	0.1	0.1	0.1	1	0.1	−

132−130 Ma							
(A) Europe	−	1	0.1	0.1	0.1	0.1	0.1
(B) Asia	1	−	0.1	0.1	0.1	0.1	0.1
(C) Africa	0.1	0.1	−	1	1	0.1	1
(D) Arabian Peninsula	0.1	0.1	1	−	1	0.1	1
(E) India	0.1	0.1	1	0.1	−	0.1	1
(F) Australia	0.1	0.1	0.1	0.1	0.1	−	0.1
(G) Madagascar	0.1	0.1	1	1	1	0.1	−
210−132 Ma							
(A) Europe	−	1	0.1	0.1	0.1	0.1	0.1
(B) Asia	1	−	0.1	0.1	0.1	0.1	0.1
(C) Africa	0.1	0.1	−	1	1	1	1
(D) Arabian Peninsula	0.1	0.1	1	−	1	1	1
(E) India	0.1	0.1	1	1	−	1	1
(F) Australia	0.1	0.1	1	1	1	−	1
(G) Madagascar	0.1	0.1	1	1	1	1	−
245.7−210 Ma							
(A) Europe	−	1	1	1	1	1	1
(B) Asia	1	−	1	1	1	1	1
(C) Africa	1	1	−	1	1	1	1
(D) Arabian Peninsula	1	1	1	−	1	1	1
(E) India	1	1	1	1	−	1	1
(F) Australia	1	1	1	1	1	−	1
(G) Madagascar	1	1	1	1	1	1	−

Supplementary Table 4. Specimens, localities and GenBank Accession Numbers (Cox1) of the *Dugesia* specimens used in the Oriental lineages analysis (Supplementary Figs 8 & 9). ✓, indicates those sequences obtained for the present study but not yet uploaded to GenBank.

Code	Species	Locality	Acc. Number	Publication
Outgroup				
Dgon	*D. gonocephala*	Germany: Niedersachsen, nr Lüder Röhrsen, river Ilmenau	✓	This work
Dcre01	*D. cretica*	Greece: Georgioupoli, Chania, Crete	KC006976.1	Solà *et al.*, 2013
Dari	*D. ariadnae*	Greece: Naxos	JN376142.1	Solà *et al.*, 2013
Dben	*D. benazzii*	Sardinia: river Lernu	FJ646977.1+FJ646933.1	Lázaro *et al.*, 2009
Ingroup				
Dno01	*D. notogaea*	Australia: North Queensland	FJ646993.1+FJ646945.1	Lázaro *et al.*, 2009
Dchi01	*D. japonica?*	China: Songxian	GQ503624.1	Zhang *et al.*, 2010
Dchi02	*D. japonica?*	China: Tongbai	GQ503625.1	Zhang *et al.*, 2010
Dchi03	*D. japonica?*	China: Weihui	GQ503626.1	Zhang *et al.*, 2010
Dchi04	*D. japonica?*	China: Gushi	GQ503627.1	Zhang *et al.*, 2010
Dchi05	*D. japonica?*	China: Shangcheng	GQ503628.1	Zhang *et al.*, 2010
Dchi06	*D. japonica?*	China: Shangcheng	GQ503629.1	Zhang *et al.*, 2010
Dchi07	*D. japonica?*	China: Xinxian	GQ503630.1	Zhang *et al.*, 2010
Dchi08	*D. japonica?*	China: Xixian	GQ503631.1	Zhang *et al.*, 2010
Dchi09	*D. japonica?*	China: Nanjian Yi	GU181291.1	Zhang *et al.*, 2010
Dchi10	*D. japonica?*	China: Lancang Lahu	GU181292.1	Zhang *et al.*, 2010
Dchi11	*D. japonica?*	China: Dengjiang	GU181293.1	Zhang *et al.*, 2010
Dchi12	*Dugesia* sp.	China: Guandong, Luoding, Silun, Fuchang Village	✓	This work
Djap01	*D. japonica?*	Japan: Gosen, Niigata	AB618487.1	Sakai & Sakaizumi, 2012
Djap02	*D. japonica?*	Japan	DQ666034.1	Álvarez-Presas *et al.*, 2008
Djap03	*D. japonica?*	Japan	D49916.1	Bessho *et al.*, 1992

Djap04	*Dugesia* sp.	Japan	✓	This work
Djap05	*Dugesia* sp.	Japan: Hyogo, Iwayadani park	✓	This work
Djap06	*Dugesia* sp.	Japan	✓	This work
Djap07	*Dugesia* sp.	Japan	✓	This work
Djap08	*Dugesia* sp.	Japan	✓	This work
Dryu01	*D. ryukyuensis?*	Japan: Naha, Okinawa	AB618488.1	Sakai & Sakaizumi, 2012
Dryu02	*D. ryukyuensis*	Japan: Okinawa	FJ646946.1	Lázaro *et al.*, 2009
Dryu03	*D. ryukyuensis*	Japan	AF178311.1	Lázaro *et al.*, 2009
Dryu04	*D. ryukyuensis*	Japan: Okinawa	✓	This work
Dryu05	*D. ryukyuensis*	Japan: Okinawa, Oyama	✓	This work
Dryu06	*D. ryukyuensis*	Japan: Okinawa	✓	This work
Dbat	*D. batuensis*	Malaysia: Batu caves	✓	This work
Dtai01	*Dugesia* sp.	Thailand	✓	This work
Dtaw01	*Dugesia* sp.	Taiwan: LuKu	✓	This work
Dtaw02	*Dugesia* sp.	Taiwan: Pinglin, Pei−shih streams	✓	This work
Dtaw03	*Dugesia* sp.	Taiwan: Caoling Historic Trail, Gongliao	✓	This work

Supplementary Table 5. Information on the temporal constraints applied for *Dugesia* migration probabilities among plates/areas constructed based on paleogeographic reconstructions. These constraints were implemented in a series of seven time slices as listed in Supplementary Table 3.

Paleogeographical Event	Areas impacted	Age	References
Extension of the Tethys Sea westwards forming the European epicontinental seaway (EES).	(Europe, Asia) − (Africa, Arabian Peninsula, Madagascar, India, Australia)	210 Ma (220–200 Ma)	Ziegler, 1988; Newton & Bottrell, 2007.
Split of Australia from the rest of Gondwana	(Africa, Arabian Peninsula, Madagascar, India) − (Australia)	132 Ma	Powell *et al.*, 1988; Müller *et al.*, 2000; Brown *et al.*, 2003
Split of India and Madagascar from Africa	(Africa) − (Madagascar, India)	130 Ma	Schettino & Scottese, 2005; Ali & Aitchison, 2008 Rabinowitz & Woods, 2006
Isolation of Madagascar	(Madagascar) − (India)	88 Ma	Besse & Courtillot, 1988, 2002; Acton, 1999
Impact of India with Eurasia	(India, Europe, Asia)	25 Ma*	Aitchison *et al.*, 2007; Ali & Aitchison, 2008; Van Hinsbergen *et al.*, 2012
Impact of the Arabian Plate with Eurasia	(Africa, Arabian Peninsula, Europe, Asia, India)	20 Ma	Robertson, 2000

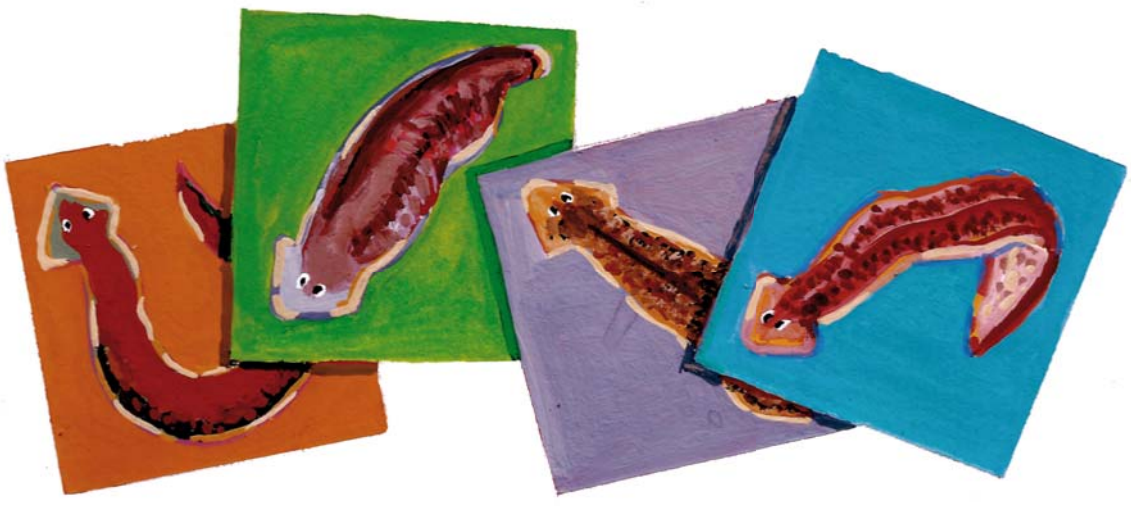
*For the impact of India we took an intermediate value between the different proposed until the present.

Chapter 3

Integrative delineation of species of Mediterranean freshwater planarians (Platyhelminthes: Tricladida: Dugesiidae)

Reference

Sluys R, Solà E, Gritzalis K, Vila-Farré M, Mateos E, Riutort M (2013). **Integrative delineation of species of Mediterranean freshwater planarians (Platyhelminthes: Tricladida: Dugesiidae).** *Zool J Linn Soc* **169**:523−547.



Summary

Dugesia morphological description is a troublesome process due to many different methodological and description difficulties. With the following paper we aimed to carry out a species delimitation approach of dugesiid specimens from both morphological and molecular perspectives. In order to achieve this challenge we sequenced a fragment of the mitochondrial gene Cox1 for some specimens from different localities across Greece. Most of these localities were also included in the biogeographical studies also contained in this thesis (Chapter 1). We used the molecular-based species delimitation General Mixed Yule-Coalescent (GMYC), a method that presents the advantatge to be based only on one mitochondrial locus instead of many loci. Although it tends to oversplit lineages, it is a proper approach for biodiversity surveys including groups of organisms for which there is scarce previous taxonomic information.

Six new species of Dugesiidae were described on the basis of morphology and supported by molecular data. Four of these species belong to the genus *Dugesia*, while two belong to a new genus that we named *Recurva*. A part from the formally described species, we also found two populations that are probably new species. This proposal is based on both molecular data and morphological incomplete but conclusive evidences. We provisionaly considered them as Confirmed Candidate Species, following the categories proposed by Vieites and collaborators (2009) for biological entities for which not all information (but molecular) is gathered to describe formally a new species. Finally, many molecularly delimitated specimens were considered as Unconfirmed Candidate Species and pointed for future morphological analyses.

From this work the diversity of the group in the region was notably increased and the molecular-based method of species delimitation GMYC has been assessed as appropriate for freshwater flatworms. Now the Aegean region is known to harbor many Dugesiidae species, probably as a consequence of its complex geological history.

The LINNEAN SOCIETY of London ZOOLOGICAL Journal of the Linnean Society

Zoological Journal of the Linnean Society, 2013, **169**, 523–547. With 18 figures

Integrative delineation of species of Mediterranean freshwater planarians (Platyhelminthes: Tricladida: Dugesiidae)

RONALD SLUYS[1,2]*, EDUARD SOLÀ[3], KONSTANTINOS GRITZALIS[4], MIQUEL VILA-FARRÉ[3,5], EDUARDO MATEOS[3] and MARTA RIUTORT[3]

[1]*Naturalis Biodiversity Center, PO Box 9514, 2300 RA Leiden, The Netherlands*
[2]*Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, The Netherlands*
[3]*Departament de Genètica, Facultat de Biologia and Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Catalonia, Spain*
[4]*Hellenic Centre for Marine Research, Institute of Inland Waters, Anavyssos, Greece*
[5]*Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany*

Received 10 April 2013; revised 1 August 2013; accepted for publication 1 August 2013

The paper presents an integrative taxonomic study on dugesiid freshwater flatworms from the north-eastern Mediterranean region by applying both morphological and molecular criteria in the formulation of stable species hypotheses. The morphological information obtained for the specimens was used in a traditional way by comparing the organismal traits of the various populations and candidate species with those of known species, as documented in the taxonomic literature and as revealed by examination of histological sections of museum specimens. In the molecular species delimitation the General Mixed Yule-Coalescent method (GMYC) was used. Results of this study (1) supported the presence of 13 *Dugesia* species in the Hellenic area (including *D. sicula* Lepori, 1948, a pan-Mediterranean species), (2) culminated in the description of four new *Dugesia* species, (3) suggested the presence of two Confirmed Candidate Species, (4) pointed to 12 GMYC-delimited units in Greece and two in Slovakia as Unconfirmed Candidate Species and (5) revealed the presence of an entirely new genus, represented by two newly described species and a third Unconfirmed Candidate Species. Our results revealed a high diversity of dugesiid species in this relatively small region. It is concluded that the morphological features used by taxonomists in comparative studies of dugesiid flatworms generally result in reliable identifications and delineations of species taxa, except in the case of cryptic species.

doi: 10.1111/zoj.12077

ADDITIONAL KEYWORDS: Aegean – candidate species – cryptic species – *Dugesia* – Dugesiidae – GMYC – integrative taxonomy – ***Recurva*** Sluys **gen. nov.** – species delimitation.

INTRODUCTION

The freshwater planarian genus *Dugesia* Girard, 1850 (Platyhelminthes, Tricladida, Dugesiidae) currently comprises about 80 nominal species that are distributed in the Afrotropical, Palearctic, Oriental and Australian biogeographical regions (cf. Sluys, Kawakatsu & Winsor, 1998). More than 20 species occur in Europe, particularly in the Mediterranean region. Generally, identification of species of *Dugesia* is difficult because they are externally very similar. The traditional source of taxonomic characters concerns features of their reproductive complex, notably their copulatory apparatus. But even in their reproductive system species may be very similar, making proper identification a time-consuming and painstaking enterprise, in addition to the fact that the

*Corresponding author. E-mail: ronald.sluys@naturalis.nl

necessary taxonomic characters can be observed only in histological sections. Another complication with identification concerns the fact that many Mediterranean populations reproduce asexually by fission and usually do not develop a copulatory apparatus, thus preventing taxonomic assignment to a particular known species or to a new species.

In our view, therefore, the genus *Dugesia* represents a highly suitable model group to explore an integrative approach to delimiting species. For this, we have obtained both molecular and morphological information for a large number of *Dugesia* and other dugesiid populations distributed in the eastern Mediterranean region that we used as data sources to formulate and test species boundary hypotheses. The phylogeographical history of most of these populations has been analysed in a companion paper (Solà *et al*., 2013).

A consensus is emerging that species are segments of separately evolving lineages of populations (cf. De Queiroz, 2007; Frankham *et al*., 2012), albeit that the problem remains of establishing where during this process the diverging groups reach species status. In some cases morphological, behavioural or ecological differences represent unequivocal signals that speciation has occurred. In other cases, only analyses based on population genetics and coalescent theory suggest lack of gene flow, thus evidencing the presence of cryptic species (Bickford *et al*., 2007; Fontaneto *et al*., 2007; Burbrink *et al*., 2011; cf. Olson, Goodman & Yoder, 2004; Carew, Pettigrove & Hoffmann, 2005; Vieites *et al*., 2009; Fujita *et al*., 2012). Genetic distances (cf. Memon *et al*., 2006; Fouquet *et al*., 2007; Vieites *et al*., 2009) and other non-coalescent molecular-based species delimitation methods are not suitable for species delimitation because they rely on highly subjective criteria (Hey, 2009).

Our methodology in species delimitation consisted of three main steps. First, hypotheses on *candidate* species were formulated based on the examination of morphological features. Second, agreements and divergences between these candidate species and *putative* species delineated by a coalescent-based molecular method were identified. Third, during an iterative process reciprocal illumination of morphological and molecular results eventually resulted in the formulation of stable species hypotheses.

The morphological information obtained for the dugesiid flatworms was used in a traditional way by comparing the organismal traits of the various populations and candidate species with those of known species, as documented in the taxonomic literature and as revealed by our examination of histological sections of relevant museum specimens. Conformity of the relevant characters with those of known species enabled taxonomic assignment of the populations sampled, while divergences of organismal attributes suggested the presence of a new species.

As our species concept for the delimitation of candidate species we have chosen the phylogenetic species concept as formulated by Cracraft (1983, 1987; see also Sluys, 1991). In practice this means that a species boundary is hypothesized when a population of organisms is characterized by the presence of one or more unique characters or by a unique combination of characters, each of which may be plesiomorphic. For Mediterranean dugesiids morphological characters were used for postulating such phylogenetic species hypotheses, which were compared with the molecular, coalescent-based delineations of putative species. In other words, our delimitation criterion for candidate species status was morphological diagnosability or distinctness, with many of the characters being derived from the reproductive system.

As our molecular species delimitation method we have applied the Yule-Coalescent transition analysis as implemented in the General Mixed Yule-Coalescent (GMYC) method (Pons *et al*., 2006; Fontaneto *et al*., 2007), using cytochrome oxidase I (COI) sequences. It has been shown that Yule-Coalescent model analysis with a single mitochondrial gene can be a meaningful and rapid approach to assess species diversity within a group of organisms (Monaghan *et al*., 2009; Talavera, 2012). This coalescent-based method allows species delimitation by distinguishing branching patterns between interspecific (Yule model; speciation and extinction) and intraspecific (coalescence of alleles) processes on a phylogenetic tree. It draws a threshold between these two processes, thus delimiting clades of individuals representing putative species. It is useful even in situations (a) with high numbers of singletons, (b) with low taxon level (3–5 species) or (c) without intraspecific coverage (Talavera, 2012). The efficiency of the GMYC method is mostly due to the fast evolving nature of the mitochondrial genes, which are presumed to coalesce faster than nuclear genes because of their smaller effective population sizes (Moore, 1995; Avise, 2000). However, there are some drawbacks in using only one marker. For example, one runs the risk of equating the gene tree with the species tree, and consequently cases of reticulated evolution, introgression or incomplete lineage sorting can mislead phylogenies and thus lead to incorrect species hypotheses (cf. Edwards, 2009; Lohse, 2009).

Our study revealed also the presence of a new dugesiid genus. The erection and description of this new genus is based on the presence of differential morphological traits and on a phylogenetic analysis of 18S rRNA and COI gene sequences. The resulting phylogenetic tree clearly demonstrates that the new

genus consitutes a monophyletic lineage separate from all other dugesiid genera.

Evidently, in any integrative study there may be a discordance between morphologically determined or candidate species taxa on the one hand and putative molecular species on the other hand. Although discordance between morphological and molecular data may be a nuisance from a taxonomic perspective, it is interesting from a biological or evolutionary point of view (Yeates *et al*., 2011). Discordances cannot always be resolved. For situations in which not all data coincide or just one kind of data is available, Vieites *et al*. (2009) proposed three different categories to describe the taxonomic status of the biological units under study. The first category concerns Unconfirmed Candidate Species (UCS), including those genealogical lineages that can be delineated by a molecular method but for which other data are not available. The second is the Confirmed Candidate Species (CCS), comprising those units that can be delimited by molecular data and are supported also by other data, such as morphology, but have not yet been formally described and named. The third category concerns Deep Conspecific Lineages (DCL), referring to lineages that have reached a certain molecular threshold but present the same or a very similar morphology. We have applied this system to indicate the taxonomic status of those biological units that do not have the status of described species (DS).

In this study we will not describe new species solely on the basis of molecular divergence and in the absence of morphological species markers. In this way we avoid the danger of overestimating the number of species as a consequence of possible oversplitting by the GMYC method, although we run the risk of overlooking morphologically cryptic species. We have chosen this taxonomic practice in view of (1) compatibility with past taxonomic practice, and (2) the situation that the current International Code of Zoological Nomenclature (ICZN, 1999) requires the description of a new species taxon to be accompanied with a description that clearly differentiates the taxon (see also Bauer *et al*., 2011), and by the deposition of type specimen(s). Formally, molecular data may be presented in a way that fulfils the requirements of the ICZN (1999) and resembles traditional descriptions (cf. Nygren & Pleijel, 2011). However, in our view a DNA barcode does not provide the in-depth information on organismal divergence that allows one to formulate and test scientifically interesting hypotheses on the evolution of structures, adaptations, functional morphology, life history and behaviour (Sluys, 2013). Therefore, here we refrain from describing new species solely on the basis of their DNA barcode.

MATERIAL AND METHODS

Collection of specimens

Freshwater planarians were collected from the type localities of eight Greek *Dugesia* species (cf. De Vries, 1984, 1988) and from other localities on the mainland as well as some islands during the spring seasons of 2009 and 2010 (cf. Solà *et al*., 2013). All individuals used in the molecular analyses, as well as information on their sampling localities, are listed in Supporting Tables S1 and S2. Specimens used for morphological studies are listed in the relevant Material Examined sections of the Systematic and Integrative Section and/or are deposited in the collections of the Naturalis Biodiversity Center, Leiden, the Netherlands.

Morphological analysis and species hypotheses

Animals for morphological studies were fixed in Steinmann's fluid and, subsequently, transferred to 70% ethanol. Specimens that had been preserved for anatomical analysis were cleared in clove oil and then embedded in paraffin wax, sectioned at intervals of 6 or 8 μm (depending on the size of the animals) and mounted on albumen-coated slides. Sections were stained in Mallory-Cason/Heidenhain (Humason, 1967; Romeis, 1989) and mounted in DPX. Reconstructions of the copulatory complex were obtained by using a camera lucida attached to a compound microscope. All material has been deposited in the collections of the Naturalis Biodiversity Center, Leiden, the Netherlands.

The species status of the animals from the various localities was assessed by applying the phylogenetic species concept as formulated by Cracraft (1983, 1987; see also Sluys, 1991) and by comparing qualitative features of their reproductive complex, in particular their copulatory apparatus, with those of known species, as documented in the taxonomic literature and revealed by examination of histological sections of relevant museum specimens housed in the collections of the Naturalis Biodiversity Center. Detailed discussions of relevant characters used to differentiate the new species are presented in the Systematic and Integrative Section. Conformity of the relevant characters with those of known species enabled taxonomic assignment of the populations sampled, while divergences of organismal attributes suggested the presence of a candidate new species.

DNA sequencing and alignment

In addition to the mitochondrial COI sequences obtained for a companion phylogeographical study (Solà *et al*., 2013), sequences of 1–3 individuals per locality were obtained, when possible, and included in

the alignments (Table S1), following the same procedure described in that paper. Furthermore, 18S rDNA nuclear gene sequences (18S) were obtained for 11 individuals (Table S4). Sequences and annealing temperatures for each pair of primers, both for COI and for 18S, are given in Table S3. 18S was aligned by using online software MAFFT, version 6 (Katoh & Toh, 2008), while ambiguous positions were removed with the program GBlocks with default settings, except the minimum number of sequences for a conserved position (set at 16) and with half of the allowed gap positions (Talavera & Castresana, 2007). The level of sequence saturation for COI sequences of different genera was analysed under the TN93 nucleotide substitution pattern model with the program DAMBE (Xia & Xie, 2001). The three positions were analysed at the same time and independently.

Phylogenetic analyses of dugesiid genera

To analyse the genetic differentiation of a candidate new genus, which we happened to encounter among our Greek material, as well as to determine its relationship to other European members of the Dugesiidae and also to the Australian species *Cura pinguis* (Weiss, 1909) (which shares some morphological similarities with the new genus), we performed phylogenetic analyses using two datasets. One dataset consisted of a concatenated set including 18S and COI. The second dataset concerned only COI because this enabled us to include *Cura pinguis*, for which 18S sequences are not available. In the concatenated analysis we compared 21 species of five genera by taking one specimen of each (Table S4). This dataset lacks the 18S for *Dugesia naiadis* Sluys sp. nov. and the COI for *Recurva conjuncta* Sluys sp. nov. because we were unable to amplify these sequences. In the COI analysis we also compared 21 species, but excluded *Recurva conjuncta* and included *Cura pinguis*. The land flatworm species *Bipalium adventitium* Hyman, 1943 (Tricladida, Geoplanidae, Bipaliinae) was used as outgroup.

All phylogenetic analyses were performed using two inference methods, namely maximum-likelihood (ML) and Bayesian inference (BI). We used jModelTest 2.1.1 (Darriba *et al*., 2012) to test which evolutionary model fitted best with our data. We used GTR + I + Γ for 18S and HKY + I + Γ for COI, excluding third positions, and set the parameter estimation as unlinked among genes in the concatenated analysis. ML analysis was run with the program RaxML 7.0.0 (Stamatakis, 2006). To obtain bootstrap support (BS), 1000 replicates were calculated. We used MrBayes (v. 3.2: Ronquist *et al*., 2012) to perform the BI analysis. In total, 1000 000 generations were run, saving a tree every 100 generations. Convergence of topologies and model parameters of both runs was surveyed by checking whether the standard deviation of the split frequencies reached a value below 0.01 (default burn-in = 25%). We also checked that likelihood values had stabilized by plotting them against the number of generations. To infer the topology and posterior probability (PP) values we used the default burn-in.

Molecular species delimitation of *Dugesia* populations

We performed a GMYC approach (Pons *et al*., 2006; Fontaneto *et al*., 2007) to compare the units delimited by this method with those identified in the morphological analysis and to detect possible cryptic species. We used the partial COI sequences of 155 individuals of *Dugesia* from 34 localities (Table S1). GMYC detects the change from population processes (coalescence of alleles) to speciation and extinction processes through analysis of branching rate patterns, setting a threshold between the inter- and intraspecific relationships. To obtain the ultrametric tree necessary for this approach, we conducted a phylogenetic analysis in BEAST v1.7.3 (Drummond & Rambaut, 2007), using a fragment of COI (745 bp) from 2–5 individuals per sampling locality (Table S1). A lognormal relaxed clock with a substitution rate of 0.017 substitutions per lineage and per million years was applied (cf. Solà *et al*., 2013). The analysis was run under a GTR + I + Γ evolutionary model. Three monophyletic clades were forced: (1) *Dugesia* species, without *D. sicula* and *D. naiadis* (used as outgroup); (2) *Dugesia* species, without *D. sicula*, *D. naiadis* and *Dugesia* from Central Europe; (3) *Dugesia* species, without *D. sicula* Lepori, 1948, *D. naiadis* Sluys sp. nov., *Dugesia* from Central Europe and *D. cretica* (Meixner, 1928; Solà *et al*., 2013). Monte Carlo Markov chains were run for 150 000 000 generations, sampling every 15 000 trees. The parameters were checked to have reached an effective sampling size (ESS) value of over 100 after a 10% burn-in with Tracer v.1.5 (Rambaut & Drummond, 2007).

The BEAST tree obtained was submitted to the SPLITS (SPecies LImits by Threshold Statistics; Ezard, Fujisawa & Barraclough, 2009) package for R (available at http://r-forge.r-project.org/projects/splits/), which implements the GMYC approach. The program also performs likelihood ratio tests (LRTs) between (a) the null and GMYC models to test whether one or multiple species are involved, and (b) single and multiple threshold options.

Abbreviations used in Figures 3–18: bc, bursal canal; cb, copulatory bursa; cg, cement glands; cod, common oviduct; cs, cyanophilic secretion; dpf, dorsal penial

fold; ed, ejaculatory duct; fl, flap; go, gonopore; in, intestine; od, oviduct; pg, penial glands; ph, pharynx; pp, penis papilla; sg, shell gland; spf, spermatphore; sv, seminal vesicle; te, testis; vd, vas deferens.

RESULTS

Morphological analysis

Analysis of the qualitative features of the reproductive complex allowed us (1) to assign the Greek populations to eight of the nine species of *Dugesia* known for Greece, namely *Dugesia aenigma* De Vries, 1984, *D. arcadia* De Vries, 1988, *D. ariadnae* De Vries, 1984, *D. cretica*, *D. damoae* De Vries, 1984, *D. elegans* De Vries, 1984, *D. malickyi* De Vries, 1984 and *D. sagitta* (Schmidt, 1861) (Table 1), and (2) to identify a sexual population from Chios (Tripes-Parparia) as *D. sicula*. Further, four new species of *Dugesia* were identified by the presence of one or more unique characters or a unique combination of characters: *Dugesia naiadis* Sluys sp. nov., *Dugesia effusa* Sluys sp. nov., *Dugesia improvisa* Sluys & Solà sp. nov. and *Dugesia parasagitta* Sluys & Solà sp. nov. (see below; Table 1, units 3, 19, 20, 33). Unfortunately, we have been unable to analyse the morphological features of several populations (Table 1, units 5–9, 12, 13, 17, 22, 23, 30, 31), due to lack of (1) fixed material, (2) sexual specimens or (3) adequate histological sections. In addition, our samplings and subsequent comparative studies revealed the presence of a new dugesiid genus, *Recurva* Sluys gen. nov., represented by two species, namely *Recurva postrema* Sluys & Solà sp. nov. and *Recurva conjuncta* Sluys sp. nov. For a possible third, as yet unnamed, species of *Recurva* no morphological information was available. Detailed accounts of the relevant characters used to differentiate the candidate new species are presented in the Systematic and Integrative Section.

Phylogenetic analysis of dugesiid genera

Saturation analysis revealed that the third positions of the COI alignment including several genera were saturated; therefore, we excluded this codon position in all subsequent analyses.

The two phylogenetic methods used (MrBayes and RaxML) yielded almost identical topologies, albeit with different supports at some nodes (Fig. 1). *Recurva* is the sister group of *Schmidtea* Ball, 1974 in both analyses, and with high bootstrap (ML)/posterior probability (BI) (89/0.98) support. In turn, these two genera form the sister group of the *Dugesia* species, with maximum support (100/1). The COI analysis including *Cura pinguis* shows that the latter is not close to *Recurva* (Fig. S1). Within the *Recurva* clade we can distinguish *R. postrema*, *R. conjuncta* and a subclade formed by three sampling localities on the island of Paros. The latter three populations are likely belong to the same species, which most probably is neither *R. postrema* nor *R. conjuncta*. However, because all specimens from Paros were asexual it has not been possible to analyse them at the morphological level. There is no resolution in the relationships among these three taxa of *Recurva*.

Molecular species delimitation of *Dugesia*

The topology of the tree found in the GMYC analysis is very similar to that obtained in an earlier companion work (cf. Solà *et al*., 2013). LRT comparison between the results of the single and multiple threshold models in GMYC revealed no significant differences ($\chi^2 = 4.39$, d.f. = 6, $P = 0.63$); therefore, we present here only the results of the single threshold model. In the GMYC analysis the likelihood ratio test of the null against the mixed model was significant (4.5e-^{08}***).

The single analysis indicated a total of 34 entities [confidence interval (CI) = 31–42], clustered as follows: 29 ML clusters of two or more individuals (CI = 27–33), and five singletons (Fig. 2; Table 1). Eleven of these clusters match with morphologically identified *Dugesia* species, four of these newly described in this paper: *D. parasagitta* Sluys & Solà sp. nov. (entity 3); *D. aenigma* (4); *D. malickyi* (11); *D. ariadnae* (18); *D. improvisa* Sluys & Solà sp. nov. (19); *D. effusa* Sluys sp. nov. (20); *D. damoae* (21); *D. elegans* (24); *D. gonocephala* (Dugès, 1830) (32); *D. naiadis* Sluys sp. nov. (33); and *D. sicula* (34). Unfortunately, we were unable to fully analyse the taxonomic status of clusters 10 and 14 as the histological sections currently available are not of the required quality. However, even from the damaged sections it is clear that these units are morphologically different from their sister clades, *D. malickyi* (entity 11) and *D. arcadia* (entities 15 and 16), respectively. The putative new species of cluster 10 differs from *D. malickyi* in the presence of (a) a central, broad ejaculatory duct, (b) a small, ventral penial fold and (c) a highly glandular ejaculatory duct. The putative new species of entity 14 differs from *D. arcadia* in the absence of a lateral fold projecting into the atrium, a structure that is characteristic for *D. arcadia*. Because we have molecular and morphological data suggesting that clades 10 and 14 do not belong to any already known species, we consider them here as CCS.

In three cases the GMYC clusters do not match the morphologically delimited candidate species. First, although *D. sagitta* splits into two clusters (1 and 2; Fig. 2, Table 1), we were unable to find any morphological difference supporting this split.

Table 1. Clusters obtained in the GMYC analysis

Entity*	Code of Solà *et al.* (2013)	Locality†	No. of individuals in the cluster	Taxonomic category	Species‡
1	27	1. Roda, Corfu, Greece	10	UCS	*Dugesia* sp.
	29	2. Kato vrisi spring, Klimatia, Corfu, Greece			
2	33	1. North of Vouniatades, Corfu, Greece	10	DS	*D. sagitta*
	34	2. Benitses, Corfu, Greece			
3	31	1. Ermones, Corfu, Greece	9	DS	*D. parasagitta*
	32	2. Ermones, slightly higher than 31, Corfu, Greece			
4	35	1. Near Agia Eirini, Kephalonia	10	DS	*D. aenigma*
	36	2. Digaleto, Cephalonia			
5	17	Potamia, Preveza, Greece	4	UCS	*Dugesia* sp.
6	17	Potamia, Preveza, Greece	1	UCS	*Dugesia* sp.
7	14	Vafkeri, Lefkada, Greece	5	UCS	*Dugesia* sp.
8	13	Varia, Aetolia-Acarnania, Greece	4	UCS	*Dugesia* sp.
9	12	Eleonas-Gravia, Phocis, Greece	4	UCS	*Dugesia* sp.
10	20	Polidrosos, Phoci, Greece	2	CCS	*Dugesia* sp.
11	19	1. Mexiates, Phthiotis, Greece	10	DS	*D. malickyi*
	18	2. Gorgopotamos, Phthiotis, Greece			
12	16	Filiates, Thesprotia, Greece	3	UCS	*Dugesia* sp.
13	23	Dorio-Psari, Peloponnese, Greece	4	UCS	*Dugesia* sp.
14	21	Tripi, Peloponnese, Greece	5	CCS	*Dugesia* sp.
15	26	Chalandritsa, Peloponnese, Greece	5	DCL	*D. arcadia*
16	25	Sella, Peloponnese, Greece	3		*D. arcadia*
17	24	1. Theisoa-Andritsaina, Peloponnese, Greece	12	UCS	*Dugesia* sp.
	22	2. Agios Floros, Peloponnese, Greece			
	13	3. Varia, Aetolia-Acarnania, Greece			
	23	4. Dorio-Psari, Peloponnese, Greece			
18	6	Apollonas, Naxos, Greece	5	DS	*D. ariadnae*
19	7	Melanes, Naxos, Greece	5	DS	*D. improvisa*
20	9	Nagos, Chios, Greece	5	DS	*D. effusa*
	10	Nagos, before the opening to the sea, Chios, Greece			
21	8	Manolates, Samos, Greece	5	DS	*D. damoae*
22	11	Kalamoudi, Euboea, Greece	2	UCS	*Dugesia* sp.
23	11	Kalamoudi, Euboea, Greece	1	UCS	*Dugesia* sp.
24	5	Petaloudes Valley, Rhodes, Greece	2	DS	*D. elegans*
25	1	Georgioupoli, Crete, Greece	1	DCL	*D. cretica*
26	1	Georgioupoli, Crete, Greece	4		*D. cretica*
27	3	Sasalos, Crete, Greece	4		*D. cretica*
28	3	Sasalos, Crete, Greece	1		*D. cretica*
29	2	Kakopetros, Crete, Greece	5		*D. cretica*
30	–	Vernár, Slovak Republic	1	UCS	*Dugesia* sp.
31	–	Ludrová, Slovak Republic Prosiek, Slovak Republic	2	UCS	*Dugesia* sp.
32	–	Limburg, Netherlands	2	DS	*D. gonocephala*
33	–	Fita-Kimpouries, Chios, Greece	4	DS	*D. naiadis*
34	–	Tripes-Parparia, Chios, Greece	5	DS	*D. sicula*

*Includes clusters and singletons.
†Locality details may be found in Supporting information Table S1.
‡On the basis of morphology.
CCS, Confirmed Candidate Species; DCL, Deep Conspecific Lineage; DS, Described Species; UCS, Unconfirmed Candidate Species.

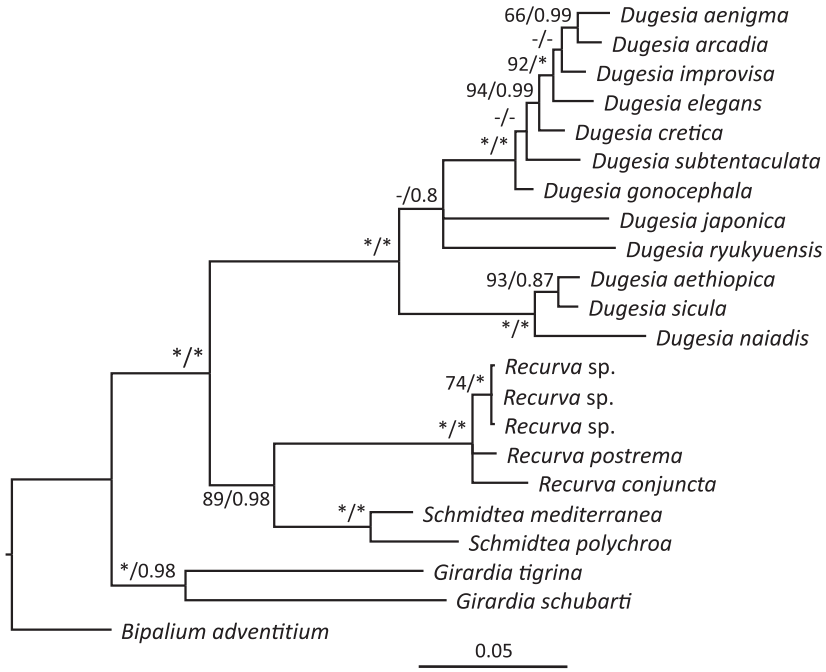


Figure 1. Bayesian tree inferred from the concatenated data set (COI + 18S). Labels correspond to species names. Node numbers correspond to bootstrap (ML)/posterior probability (BI); values are only indicated when >50/ >0.80; '*' indicates maximum support. The scale bar indicates substitutions per site.

Second, *Dugesia arcadia* was identified from Sella and Chalandritsa in the northern Peloponnisos (localities 25 and 26, or entities 15 and 16), but these two localities with morphologically identical individuals are split in the GMYC analysis. However, the divergence of these two populations almost coincides with the GMYC threshold.

A third case concerns specimens from Crete. All individuals that were examined from three sampling localities on this island presented the diagnostic features of *D. cretica*, although the coalescent-based tree splits them into no fewer than five units (entities 25–29), comprising three clusters and two singletons (Fig. 2), which we here consider DCL (Table 1).

Furthermore, 11 clusters and three singletons concern specimens that could not be checked morphologically. Among these cases is a large cluster (entity 17; 12 individuals) that includes three sampling sites from the Peloponnisos (Theisoa-Andritsaina, Agios Floros and Dorio-Psari) and also one individual from Lake Trichonida (Varia, Aetolia-Acarnania) in Central Greece. Another case is an individual from the Potamia locality in Preveza (entity 6), constituting a singleton that groups with high support with a different clade than the other four specimens from the same locality (entity 5), thus suggesting the presence of two different species at the same site. Finally, three individuals from Euboea also split in two different clusters (22 and 23). All of these clusters and singletons for which we lack morphological data are here considered as Unconfirmed Candidate Species.

SYSTEMATIC AND INTEGRATIVE SECTION

Order Tricladida Lang, 1884

Family Dugesiidae Ball, 1974

Genus *Dugesia* Girard, 1850

Dugesia effusa Sluys **sp. nov.** (Figs 3–5)

Material examined: Holotype: ZMA V.Pl. 7114.1, river just before opening into the sea, Nagos, Chios, Greece, 38°33′32.31″N, 26°4′59.42″E, 30 April 2010, coll. M. Vila-Farré, sagittal sections on seven slides.

Paratypes: ZMA V.Pl. 7114.2, ibid., sagittal sections on six slides; V.Pl. 7114.3 (RS 221-3), ibid., horizontal sections on three slides.

Other material: ZMA V.Pl. 7115.1, river, Nagos, Chios, Greece, 38°33′27.57″N, 26°4′51.61″E, 30 April 2010, coll. M. Vila-Farré, sagittal sections on five slides; V.Pl. 7115.2, ibid., sagittal sections on five slides; V.Pl. 7115.3, ibid., horizontal sections on three slides.

Etymology: The specific epithet is derived from the Latin adjective *effusus*, generous, abundant, and alludes to the highly glandularized penis papilla.

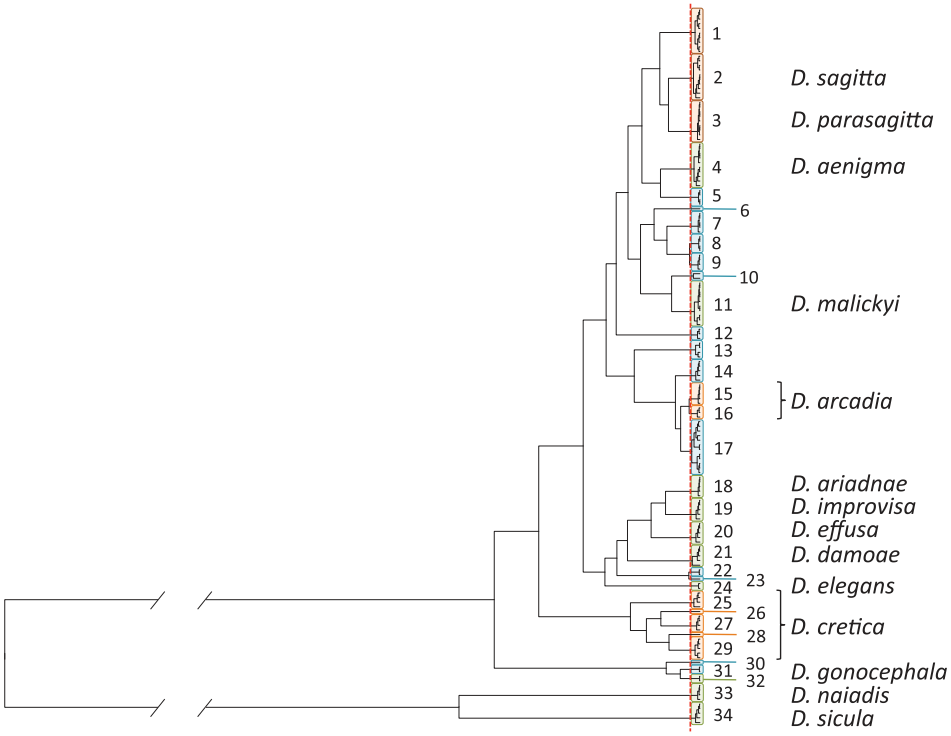


Figure 2. Result of the GMYC analysis. Threshold-delimiting speciation and coalescent processes plotted as a broken line. Numbers indicate molecular-based entities; labels correspond to species names. Entities in green show correspondence between the molecular species delimitation method and the morphologically identified species. In orange are shown groupings where there is conflict between morphological and molecular methods. In blue are shown the groupings for which only molecular data are available.

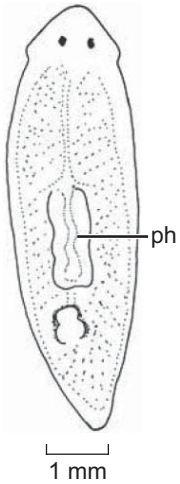


Figure 3. ***Dugesia effusa*** Sluys **sp. nov**. Dorsal view of preserved specimen.

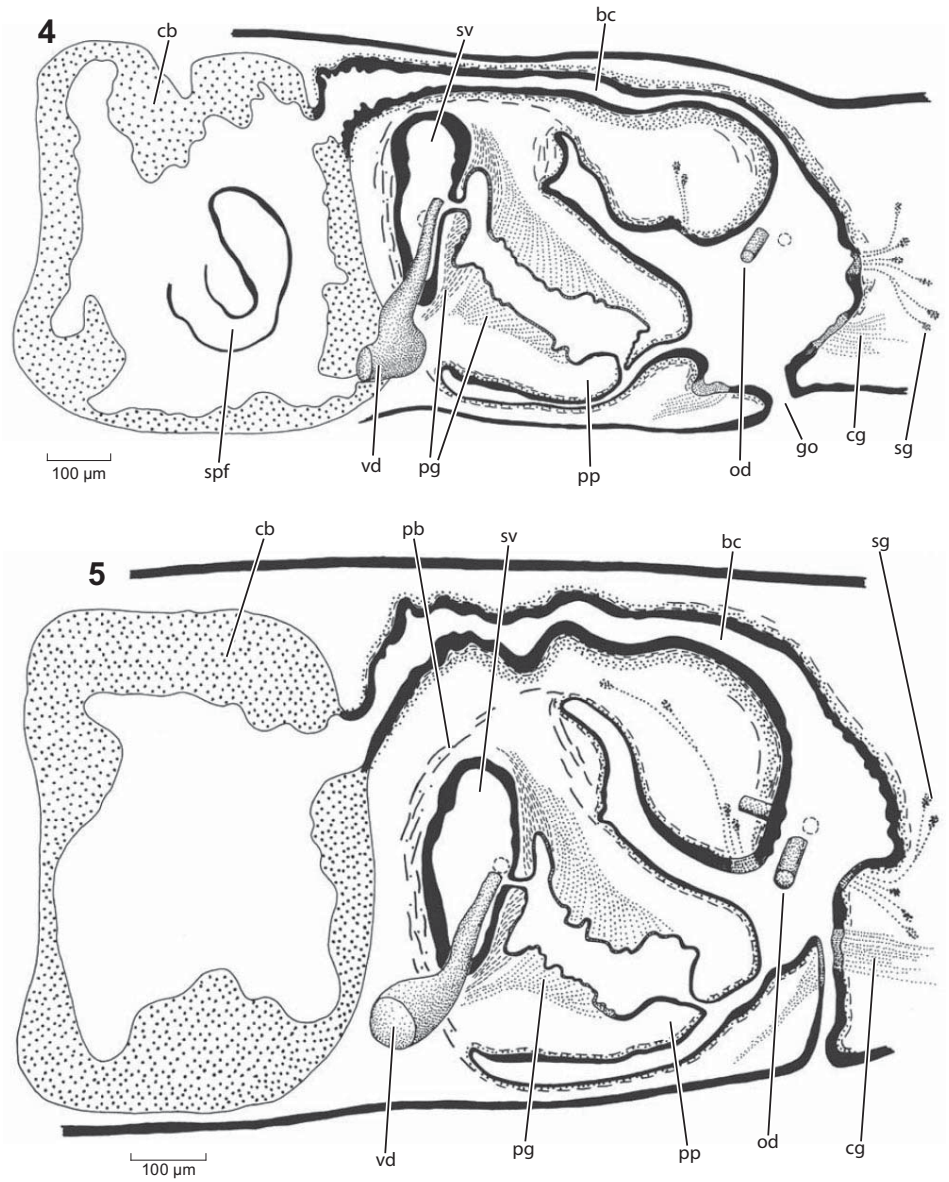
Diagnosis: Dugesia effusa is characterized by the combination of the following features: presence of a small, dorsal penial fold; central ejaculatory duct; short, valve-like diaphragm; large, intrabulbar seminal vesicle; highly glandularized penis papilla; a bursal canal that widens considerably at its communication with the atrium; ectal reinforcement of the bursal canal confined to the vaginal region.

Ecology and distribution: The species is known only from two sites in the same river, i.e. the type locality close to the opening into the sea and another site further upstream.

Description: Preserved specimens up to 9 × 2.25 mm, with low-triangular head with rounded auricles; tail obtusely pointed (Fig. 3). Dorsal surface pale brown; ventral surface pale. Two eyes, situated in pigment-free patches.

Pharynx situated in the mid-region of the body, measuring between one-quarter and one-sixth of the body length. Mouth opening located at the posterior end of the pharyngeal pocket.

The testes are located dorsally and extend from the level of the ovaries into the posterior end of the body. The vasa deferentia penetrate the ventro-lateral wall of the penis bulb and open into the seminal vesicle at a point very close to the diaphragm. The ovoid or



Figures 4, 5. ***Dugesia effusa*** Sluys **sp. nov. 4.** ZMA V.Pl. 7114.2. Sagittal reconstruction of the copulatory apparatus. **5.** ZMA V.Pl. 7114.1. Sagittal reconstruction of the copulatory apparatus.

pear-shaped seminal vesicle fills the major part of the penis bulb and is lined with a columnar, nucleated epithelium. Through a very narrow diaphragm this seminal vesicle opens into the funnel-shaped, proximal section of the ejaculatory duct (Fig. 4). The short, stubby lips of the valve-like diaphragm, as well as the funnel-shaped section of the ejaculatory duct, receive the finely granular and dark red staining secretion of erythrophil penis glands. The broad ejaculatory duct follows a slightly ventrally displaced course through the penis papilla and opens at the blunt tip of the penis papilla, the actual opening being rather narrow. Along the major part of its length the lining epithelium of the ejaculatory duct is pierced by the numerous openings of abundant penis glands that produce an orange–brown secretion.

The plug-shaped penis papilla is lined with a nucleated epithelium and is provided with a subepithelial layer of circular muscles, followed by a layer of longitudinal muscles. A penial fold is located symmetrically at the dorsal base of the penis papilla; the fold is traversed by some longitudinal muscle fibres.

The ovaries are situated directly medially to the ventral nerve cords and are located at one-third to

one-quarter of the distance between the brain and the root of the pharynx. The oviducts are lined with an infranucleated epithelium and are surrounded by a well-developed coat of circular muscles. The oviducts open separately into the ventral-most, widened section of the bursal canal, close to the point where the canal communicates with the atrium. Shell glands discharge their secretion into the bursal canal ventrally to the oviducal openings.

The bursal canal is lined with a nucleated, cuboidal-columnar epithelium. The diameter of the bursal canal increases considerably near its point of communication with the atrium. Notably the most ventral section of the canal, at the level of the oviducal openings, shows a widening into posterior direction (Fig. 5). The bursal canal is overlain with a thin layer of circular muscles, the latter being particularly developed in the vaginal region. Ectal reinforcement in the form of outer longitudinal muscle fibres is present in the vaginal area and extends towards the point where the bursal canal bends forwards. The copulatory bursa is a voluminous sac-shaped structure that fills the entire dorsoventral space of the body. In several specimens remnants of a spermatophore are present in the bursa.

Discussion

A dorsal penial fold of similar size and location as in this species *D. effusa* is present also in *D. sagitta* (some specimens have only one, dorsal fold), *D. malickyi*, *D. benazzii* Lepori, 1951, *D. elegans* and *D. leporii* Pala, Stocchino, Corso & Casu, 2000. In *D. elegans* the openings of the vasa deferentia into the seminal vesicle are far removed from the diaphragm, contrasting with the location of the openings immediately anterior to the diaphragm in all other species mentioned. In addition, the penial fold of *D. elegans* is more developed and more strongly muscular than in *D. effusa*. (cf. De Vries, 1984).

Dugesia leporii differs from *D. effusa* in the presence of a pointed diaphragm and small intrabulbar seminal vesicle, and in the fact that its ectal reinforcement extends from the vaginal area far anterior along the bursal canal (cf. Pala *et al*., 2000). In contrast to *D. effusa*, *D. benazzii* is provided with a small intrabulbar seminal vesicle and a pointed diaphragm (cf. Lepori, 1951; De Vries, 1984).

The gross morphology of *D. effusa* is very similar to that of *D. malickyi* and *D. sagitta*. But *D. malickyi* differs from *D. effusa* in the presence of (1) a considerably bigger penial fold that also has a distinctly lateral position, and (2) a much narrower and distinctly ventrally displaced ejaculatory duct, the latter being devoid of the high glandularization that occurs in *D. effusa*. Such a highly glandular papilla, however, is also characteristic of *D. sagitta* (cf. De Vries, 1984) and also of *D. improvisa* Sluys & Solà sp. nov. The last-mentioned species lacks the penial fold as well as the widening of the bursal canal in the vaginal area, while its seminal vesicle is highly glandular, in contrast to the conditions in *D. effusa*.

The GMYC analysis supports *D. effusa* as a different species (Fig. 2, Table 1), clearly delimitating the specimens from Chios as entity 20. Furthermore, *D. effusa* is not close to *D. sagitta* in the phylogenetic tree of Solà *et al*. (2013). Nevertheless, *D. effusa* shares with *D. sagitta* the 'V-shaped' glandular zone that surrounds the ejaculatory duct (cf. De Vries, 1984: 106). In *D. sagitta* there are usually two penial folds, the ventral one being smaller than the dorsal one; the ventral fold may also be completely absent. However, in relation to the size of the penis papilla, the penial fold of *D. sagitta* is considerably bigger than that in *D. effusa*. Furthermore, the dorsal penial fold of *D. sagitta* is traversed by a cyanophilic secretion, which is discharged through its lining epithelium; such is not the case in *D. effusa*.

DUGESIA IMPROVISA SLUYS & SOLÀ **SP. NOV.** (FIGS 6–9)

Material examined: Holotype: ZMA V.Pl. 7116.1, Melanes, Naxos, Greece, 37°5′3.38″N, 25°26′59.40″E, alt. 199 m, 9 April 2009, coll. Eduardo Mateos & Eduard Solà, sagittal sections on nine slides.

Paratypes: ZMA V.Pl. 7116.2, ibid., sagittal sections on ten slides; V.Pl. 7116.3, ibid., horizontal sections on four slides; V.Pl. 7116.4, ibid., sagittal sections on eight slides.

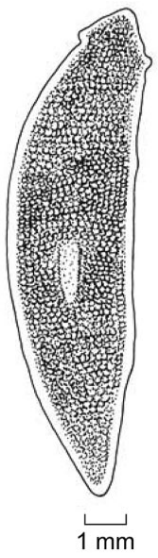
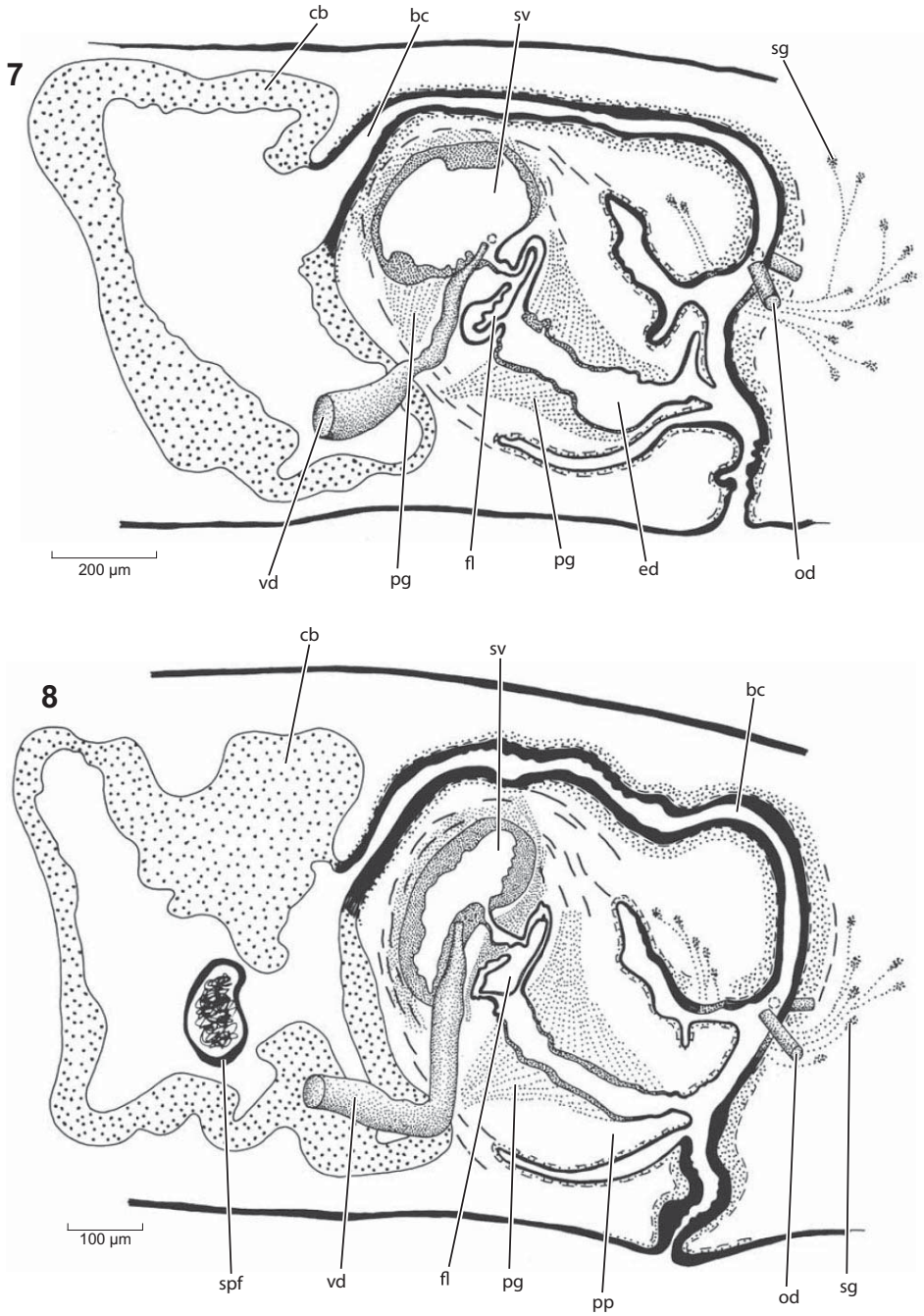


Figure 6. ***Dugesia improvisa*** Sluys & Solà **sp. nov.** Dorsal view of preserved specimen.



Figures 7, 8. ***Dugesia improvisa*** Sluys & Solà **sp. nov. 7.** ZMA V.Pl. 7116.2. Sagittal reconstruction of the copulatory apparatus. **8.** ZMA V.Pl. 7116.1. Sagittal reconstruction of the copulatory apparatus.

Etymology: The specific epithet is derived from the Latin adjective *improvisus*, unexpected, and alludes to our surprise in finding a second and new species of *Dugesia* on such a small island as Naxos.

Diagnosis: Dugesia improvisa is characterized by: an acentral, ventrally displaced ejaculatory duct, opening at the tip of the penis papilla; a short diaphragm; ectal reinforcement being confined to the posterior wall of the ascending portion of the bursal canal; vasa deferentia separately opening into the anterior section of the seminal vesicle, at a point close to the diaphragm; broad zone of abundant penis glands traversing the penial papilla and opening into the ejaculatory duct.

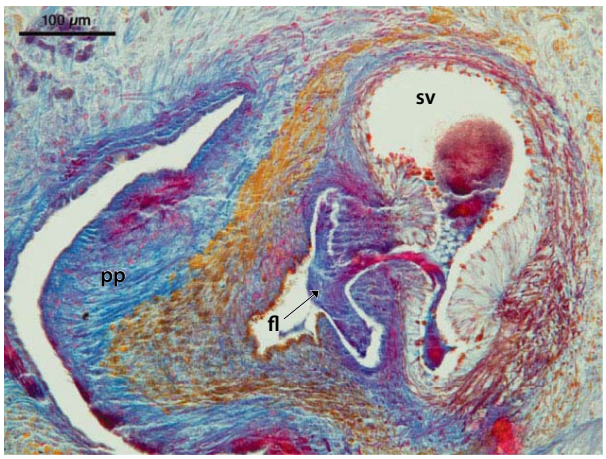


Figure 9. ***Dugesia improvisa*** Sluys & Solà **sp. nov.** Photomicrograph of penial complex of specimen ZMA V.Pl. 7116.4, showing the sickle-shaped flap of tissue or secretion.

Ecology and distribution: Specimens were collected from under stones in a small, shallow pool, receiving the outflow of water from a concrete pipe. The species is known only from this type locality.

Description: Preserved specimens up to about 12.5 × 3 mm. Triangular head with distinct, blunt auricles. Posterior end obtusely pointed. Dorsal surface pale brown, with the pigment arranged in a finely reticulated pattern and with a concentration of pigment following the outline of the pharyngeal pocket (Fig. 6). Dorsal body margin and ventral surface pale. The two eyes are situated in conspicuous pigment-free patches.

The pharynx is located in the posterior half of the body and measures about 1/8th of the body length in preserved specimens. The mouth opening is located at the posterior end of the pharyngeal pocket.

The testes are located dorsally and extend from the level of the ovaries to the posterior end of the body.

The vasa deferentia penetrate the antero-lateral wall of the intrabulbar seminal vesicle; the ducts open separately into the vesicle at a position very close to the diaphragm (Fig. 7). The intrabulbar seminal vesicle is lined with an epithelium, consisting of columnar cells, that is pierced by the numerous openings of penis glands, the latter producing a granular, erythrophil secretion. At the free end of the lining epithelium of the seminal vesicle this secretion projects into the lumen as relatively large, pear-shaped, granular drops. Through a short, stubby diaphragm the seminal vesicle opens into the proximal, funnel-shaped section of the ejaculatory duct.

The diaphragm is short. The proximal funnel-shaped section of the ejaculatory duct, immediately adjacent to the diaphragm, houses a sickle-shaped flap of tissue or secretion (Figs 7–9). This flap seems to be attached to the rest of the diaphragm by only a minute piece of tissue. The lining epithelium of the flap is pierced by the openings of the erythrophil penis glands that open into the seminal vesicle and also penetrate the epithelium of the rest of the diaphragm. The flap was observed in all four specimens examined and its histology suggested true mesenchyme, surrounded by an epithelium.

The ejaculatory duct runs slightly acentrally, i.e. ventrally displaced, through the penis papilla, opening at its tip. The major portion of the ejaculatory duct receives the conspicuous, abundant and granular secretion of erythrophil penis glands, which are located outside of the penis.

The penis papilla is a broad, pointed or blunt cone. The papilla is covered with a nucleated epithelium and is underlain with a subepithelial layer of circular muscles, followed by a layer of longitudinal muscles. The penis bulb is well developed and muscular.

The small, paired ovaries are situated at about 1/3rd of the distance between the brain and the root of the pharynx and are positioned directly medially to the ventral nerve cords. The oviducts arise from the dorsal surface of the ovaries and run backwards immediately dorsally to the ventral nerve cords. At the level of the copulatory apparatus the oviducts curve dorsomedially to open separately into the most proximal, posterior, section of the bursal canal, i.e. close to the point where the duct communicates with the atrium. Erythrophil shell glands open into the bursal canal immediately ventrally to the openings of the oviducts.

The bursal canal is lined with a cuboidal, nucleated epithelium and is surrounded by a reversed musculature: a thin subepithelial layer of longitudinal muscle, followed by a thicker layer of circular muscle. Around the proximal, posterior, section of the bursal canal this circular muscle layer is rather thick, but it becomes gradually thinner towards the copulatory bursa. Ectal reinforcement of the bursal canal musculature is only present along the proximal section of the canal, i.e. from its opening into the atrium to about the point where the duct curves anteriad. However, this ectal reinforcement is only present as a single layer of longitudinal muscle along the posterior wall of the ascending portion of the bursal canal; it was not observed along the anterior wall of this part of the canal. The bursal canal communicates with a large, sac-shaped copulatory bursa, which occupies most of the dorso-ventral space of the body. In two specimens the bursa contained remnants of a sclerotic spermatophore.

Discussion

The presence of a peculiar flap of tissue on the diaphragm sets *D. improvisa* immediately apart from

any of the known species of *Dugesia*. However, in specimens of other species of *Dugesia* a more or less crescent-shaped stretch of secretion may be present in precisely the same position, albeit less clearly attached to the epithelium, while in these specimens its staining properties clearly suggest a glandular origin. In these animals, and also in *D. improvisa*, this flap or stretch of secretion may be related to the formation of the spermatophore (which is formed in the ejaculatory duct) or to the transfer of sperm into the latter. However, in *D. improvisa* the flap did not resemble a spermatophore *in statu nascendi* but suggested true mesenchyme surrounded by an epithelium. We are hesitant to consider this feature as a diagnostic character of *D. improvisa*, but would first prefer to check the presence of this flap in another series of individuals of *D. improvisa*. Unfortunately, additional material is not presently available. However, *D. improvisa* also presents a combination of other characters that makes it different from its congeners.

In the fact that the vasa deferentia open into the seminal vesicle at a point close to the diaphragm, *D. improvisa* resembles a good number of other species of *Dugesia* (cf. Sluys *et al*., 1998, table II). However, in other features these species differ much from *D. improvisa*, for example in the presence of penial or atrial folds, except *Dugesia subtentaculata* (Draparnaud, 1801) and *D. burmaensis* (Kaburaki, 1918). However, the atrium of *D. subtentaculata* shows a distinct musculo-glandular area (cf. De Vries, 1986), which is absent in *D. improvisa*. Furthermore, *D. subtentaculata* also possesses a ring of spongiose mesenchymatic tissue in the penis papilla that is absent in *D. improvisa*. In addition, in *D. subtentaculata* the ectal reinforcement along the bursal canal is much more developed and extends much farther anteriad.

The gross morphology of the copulatory apparatus of *D. burmaensis* is very similar to that of *D. improvisa*. However, for *D. burmaensis* it has been reported that the oviducts arise from the anterolateral wall of the ovaries, contrasting with their dorsal origin in *D. improvisa*. *Dugesia burmaensis* resembles *D. improvisa* in the presence of highly developed penis glands, discharging their abundant secretion into the ejaculatory duct. Such a broad zone with abundant secretion traversing the penis papilla is also characteristic of *D. sagitta* from Corfu. However, there are a number of clear differences between *D. sagitta* and *D. improvisa*.

In *D. sagitta* the penis papilla is blunt and provided with distinct, asymmetric penial folds at both the dorsal and the ventral side of its base (cf. De Vries, 1984), which are absent in *D. improvisa*. Furthermore, in *D. sagitta* the ejaculatory duct follows a central course through the penis papilla, whereas it has a ventrally displaced trajectory in *D. improvisa*. In addition, the ectal reinforcement of the bursal canal extends much farther anterior in *D. sagitta*.

In all molecular analyses *D. improvisa* is the sister species of *D. ariadnae* (Fig. 2; Solà *et al*., 2013), the latter also restricted in its distribution to the island of Naxos. However, the two species are clearly delimited in the GMYC analysis, while morphologically *D. ariadnae* is very different from *D. improvisa*. In particular, *D. ariadnae* is characterized by two well-developed adenodactyls that are suspended from the dorsal atrial wall, one on either side of the base of the penis. On the basis of our comparative and integrative analysis, as presented above, we conclude that *D. improvisa* concerns a new species.

DUGESIA NAIADIS SLUYS SP. NOV.
(FIGS 10–12)

Material examined: Holotype: ZMA V.Pl. 7117.1, 650 m before Kipouries (coming from Fita), Chios, Greece, 38°30′43.31″N, 25°59′55.06″E, 30 April 2010, coll. M. Vila-Farré, sagittal sections on 12 slides.

Paratype: ZMA V.Pl. 7117.2, ibid., sagittal sections on nine slides.

Etymology: The specific epithet is derived from the Latin *naias*, water nymph, and alludes to the small freshwater stream from which the specimens were collected.

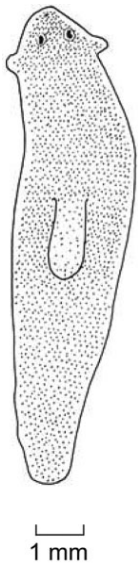
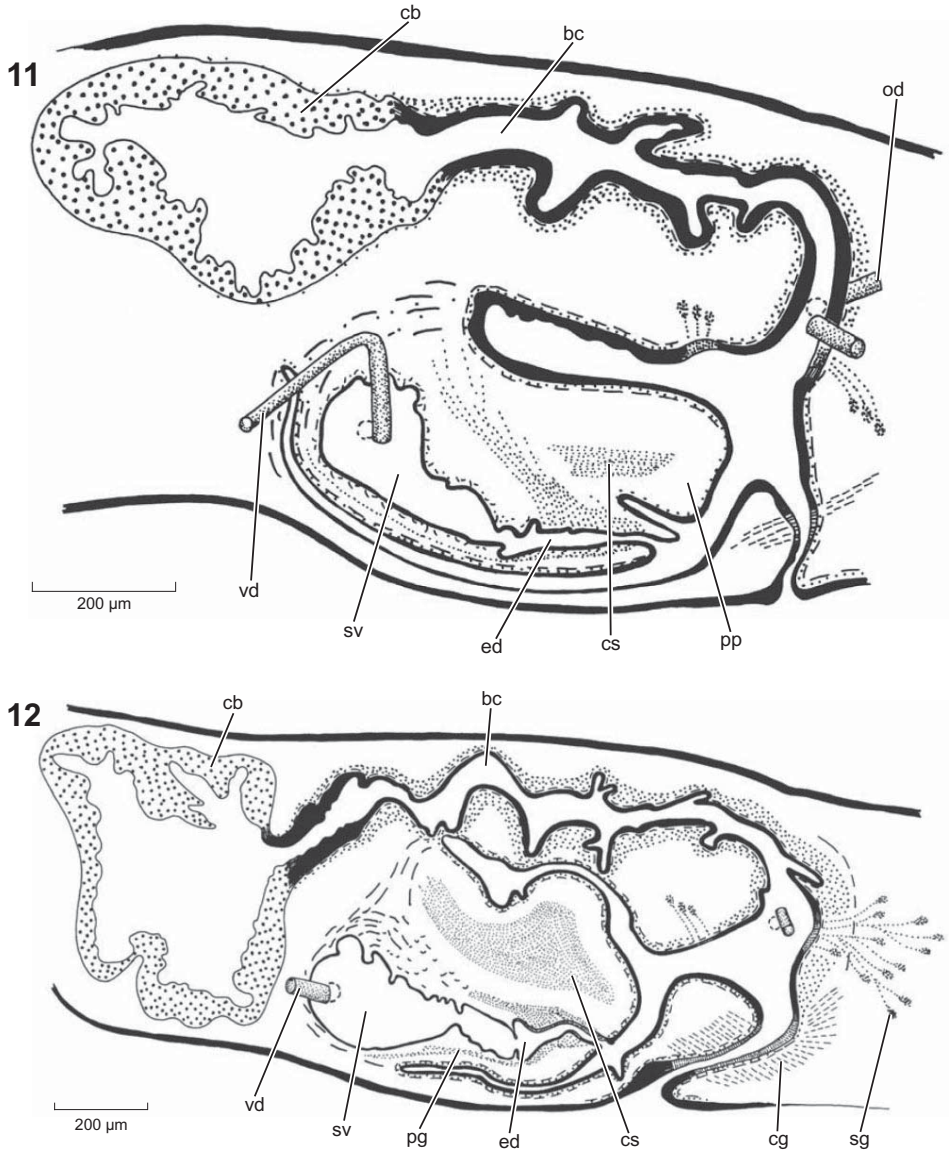


Figure 10. ***Dugesia naiadis*** Sluys **sp. nov.** Dorsal view of preserved specimen.



Figures 11, 12. ***Dugesia naiadis*** Sluys **sp. nov. 11.** ZMA V.Pl. 7117.2. Sagittal reconstruction of the copulatory apparatus. **12.** ZMA V.Pl. 7117.1. Sagittal reconstruction of the copulatory apparatus.

Diagnosis: Dugesia naiadis is characterized by: vasa deferentia that open into the proximal, anterior section of the seminal vesicle; a short diaphragm; an acentral, ventrally displaced ejaculatory duct, opening terminally at the tip of a blunt penis papilla; a broad zone of cyanophilic secretion in the dorsal section of the penis papilla; oviducts that open symmetrically into the most proximal section of the bursal canal; a bursal canal provided with many irregular pleats and folds, surrounded by a well-developed coat of circular muscle and a zone of mesenchymatic, erythrophil gland cells; hyperplasic ovaries; lack of testes.

Ecology and distribution: Specimens were collected from a small creek; the species is known only from the type locality.

Description: Preserved specimens with low triangular head and rounded auricles (Fig. 10), measuring up to 11 mm in length and 2.5 mm in width. Dorsal body surface pale brown; ventral surface pale. A pair of eyes is present and somewhat smaller additional eyes are present also in the sectioned specimens.

Pharynx located in the middle of the body, measuring about 1/6th of the body length. The mouth opening is located at the posterior end of the pharyngeal pocket.

Testes are completely absent. The ovaries are hyperplasic: ovarian tissue fills the entire dorso-ventral space over a distance of about 750 μm. The midpoint of the hyperplasic ovaries is located at about 1/4th the distance between the brain and the root of the pharynx.

The oviducts open separately and symmetrically into the most proximal section of the bursal canal, i.e. close to the point where the canal communicates with the atrium (Fig. 11). Erythrophil shell glands discharge their secretion into the bursal canal, immediately ventrally to the oviducal openings.

The bursal canal is lined with a nucleated epithelium; it follows a somewhat undulating course towards the copulatory bursa, while giving rise to a number of irregular pleats or folds that project into the surrounding mesenchyme (Fig. 12). The canal is surrounded by a very thin, subepithelial layer of longitudinal muscle, followed by a thick layer of circular muscle. Ectally to its surrounding coat of muscles the bursal canal is surrounded by a zone of mesenchymatic, erythrophil gland cells, which discharge their secretion into the lining epithelium of the canal. Only in specimen ZMA V.Pl. 7117.1 (Fig. 12) could ectal reinforcement by some longitudinal muscles be detected on the posterior wall of the bursal canal, in the region of the oviducal openings.

In specimen ZMA V.Pl. 7117.1 (Fig. 12) the copulatory bursa is a large sac-shaped structure that fills the entire dorso-ventral space, but in ZMA V.Pl. 7117.2 (Fig. 11) the bursa is much smaller and also lined with cells with a more densely stained content.

Although the oviducts run from the level of the copulatory apparatus to the ovaries, vasa deferentia could be traced only in the vicinity of the penis bulb. After having penetrated the ventro-lateral wall of the penis bulb, the vasa deferentia open separately into the proximal, anterior section of the seminal vesicle. The latter gradually narrows towards a small diaphragm, through which it communicates with the ejaculatory duct. Seminal vesicle and ejaculatory duct are positioned in the ventral region of the penis papilla, which therefore is asymmetrical: its dorsal section is much larger that the ventral section. The ejaculatory duct receives the secretion of numerous erythrophil penis glands and opens terminally at the blunt tip of the penis papilla. The latter is a plug-shaped structure that fills most of the male atrium. The penis papilla is covered with a nucleated epithelium that is underlain by a thin layer of circular muscle, followed by an equally thin layer of longitudinal muscle. The dorsal section of the penis papilla is traversed by a broad zone of strands of cyanophilic secretion that does not seem to open into the ejaculatory duct or through the covering epithelium of the papilla. The spaces present in the penial mesenchyme, near the tip of the papilla, seem to result from clefts in torn tissue.

Discussion

Presence of hyperplasic ovaries and complete absence of testes are signs that these animals probably concern sexualized specimens from an otherwise asexually reproducing population. Such sexualization may be induced either spontaneously (as was the case with these animals from Chios) or experimentally and has been reported for 11 species of *Dugesia* (cf. Charni *et al*., 2004 and references therein; Stocchino, Sluys & Manconi, 2012; Harrath *et al*., 2013). Furthermore, hyperplasic ovaries and poorly developed testes have been found also in ex-fissiparous specimens of *Phagocata morgani* (Stevens & Boring, 1906; Benazzi & Ball, 1972).

The fortunate circumstance that animals of an otherwise asexually reproducing population sometimes develop reproductive organs enables taxonomic identification of such specimens. In that context, the animals from Chios should be compared with other species for which a ventrally displaced ejaculatory duct has been reported, forming a presumably monophyletic subset within the genus *Dugesia* (Sluys *et al*., 1998). This comparison should be restricted to those species in which the ventrally displaced ejaculatory duct opens terminally at the tip of the penis papilla, thus excluding species with a subterminal opening. This immediately excludes *D. sicula*, *D. aethiopica* Stocchino *et al*., 2013 and *Dugesia arabica* Harrath & Sluys, 2013 as candidate species because these have a subterminal opening of the ejaculatory duct. However, both *D. aethiopica* and *D. arabica* resemble the Chios specimens in the presence of a bursal canal with many elaborate folds, a feature that has been reported also for *D. biblica* (cf. Benazzi & Banchetti, 1972), albeit that in the latter it is much less developed in comparison with *D. aethiopica*, *D. arabica* and the Chios specimens of *D. naiadis*. For *D. biblica* Benazzi & Banchetti (1972) describe the bursal canal as having '. . . un diametro alquante irregolare . . .' [a considerably irregular diameter], which agrees with our observations on specimens from Israel (ZMA V.Pl. 698.1, V.Pl. 699.1).

Another difference between the Chios animals and *D. aethiopica* and *D. sicula* concerns the openings of the oviducts into the bursal canal. In both *D. sicula* and *D. arabica* the oviducal openings are highly asymmetrical, in contrast to the symmetrical openings in *D. naiadis* (cf. Sluys, 2007; Harrath *et al*., 2013). In the specimens of *D. aethiopica* from Ethiopia the situation is different in that the oviducts open symmetrically into the ventral part of the horizontally running section of the bursal canal. In these

type specimens the proximal section of the bursal canal approaches the atrium by running more or less parallel to the body surface, thus contrasting with the course of the canal in *D. naiadis*.

In the presence of mesenchymal glands around the bursal canal and the patch of cyanophilic secretion in the penis papilla *D. naiadis* resembles *D. sicula*, *D. biblica* and the presumed *biblica* specimens from Bucak, Turkey (ZMA V.Pl. 813). However, in other features *D. naiadis* differs from these taxa.

The phylogenetic analysis (Fig. 1) shows that *D. naiadis* belongs to the *sicula–aethiopica* clade (as defined in Lázaro *et al*., 2009) with maximum support (100/1), being the sister group of *D. aethiopica* and *D. sicula*. The fact that the GMYC method (Fig. 2, Table 1, entity 33) delimits the four specimens of *D. naiadis* as a differentiated species supports the description of this new species. Interestingly, *D. naiadis* does not present the duplication in the nuclear ribosomal internal transcribed spacer-1 (ITS-1) molecule that *D. aethiopica* and *D. sicula* share (data not shown; cf. Baguñà *et al*., 1999; Lázaro *et al*., 2009).

On the basis of their gene identity we have been able to assign several asexual *Dugesia* populations from Chios to either *D. naiadis* or *D. sicula* (Table S5).

DUGESIA PARASAGITTA SLUYS & SOLÀ **SP. NOV.** (FIG. 13)

Material examined: Holotype: ZMA V.Pl. 7118.1, Ermones, Corfu, Greece, 39°36′37.98″N, 19°46′41.64″E, somewhat higher upstream than ZMA V.Pl. 7119, 20 April 2009, coll. R. Sluys, sagittal sections on 13 slides.

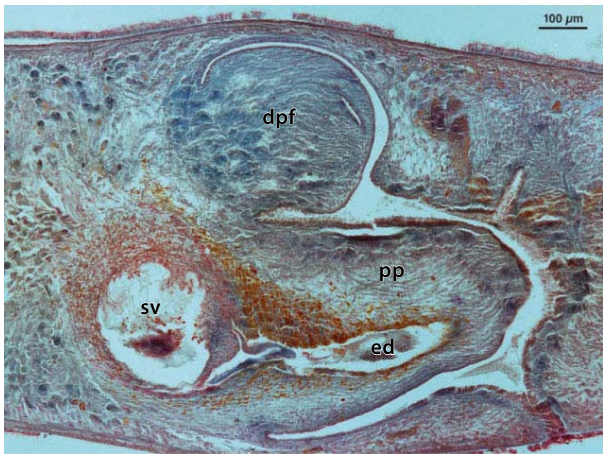


Figure 13. ***Dugesia parasagitta*** Sluys & Solà **sp. nov.** Photomicrograph of large dorsal penial fold in specimen ZMA V.Pl. 7118.1.

Paratypes: ZMA V. Pl. 7118.2, ibid., horizontal sections on eight slides; V.Pl. 7118.3, ibid., sagittal sections on six slides.

Other material examined: ZMA V.Pl. 7119.1, Ermones, Corfu, Greece, 39°36′41.93″N, 19°47′1.40″E, outflow of river into the sea, 20 April 2009, coll. R. Sluys, sagittal sections on five slides; V.Pl. 7119.3, ibid., horizontal sections on six slides; V.Pl. 7119.4, ibid., sagittal sections on 18 slides, V.Pl. 7119.5, ibid., sagittal sections on 14 slides; V.Pl. 7119.6, ibid., sagittal sections on 17 slides.

Etymology: The specific epithet is based on the prefix *para* (somewhat resembling, related to) and the specific epithet of the species *D. sagitta*.

Diagnosis: The species differs morphologically from its closest relative, *D. sagitta*, in the presence of a very large dorsal penial fold, very small ventral fold and a ventrally displaced ejaculatory duct.

Ecology and distribution: The species is known only from two sites in the same river. One site is close to the opening of this river into the sea, while the type locality is located slightly farther upstream.

Comparative discussion: The taxonomic status of *D. sagitta* (Schmidt, 1861) from Corfu as a valid and separate species was clarified by De Vries (1984). Prior to her study, the *Dugesia* populations from Corfu were usually considered to be conspecific with *D. gonocephala*, following a conclusion reached by Komárek (1925). To avoid future taxonomic confusion, De Vries (1984) fortunately designated a series of neotypes for *D. sagitta*. Although the International Code of Zoological Nomenclature (ITZN, 1985; ICZN, 1999) restricts designation of a neotype to only one specimen that forms the new name-bearing type of a nominal species and thus does not allow it to be a *series* of animals, the neotype specimens specified by De Vries (1984: 104) represent a morphologically homogeneous set of animals. As neotype locality was chosen Messonghi River, just west of Messonghi.

The Ermones population was first mentioned by Ball (1979), who attributed it to *D. gonocephala*. In the same paper the karyotype of presumed *D. gonocephala* from Corfu was analysed but it is not clear which population was studied, either the one from Ermones or the animals from Messonghi River. However, De Vries (1984) writes that animals from the neotype locality of *D. sagitta*, i.e. Messonghi River, were analysed.

Our integrative analysis of the populations that we sampled from Corfu revealed an unexpected and interesting situation. Molecular analysis of both COI and ITS-1 grouped the various populations sampled

into three clades (cf. Solà *et al*., 2013). These three clades are also identified as separate entities in the GMYC analysis (Fig. 2, entities 1, 2 and 3). One clade was formed by populations 27, 28, 29 and 30 (i.e. north of the San Salvador mountain range). The second clade consisted of populations 33 and 34. The third clade consisted of two samples from basically the same locality, namely Ermones (localities 31 and 32) (Fig. S2).

On the basis of morphological analysis of the populations from Corfu we were able to differentiate between only two types. The majority of the populations sampled conformed to the classical diagnosis of *D. sagitta*, notably in the presence of a well-developed dorsal fold and a very small or absent ventral fold, and with a central ejaculatory duct. This also holds true for populations that we have not re-collected, but of which material is present in the collections of the NBC: Messonghi River, Marbella beach (now called Par. Ag. Ioannis Peristeron) and Mesaria. However, the population from Ermones (ZMA V.Pl. 7118 + V.Pl. 7119) is characterized by a very large dorsal penial fold, very small ventral fold and a ventrally displaced ejaculatory duct (Fig. 13). Thus, coincidence of molecular and morphological results suggests that at least the population from Ermones is well differentiated from other populations on Corfu. Therefore, we do here describe this population as the new species *D. parasagitta*.

It remains remarkable that the populations that are geographically closest to the *D. sagitta* type locality, namely ZMA V.Pl. 7120 from near Vouniatades (locality 33) and ZMA V.Pl. 7121 from near Benitses (34) (entity 2, Fig. 2), differ molecularly so much from the populations in the northern part of the island (entity 1, Fig. 2), whereas morphologically they cannot be distinguished from each other, nor from the neotype population.

After the separation and description of *D. parasagitta*, the nominal species *D. sagitta* actually forms a paraphyletic taxon, according to all molecular analyses done so far (cf. Solà *et al*., 2013; COI gene tree, Fig. 2). Furthermore, the geographical distribution of the various populations (Fig. S2) suggests that these two units form two independent lineages. In view of the definition of a species as an independently evolving lineage, this suggests that these lineages are actually two different species. We do take a conservative approach to taxonomy and do not assign formal species status to these taxa, pending the availability of further data. However, we do suggest that entity 2 (from localities 33 and 34, i.e. in the proximity of the neotype locality of *D. sagitta*) is assigned to the nominal species *D. sagitta*, and that entity 1 (from localities 27 and 29) represents a UCS (Fig. S2, Table 1).

GENUS *RECURVA* SLUYS **GEN. NOV.**

Diagnosis: Dugesiidae with very slender body and rotund head. Asymmetrical penis papilla with oblique or almost vertical orientation, when non-extended. Ejaculatory duct with a distinctly subterminal opening at the anterior or antero-ventral side of the penis papilla and surrounded by a well-developed coat of circular muscle. Testes dorsal, distributed throughout the body length. Intrabulbar seminal vesicle surrounded by well-developed coat of interwoven muscle. Common oviduct, opening onto ventral, horizontal and broadened section of the bursal canal, which receives the openings of shell glands anteriorly to the oviducal opening. Bursal canal covered with a coat of circular muscle.

Type species: *Recurva postrema* Sluys & Solà sp. nov.

Etymology: The generic name is derived from the Latin adjective *recurvus*, bent backwards, and alludes to the situation that the ejaculatory duct curves backwards to such an extent that its opening is located at the antero-ventral side of the penis papilla.

Gender: female.

RECURVA POSTREMA SLUYS & SOLÀ **SP. NOV.** (FIGS 14–16)

Material examined: Holotype: ZMA V.Pl. 7122.1, NE Laerma, Rhodes, Greece, 36°10′6.76″N, 27°57′34.55″E, alt. 135 m, 5 April 2009, coll. Eduardo Mateos and Eduard Solà, sagittal sections on six slides.

Paratypes: ZMA V.Pl. 7122.2, ibid., sagittal sections on four slides (not fully mature specimen); V.Pl. 7122.3, ibid., sagittal sections on six slides; V.Pl.7122.4, ibid., sagittal sections on four slides; V.Pl. 7122.5, ibid., sagittal sections on seven slides; V.Pl. 7122.6, ibid., horizontal sections on four slides; V.Pl. 7122.7, ibid., sagittal sections on six slides; V.Pl. 7122.8, ibid., sagittal sections on eight slides; V.Pl. 7122.9, ibid., sagittal sections on six slides.

Etymology: The specific epithet is derived from the Latin adjective *postremus*, located posteriorly, and alludes to the far posteriorly located position of the copulatory apparatus.

Diagnosis: Animals slender, with rotund head. Pharynx and copulatory apparatus situated in the far posterior end of the body. Dorsal testes, distributed throughout the body length but anteriormost testes located at a considerable distance behind the brain. Vasa deferentia open asymmetrically into intrabulbar seminal vesicle. Penis papilla asymmetrical, with more or less vertical orientation in the male atrium.

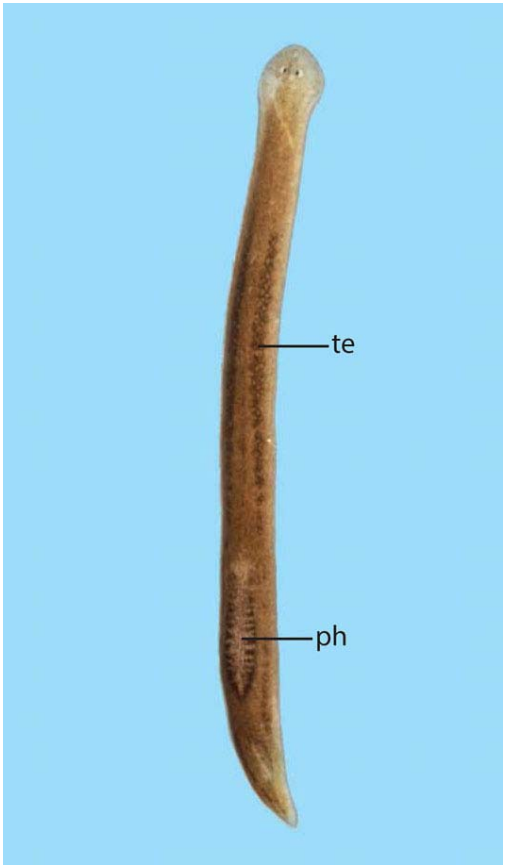


Figure 14. ***Recurva postrema*** Sluys & Solà **sp. nov.** Photograph of external features (scale bar not available).

Ejaculatory duct opening at the anterior or ventro-anterior side of the penis papilla. Ventral or ventro-anterior, muscular penial fold present at the point of insertion of the penis papilla. Ovaries located at about 1/4th the distance between the brain and the root of the pharynx. Distal, posterior parts of the oviducts increase in diameter before communicating with an equally wide common oviduct. Bursal canal is surrounded by a well-developed coat of circular muscle.

Ecology and distribution: The species is known only from the type locality, where it was collected from stagnant water in a rather dry creek. Specimens were found in high numbers, gliding on the substrate, together with other small, white flatworms of an unknown species.

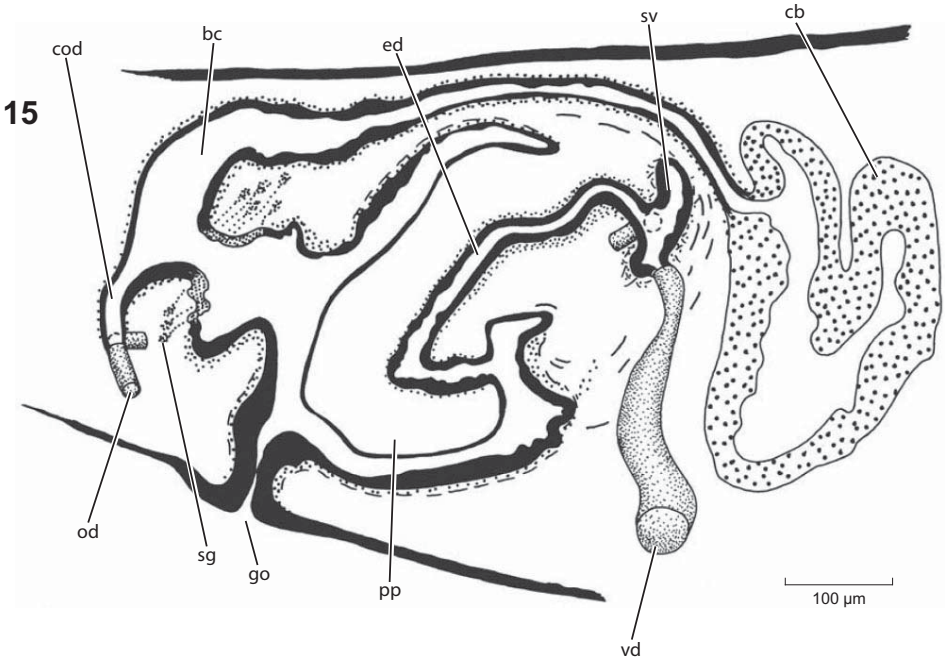
Description: Preserved specimens measure up to 9.5 mm in length and 2.25 mm in width. Notably live specimens are very slender (Fig. 14), with a rotund head that is provided with a pair of close-set eyes, situated in pigment-free patches. Each eye cup houses numerous retinal cells. Behind the eyes, along the lateral margins of the body, there is an auricular streak on either side, at the level of which the head narrows so that there is a more slender neck region. The dorsal surface is finely pigmented pale brown, with notable accumulations of pigment around the pharyngeal pocket. Ventral surface pale.

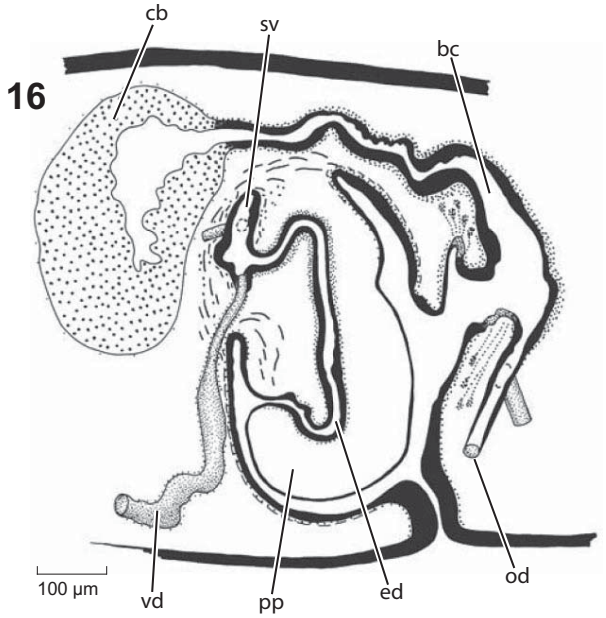
The pharynx measures between 1/6th and 1/8th of the body length and is positioned far into the posterior part of the body. The musculature of the pharynx conforms to the planariid type. This highly posterior location of the pharynx means that the copulatory apparatus is pushed far into the tail end of the animal. The mouth opening is located at the posterior end of the pharyngeal pocket.

The testes are located dorsally, extending from directly behind the ovaries to almost the posterior margin of the body. After having penetrated the penis bulb, the vasa deferentia open separately into the intrabulbar seminal vesicle. The openings of the seminal ducts are asymmetrical in that one vas deferens opens into the ventral section and the other in a more dorsal section of the seminal vesicle (Figs 15, 16). The latter, lined with a nucleated epithelium and surrounded by a coat of intermingled muscle, communicates with the ejaculatory duct, which in most of the specimens examined exhibits an S-shaped loop before curving downwards to follow its central course through the penis papilla. The papilla is more or less cylindrical in shape and has a more or less vertical orientation in the male atrium. The penis papilla is highly asymmetrical in the sense that in its distal, ventral section the ejaculatory duct shows a sharp, anteriorly directed, knee-shaped bend, after which it opens at the anterior or ventro-anterior side of the penis papilla (Figs 15, 16). This course of the ejaculatory duct results in the situation that the anterior portion or lip of the penis papilla is shorter and smaller, in some specimens much shorter and smaller, than the posterior section. At the base of this anterior or ventro-anterior lip of the penis papilla, at its point of insertion, a penial fold is present. This fold is characterized by a more or less developed outbulging and is provided with its own musculature. It is a penial fold, in contrast to an atrial fold, because it is located entally to the point of attachment of the musculature of the penis bulb. The penis papilla is covered with a thin, nucleated epithelium.

The ovaries are located at about 1/4th the distance between the brain and the root of the pharynx. This implies that also the row of testes starts at a considerable distance posterior to the brain, as may be observed even in living specimens (Fig. 14).

Directly posterior to the gonopore the oviducts turn dorso-medially, while their diameter increases





Figures 15, 16. ***Recurva postrema*** Sluys & Solà **sp. nov. 15.** ZMA V.Pl. 7122.4. Sagittal reconstruction of the copulatory apparatus. **16.** ZMA V.Pl. 7122.1. Sagittal reconstruction of the copulatory apparatus.

considerably. Subsequently, the oviducts fuse to form a common oviduct, with an equally wide diameter, that opens into the ventral section of the bursal canal. The latter starts at the copulatory bursa as a rather narrow duct that gradually widens and posterior to the gonopore makes a sharp anteriorly directed bend before opening into the rather dorsal section of the atrium. The more or less horizontally running and widened part of the bursal canal receives the openings of the shell glands anteriorly to the opening of the common oviduct. The nucleated bursal canal is surrounded by a well-developed coat of circular muscle. The copulatory bursa sits immediately anterior to the penis bulb.

RECURVA CONJUNCTA SLUYS **SP. NOV.**
(FIG. 17)

Material examined: Holotype: ZMA V.Pl. 7123.1, near Agios Georgios, Kefalonia, Greece, 38°6′0.72″N, 20°44′55.50″E, 26 April 2009, coll. R. Sluys, sagittal sections of the anterior, prepharyngeal end of the animal on six slides; V.Pl. 7123.1, ibid., sagittal sections of the posterior end (including the pharynx) of the same animal on six slides.

Etymology: The specific epithet is derived from the Latin adjective *coniunctu*s, connected, and alludes to the genito-intestinal connection present in this species.

Diagnosis: Animals slender, with rotund head. Dorsal testes, distributed throughout the body length. Vasa deferentia narrow when penetrating the ventro-lateral side of the penis bulb, subsequently expanding again and opening into the mid-lateral section of the intrabulbar seminal vesicle. Asymmetrical penis papilla, with an oblique, ventro-posterior orientation. Ejaculatory duct opening at the antero-ventral side of the penis papilla. Common oviduct surrounded by a coat of circular muscle. Copulatory bursa communicating with a branch of the intestine. Bursal canal surrounded by a layer of circular muscle.

Ecology and distribution: The species is known only from its type locality, where it was found under stones in an almost dry, muddy stream flowing beneath a concrete bridge.

Description: In the field the two specimens collected (one immature) were identified as *Schmidtea*-like animals, i.e. with a rounded head. The animals were very slender, the holotype specimen measuring up to 2 cm in length when fully stretched and moving. Dorsal surface pigmented, ventral surface pale (as deduced from examination of the sections). Each eye cup houses numerous retinal cells.

The pharynx measures about 1/9th of the body length, its root being situated about half-way along the body length. The mouth opening is located at the posterior end of the pharyngeal cavity.

The testes are situated dorsally, extending from directly behind the brain into the posterior end of the body. The vasa deferentia, which are expanded to spermiducal vesicles, narrow considerably when they penetrate the ventro-lateral side of the penis bulb. Once within the bulb, the ducts expand again in diameter and, subsequently, open into the mid-lateral section of the intrabulbar seminal vesicle. The latter is lined with a nucleated epithelium and surrounded by a rather thick coat of interwoven muscles.

The ejaculatory duct arises from the dorsal section of the seminal vesicle and immediately thereafter

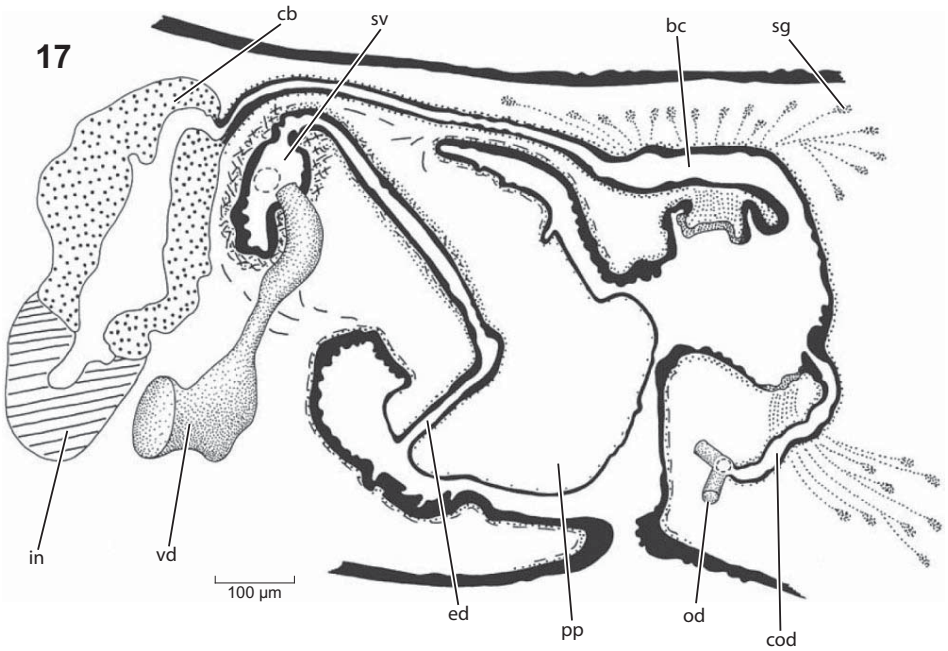


Figure 17. ***Recurva conjuncta*** Sluys **sp. nov.** ZMA V.Pl. 7123.1. Sagittal reconstruction of the copulatory apparatus.

sharply curves postero-ventrally to continue its more or less central course through the penis papilla. However, at some point the ejaculatory duct makes another sharp, hooked bend towards the antero-ventral surface of the body (Fig. 17). This results in the situation that the duct opens at the antero-ventral side of the penis papilla. The papilla has an oblique, ventro-posterior orientation and is covered with a nucleated epithelium. Because of the peculiar course of the ejaculatory duct, the distal section of the penis papilla is highly asymmetrical, with a short ventral lip and a bulky dorsal lip. In fact, the tip of the papilla is to some extent also curved towards the lateral side of the male atrium. Therefore, the opening of the ejaculatory duct is not only displaced towards the antero-ventral side of the penis papilla but also to a more lateral position. This lateral twist of the tip of the penis papilla may be due to a preservation artefact. The major portion of the ejaculatory duct is surrounded by a relatively thick layer of mostly circular muscle fibres.

The paired ovaries are situated directly behind the brain. Immediately posterior to the gonopore the oviducts turn medially and fuse to form a common oviduct, which opens at the postero-ventral section of the bursal canal. The common oviduct is surrounded by a coat of circular muscle.

The bursal canal arises as a broad duct from the mid-posterior wall of the atrium. This first, broad section of the canal runs more or less horizontally and receives the openings of the abundant shell glands, which open anteriorly to the opening of the common oviduct. This broad part of the bursal canal narrows considerably and, subsequently, curves forwards to continue its course immediately dorsally to the male atrium and the penis bulb. Half-way along its course the canal becomes even narrower before communicating with the copulatory bursa. The entire bursal canal is lined with a nucleated epithelium and is surrounded by a layer of circular muscle.

The copulatory bursa lies immediately anterior to the penis bulb, while its ventral part is connected with a branch of the intestine.

Comparative discussion of *Recurva*

The new genus *Recurva* shows a combination of morphological features that sets it apart from all dugesiid genera known at present, albeit that the rounded head, the muscular intrabulbar seminal vesicle and the muscular ejaculatory duct remind one of the genus *Schmidtea*. However, *Schmidtea* is characterized (a) by two muscular seminal vesicles, while (b) its bursal canal is surrounded by a coat of intermingled muscles (characters 18-1 and 22-2, respectively in Sluys, 2001: fig. 7.15), and (c) by separate oviducal openings into the bursal canal. *Recurva* again resembles *Schmidtea* in the dorsally displaced opening of the bursal canal into the atrium. However, such a dorsally displaced communication between bursal canal and atrium is also present in the genus *Cura* Strand, 1942. There are also some other resemblances between *Cura* and *Recurva*, notably (1) the presence of a common oviduct, and (2) the situation that the shell glands open into the section of the bursal canal that lies between its point of communication with the atrium and the point where the canal receives the opening of the common oviduct. However, in other features there is not much resemblance between *Recurva* and *Cura*.

The phylogenetic analyses based on 18S + COI (Fig. 1) and COI alone (Fig. S1) also clearly show that *Recurva* groups independently from the genera *Dugesia*, *Schmidtea*, *Girarda* and *Cura*, and that the species from Rhodes groups closely with the species from Kefalonia. Interestingly, asexual specimens from Paros form the sister group of *Recurva postrema* and *R. conjuncta*, thus constituting an independent lineage. Although we have not performed a molecular species delimitation analysis, this situation nevertheless suggests the presence of a third species of *Recurva* on this island. The external appearance of the Paros animals is very similar to *R. postrema* and *R. conjuncta* in that the animals are also very slender, with rounded head. The Paros specimens (Fig. 18) have their pharynx located in the far posterior region of the body, as is the case also in *R. postrema*. We do here consider the putative third species of *Recurva* from Paros to be a UCS.

A comparison between *Recurva postrema* and *R. conjuncta* reveals clearly that they represent different, species-specific variations on the Bauplan of the genus *Recurva*. Animals of *R. postrema* from Rhodes can be differentiated immediately by the situation that the pharynx and the copulatory apparatus are shifted very far into the posterior end of the body; such is not the case in *R. conjuncta*. Furthermore, *R. postrema* possesses a ventral penial fold, which is absent in *R. conjuncta*. Other differences between the two species concern the asymmetrical openings of the vasa deferentia into the seminal vesicle of *R. postrema*, the fact that the distal sections of its oviducts expand before communicating with the equally wide common oviduct, and the presence of a genito-intestinal connection in *R. conjuncta*.

GENERAL DISCUSSION

Although in the past several papers have been published on the biodiversity of dugesiid freshwater planarians in the Mediterranean region (see above), our study of only the north-eastern Mediterranean

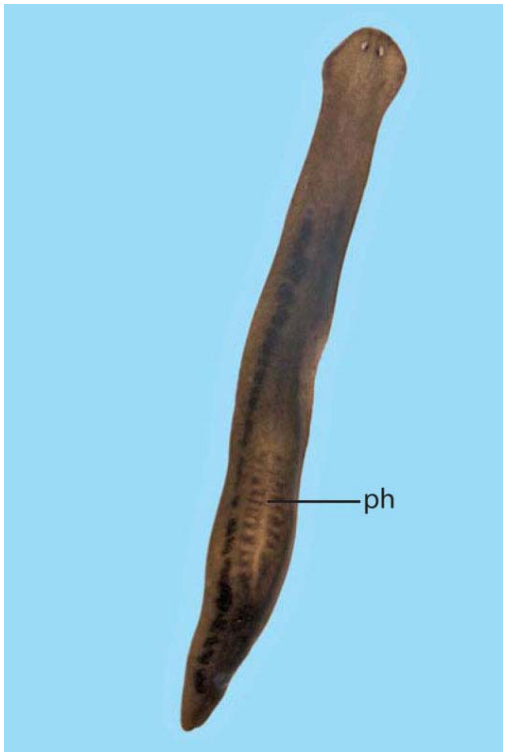


Figure 18. *Recurva* sp. Photograph of external features of specimen from Paros (scale bar not available).

raised the world total of *Dugesia* species with four newly described species, two CCS, 12 UCS from Greece and two more from Slovakia, and at the same time increased the number of dugesiid genera with one new genus, currently comprising two newly described species and one UCS.

Evidently, there is no single objective procedure to delimit higher level taxonomic groups, such as the new genus *Recurva* in the present study. However, the use of genes with a level of variability that results in well-supported and resolved phylogenetic trees (such as 18S rDNA and COI) generally suffices to detect lineage independence. Molecular monophyly combined with the presence of distinct morphological differences subsequently allows for a robust delimitation of higher taxonomic groups, as was the case with *Recurva*.

Although it was not the focus of our present study, it is noteworthy that our results suggest the presence of two UCS of *Dugesia* in Central Europe (Slovakia; entities 30 and 31) that are different from the *D. gonocephala* specimens included in our analysis (Fig. 2, Table 1). Generally, central and northern European specimens of *Dugesia* are assigned to the species *D. gonocephala* as the species has been established to occur with certainty in Denmark, Germany, the Netherlands, Belgium, France, Austria, Bulgaria and the Former Yugoslav Republic of Macedonia (De Vries, 1986). But the fact that Ude's (1908) presumed *D. gonocephala* specimens from Austria differed in certain respects from *D. gonocephala sensu stricto* may have already foreshadowed the possible presence of other species of *Dugesia* in Central Europe, as is now also suggested by our study.

Our study also shows that the planarian diversity of a rather well-researched region such as the Mediterranean remains grossly underestimated and that such must apply to an even greater degree to the global species richness of these animals.

The integrative approach detailed above revealed the beneficial effect of reciprocal illumination of morphological and genetic data in triclads. These different types of data complement each other by pointing out ambiguities or unstable hypotheses on the basis of only a single character set. For example, the gross morphology of *D. effusa* is very similar to that of *D. sagitta* and *D. improvisa*. However, the GMYC analysis delimits *D. effusa* as a different species from *D. improvisa*, while in the phylogenetic trees *D. effusa* is not closely related to *D. sagitta*. Thus, molecular information supports the presumed species status of *D. effusa* that was suggested by the morphological data.

In another case, the opposite situation applied. Molecular data suggested a separate identity for *D. parasagitta* populations on Corfu. As a consequence, more detailed morphological investigations were started, which uncovered some divergent morphological characters with *D. sagitta* as described in the literature and as revealed after examination of both new material and museum specimens. The two data sets thus reinforced each other and induced us to describe the new species *D. parasagitta*.

Although conflicts between datasets can be expected in an integrative taxonomic study because speciation is not always accompanied by simultaneous character change at all levels (Padial & de la Riva, 2009; Padial *et al.*, 2010), our analysis of *Dugesia* actually revealed in many cases a good correspondence between species boundaries hypothesized on morphological data and those suggested by molecular data. As these different lines of evidence generally converged in the delimitation of the same units of biotic diversity, the species taxa recognized can be considered stable systematic hypotheses. For example, we found full correspondence between the GMYC analysis and the morphology-based species hypotheses concerning units from the Eastern and Central Aegean region, namely *D. ariadnae*, *D. damoae*, *D. effusa*, *D. elegans* and *D. improvisa*. However, in other cases we have found DCLs (as in *D. cretica*) or potential cryptic species (as in *D. sagitta*) in which morphology and molecules do not fully correspond. The situation that GMYC can potentially overestimate the number of species (Lohse,

2009) and that we used only a single gene marker has made us refrain from proposing new species solely on the basis of molecular divergence.

Two important conclusions can be drawn from our study. First, despite the fact that we used only a single molecular marker in the present study, GMYC analysis with COI turns out to form a good strategy for detecting potentially new species and for testing the taxonomic status of known species. Second, the morphological features generally used by taxonomists in their comparative studies of dugesiid flatworms indeed result in reliable identifications and delineations of species taxa, at least when no cryptic species are involved, in which case the use of other types of data is unavoidable. This is a comforting insight because it is to be expected that morphological characters will '. . . retain an outstanding role in taxonomy . . .' (Padial & de la Riva, 2010: 753).

ACKNOWLEDGEMENTS

This research was supported by Grant BES-2009-022530 from the Ministerio de Ciencia e innovación to E.S., and by grants CGL2008-00378 and CGL2011-23466 to M.R. Completion of this study was made possible by a grant from the Naturalis Biodiversity Center to R.S. We are indebted to Núria Bonada, Joan Solà, Dora Vázquez and Caterina Rodríguez who provided material and helped with sampling. We are also indebted to Marta Álvarez-Presas for providing the 18S *D. sicula* sequence and to Gema Blasco and Eva Làzaro for laboratory assistance. Professor Dr. M. Kawakatsu (Sapporo) is thanked for nomenclatural advice. We are grateful to Mr J. van Arkel (University of Amsterdam) for the digital rendering of the figures. Mr N. Korenhof (Naturalis Biodiversity Center) is thanked for digitally finalizing the plates.

REFERENCES

Avise JC. 2000. *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard University Press.

Baguñà J, Carranza S, Pala M, Ribera C, Giribet G, Arnedo M, Ribas M, Riutort M. 1999. From morphology and karyology to molecules. New methods for taxonomical identification of asexual populations of freshwater planarians. A tribute to Professor Mario Benazzi. *Italian Journal of Zoology* **66:** 207–214.

Ball IR. 1974. A contribution to the phylogeny and biogeography of freshwater triclads (Platyhelminthes: Turbellaria). In: Riser NW, Morse MP, eds. *Biology of the Turbellaria*. New York: McGraw-Hill, 339–401.

Ball IR. 1979. The karyotypes of two *Dugesia* species from Corfu, Greece (Platyhelminthes, Turbellaria). *Bijdragen tot de Dierkunde* **48:** 187–190.

Bauer AM, Parham JF, Brown RM, Stuart BL, Grismer L, Papenfuss TJ, Böhme W, Savage JM, Carranza S, Grismer JL, Wagner P, Schmitz A, Ananjeva NB, Inger RF. 2011. Availability of new Bayesian-delimited gecko names and the importance of character-based species descriptions. *Proceedings of the Royal Society B* **278:** 490–492.

Benazzi M, Ball IR. 1972. The reproductive apparatus of sexual specimens from fissiparous populations of *Fonticola morgani* (Tricladida, Paludicola). *Canadian Journal of Zoology* **50:** 703–704.

Benazzi M, Banchetti R. 1972. Descrizione di *Dugesia biblica*, nuova microspecie del 'gruppo *Dugesia gonocephala*' trovata nel fiume Giordano (Israele). *Atti della Società Toscane di Scienze Naturali, Memorie, Series B* **79:** 83–91.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I. 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* **22:** 148–155.

Burbrink FT, Yao H, Ingrasci M, Bryson JRW, Guiher TJ, Ruane S. 2011. Speciation at the Mogollon Rim in the Arizona Mountain Kingsnake (*Lampropeltis pyromelana*). *Molecular Phylogenetics and Evolution* **60:** 445–454.

Carew ME, Pettigrove V, Hoffmann AA. 2005. The utility of DNA markers in classical taxonomy: using cytochrome oxidase I markers to differentiate Australian *Cladopelma* (Diptera: Chironomidae) midges. *Annals of the Entomological Society of America* **98:** 587–594.

Charni M, Harrath AH, Sluys R, Tekaya S, Zghal F. 2004. The freshwater planarian *Dugesia sicula* Lepori, 1948 (Platyhelminthes, Tricladida) in Tunisia: ecology, karyology, and morphology. *Hydrobiologia* **517:** 161–170.

Cracraft J. 1983. Species concepts and speciation analysis. *Current Ornithology* **1:** 159–187.

Cracraft J. 1987. Species concepts and the ontology of evolution. *Biology and Philosophy* **2:** 329–346.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9:** 772.

De Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* **56:** 879–886.

De Vries EJ. 1984. On the species of the *Dugesia gonocephala* group (Platyhelminthes, Turbellaria, Tricladida) from Greece. *Bijdragen tot de Dierkunde* **54:** 101–126.

De Vries EJ. 1986. On the taxonomic status of *Dugesia gonocephala* and *Dugesia subtentaculata* (Turbellaria, Tricladida, Paludicola). *Journal of Zoology, London* **209:** 43–59.

De Vries EJ. 1988. On the identity and occurrence of *Dugesia sicula* and *D. biblica* (Platyhelminthes, Tricladida, Paludicola) in the Mediterranean region. *Fortschritte der Zoologie* **36:** 405–411.

Draparnaud J. 1801. *Tableau des Mollusques terrestres et fluviatiles de la France*. Paris: Montpellier.

Drummond A, Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7:** 214.

Dugès A. 1830. Aperçu de quelques observations nouvelles

sur les Planaires et plusieurs genres voisins. *Annales des Sciences Naturelles* **21:** 70–91.

Edwards SV. 2009. Is a new and general theory of molecular systematics emerging? *Evolution* **63:** 1–19.

Ezard T, Fujisawa T, Barraclough T. 2009. *Splits: species' limits by threshold statistics*. R package version 1.0. Available at: http://R-Forge.R-project.org/projects/splits/

Fontaneto D, Herniou EA, Boschetti C, Caprioli M, Melone G, Ricci C, Barraclough TG. 2007. Independently evolving species in asexual bdelloid rotifers. *PLoS Biology* **5:** e87.

Fouquet A, Gilles A, Vences M, Marty C, Blanc M, Gemmell NJ. 2007. Underestimation of species richness in Neotropical frogs revealed by mtDNA analyses. *PLoS ONE* **2:** e1109.

Frankham R, Ballou JD, Dudash MR, Eldridge MDB, Fenster CB, Lacy RC, Mendelson III JR, Porton IJ, Ralls K, Ryder OA. 2012. Implications of different concepts for conserving biodiversity. *Biological Conservation* **153:** 25–31.

Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology and Evolution* **27:** 480–488.

Girard C. 1850. A brief account of the freshwater Planariae of the United States. *Proceedings of the Boston Society of Natural History* **3:** 264–265.

Harrath AH, Sluys R. 2013. In: Harrath AH, Sluys R, Aldahmash W, Errazagui AK, Ibrahim K, Alwasel S. 2013. Reproductive strategies, karyology, parasites, and taxonomic status of *Dugesia* populations from Yemen (Platyhelminthes, Tricladida, Dugesiidae). *Zoological Science* **30:** 502–508.

Harrath AH, Sluys R, Aldahmash W, Errazagui AK, Ibrahim K, Alwasel S. 2013. Reproductive strategies, karyology, parasites, and taxonomic status of *Dugesia* populations from Yemen (Platyhelminthes, Tricladida, Dugesiidae). *Zoological Science* **30:** 502–508.

Hey J. 2009. On the arbitrary identification of real species. In: Butlin RK, Bridle J, Schluter D, eds. *Speciation and patterns of diversity*. Cambridge: Cambridge University Press, 15–28.

Humason GL. 1967. *Animal tissue techniques*, 2nd edn. San Francisco: W. H. Freeman.

Hyman L. 1943. Endemic and exotic land planarians in the United States with a discussion of necessary changes of names in Rhynchodemidae. *American Museum Novitates* **1667:** 1–21.

ICZN, International Commission on Zoological Nomenclature. 1999. *International code of zoological nomenclature*, 4th edn. London: The International Trust for Zoological Nomenclature.

ITZN, International Trust for Zoological Nomenclature. 1985. *International code of zoological nomenclature*, 3rd edn. Berkeley and Los Angeles, CA: University of California Press.

Kaburaki T. 1918. Freshwater triclads from the basin of the Inle Lake. *Records of the Indian Museum* **14:** 187–194 + 1 pl.

Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9:** 286–298.

Komárek J. 1925. Die O. Schmidt'schen Süsswassertricladen von Corfu und Cephalonia, gesammelt von Prof. Wilhelmi. *Zoologischer Anzeiger* **63:** 322–328.

Lang A. 1884. *Die Polycladen (Seeplanarien) des Golfes von Neapel und der angrenzenden Meeresabschnitte*. Fauna und Flora des Golfes von Neapel, Monographie 11. Leipzig: W. Engelmann.

Lázaro EM, Sluys R, Pala M, Stocchino GA, Baguñà J, Riutort M. 2009. Molecular barcoding and phylogeography of sexual and asexual freshwater planarians of the genus *Dugesia* in the Western Mediterranean (Platyhelminthes, Tricladida, Dugesiidae). *Molecular Phylogenetics and Evolution* **52:** 835–845.

Lepori NG. 1948. Descrizione di *Dugesia sicula*, nuova sp. di Triclade d'acqua dolce dei dintorni di Catania. *Archivio Zoologico Italia* **33:** 461–472.

Lepori NG. 1951. Sulla caratteristiche morfologiche e sulla posizione sistematica della planaria di Sardegna e Corsica già ascritta a *Dugesia gonocephala* (Dugès). *Atti della Società Toscane di Scienze Naturali, Memorie, Series B* **58:** 28–47.

Lohse K. 2009. Can mtDNA barcodes be used to delimit species? A response to Pons *et al*. (2006). *Systematic Biology* **58:** 439–442.

Meixner J. 1928. Der Genitalapparat der Tricladen und seine Beziehungen zu ihrer allgemeinen Morphologie, Phylogenie, Ökologie und Verbreitung. *Zeitschrift für Morphologie und Ökologie der Tiere* **11:** 570–612.

Memon N, Meier R, Manan A, Su KF. 2006. On the use of DNA sequences for determining the species limits of a polymorphic new species in the stink bug genus *Halys* (Heteroptera: Pentatomidae) from Pakistan. *Systematic Entomology* **31:** 703–710.

Monaghan MT, Wild R, Elliot M, Fujisawa T, Balke M, Inward DJG, Lees DC, Ranaivosolo R, Eggleton P, Barraclough TG, Vogler AP. 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology* **58:** 298–311.

Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial gene trees versus nuclear-gene trees. *Evolution* **49:** 718–726.

Nygren A, Pleijel F. 2011. From one to ten in a single stroke – resolving the European *Eumida sanguinea* (Phyllodocidae, Annelida) species complex. *Molecular Phylogenetics and Evolution* **58:** 132–141.

Olson LE, Goodman SM, Yoder AD. 2004. Illumination of cryptic species boundaries in long-tailed shrew tenrecs (Mammalia: Tenrecidae; Microgale), with new insights into geographic variation and distributional constraints. *Biological Journal of the Linnean Society* **83:** 1–22.

Padial J, Miralles A, de la Riva I, Vences M. 2010. The integrative future of taxonomy. *Frontiers in Zoology* **7:** 16.

Padial JM, de la Riva I. 2009. Integrative taxonomy reveals cryptic Amazonian species of *Pristimantis* (Anura: Strabomantidae). *Zoological Journal of the Linnean Society* **155:** 97–122.

Padial JM, de la Riva I. 2010. A response to recent proposals for integrative taxonomy. *Biological Journal of the Linnean Society* **101:** 747–756.

Pala M, Stocchino GA, Corso G, Casu S. 2000. A new species of freshwater planarian belonging to the genus *Dugesia* (Platyhelminthes, Tricladida) from Sardinia. *Italian Journal of Zoology* **67:** 379–383.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* **55:** 595–609.

Rambaut A, Drummond AJ. 2007. Tracer v1.4. Available at: beast.bio.ed.ac.uk/Tracer

Romeis B. 1989. *Mikroskopische Technik*, 17th edn. München: Urban und Schwarzenberg.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61:** 539–542.

Schmidt O. 1861. Untersuchungen über Turbellarien von Corfu und Cephalonia, nebst Nachträgen zu früheren Arbeiten. *Zeitschrift für wissenschaftliche Zoologie* **11:** 1–32.

Sluys R. 1991. Species concepts, process analysis, and the hierarchy of nature. *Experientia* **47:** 1162–1170.

Sluys R. 2001. Towards a phylogenetic classification and characterization of dugesiid genera (Platyhelminthes, Tricladida, Dugesiidae): a morphological perspective. In: Littlewood DTJ, Bray RA, eds. *Interrelationships of the Platyhelminthes*, Syst. Ass. Spec. Vol. ser. 60. London: Taylor & Francis, 57–73.

Sluys R. 2007. Annotations on freshwater planarians (Platyhelminthes Tricladida Dugesiidae) from the Afrotropical Region. *Tropical Zoology* **20:** 229–257.

Sluys R. 2013. The unappreciated, fundamentally analytical nature of taxonomy and the implications for the inventory of biodiversity. *Biodiversity and Conservation* **22:** 1095–1105.

Sluys R, Kawakatsu M, Winsor L. 1998. The genus *Dugesia* in Australia, with its phylogenetic analysis and historical biogeography (Platyhelminthes, Tricladida, Dugesiidae). *Zoologica Scripta* **27:** 273–289.

Solà E, Sluys R, Gritzalis K, Riutort M. 2013. Fluvial basin history in the northeastern Mediterranean Region underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae). *Molecular Phylogenetics and Evolution* **66:** 877–888.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22:** 2688–2690.

Stevens NM, Boring AM. 1906. *Planaria morgani* n. sp. *Proceedings of the Academy of Natural Sciences of Philadelphia* **58:** 7–9 + 1 pl.

Stocchino GA, Corso G, Manconi R, Pala M. 2002. African planarians: *Dugesia aethiopica* sp. n. (Platyhelminthes, Tricladida) from Lake Tana (NW Ethiopia). *Italian Journal of Zoology* **69:** 45–51.

Stocchino GA, Sluys R, Manconi R. 2012. A new species of *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae) from the Afromontane forest in South Africa, with an overview of freshwater planarians from the African continent. *Zootaxa* **3551:** 43–58.

Strand E. 1942. Miscellanea nomenclatoria zoologica et palaeontologica. *Folia Zoologica et Hydrobiologica* **11:** 386–402.

Talavera G. 2012. Phylogenetic inference at different insect taxonomic levels. Unpublished D. Phil. Thesis, Universitat Autònoma de Barcelona.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* **56:** 564–577.

Ude J. 1908. Beiträge zur Anatomie und Histologie der Süßwassertricladen. *Zeitschrift für wissenschaftliche Zoologie* **89:** 308–370 + 3 pls.

Vieites DR, Wollenberg KC, Andreone F, Köhler J, Glaw F, Vences M. 2009. Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proceedings of the National Academy of Sciences USA* **106:** 8267–8272.

Weiss A. 1909. Turbellaria, Rhabdocoelida and Tricladida paludicola. In: Michaelsen W, Hartmeyer R, eds. *Die Fauna Südwest-Australiens: Ergebnisse der Hamburger südwest-australischen Forschungsreise 1905*. Jena: Verlag von Gustav Fischer, 403–410.

Xia X, Xie Z. 2001. DAMBE: software package for data analysis in molecular biology and evolution. *Journal of Heredity* **92:** 371–373.

Yeates DK, Seago A, Nelson L, Cameron SL, Joseph L, Trueman JWH. 2011. Integrative taxonomy, or iterative taxonomy? *Systematic Entomology* **36:** 209–217.

SUPPORTING INFORMATION

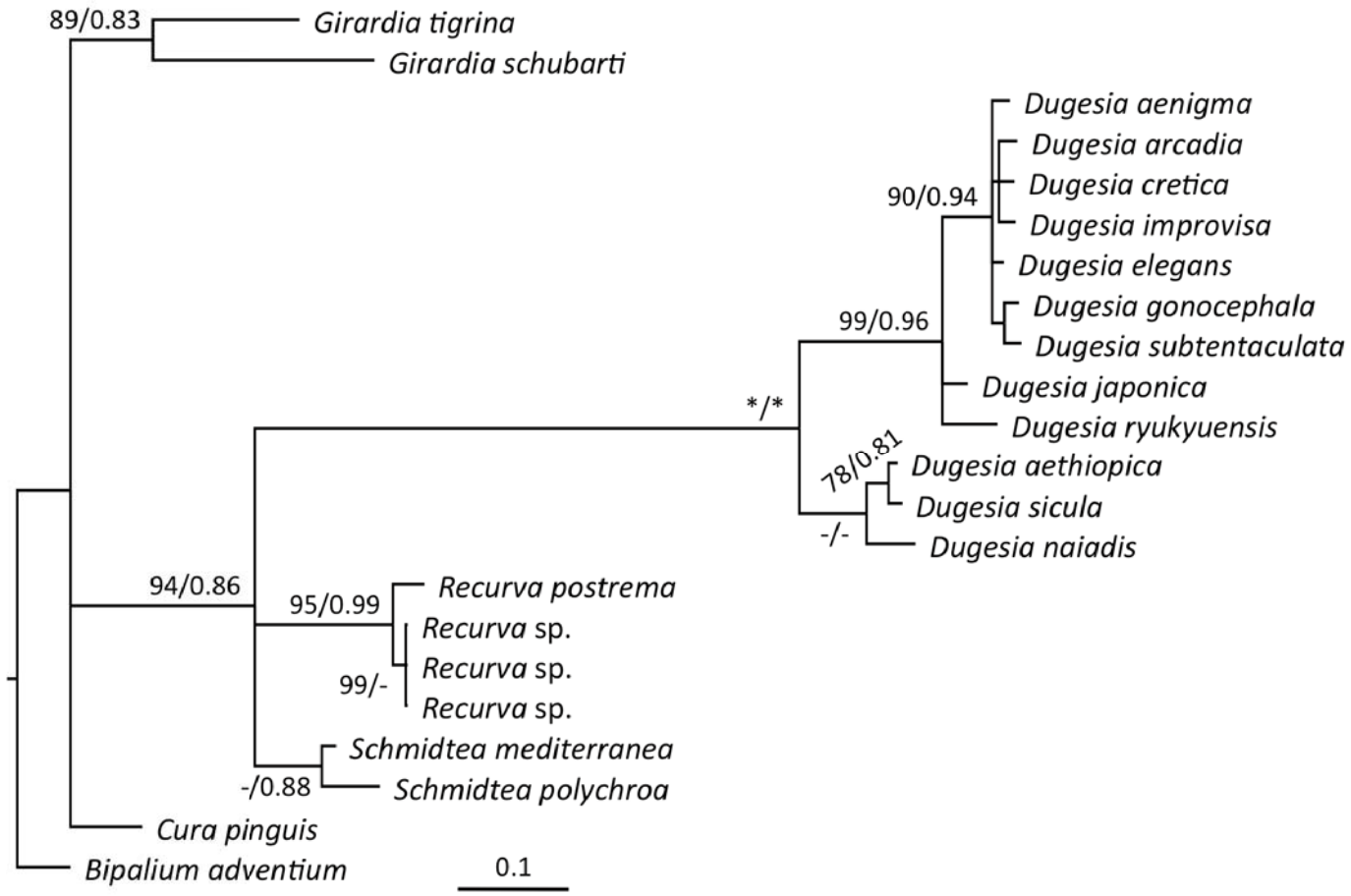
Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Sampling localities and their various codes.
Table S2. Sampling localities of specimens of ***Recurva*** Sluys **gen. nov**.
Table S3. Forward and reverse primers used in amplification and sequencing.
Table S4. Species and genes used in the phylogenetic analysis.
Table S5. Species status of asexual *Dugesia* from Chios.
Figure S1. Bayesian tree inferred from the COI data set.
Figure S2. Location of the *Dugesia* sampling sites on Corfu.

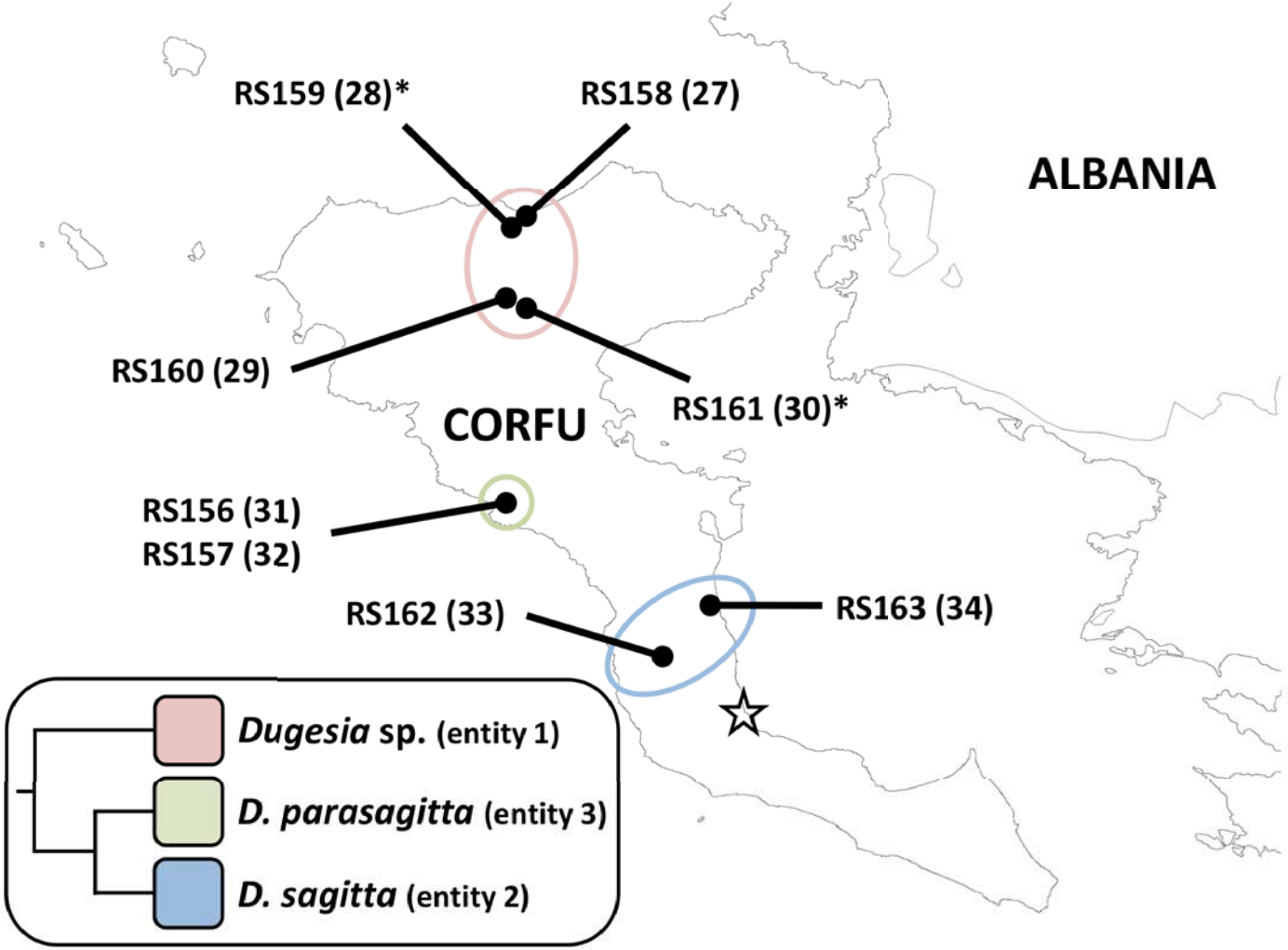
Supplementary Information

Supplementary Figures & Supplementary Tables

Supporting Figure 1. Bayesian tree inferred from the COI dataset. Labels correspond to species names. Node numbers correspond to bootstrap (ML)/posterior probability (BI); values are only indicated when >50/>0.80, "*" indicates maximum support. Scale bar indicates number of substitutions per site.



Supporting Figure 2. Location of the *Dugesia* sampling sites on Corfu, Greece. Populations are labeled with the field code and with locality codes (in parentheses) as used in a companion paper (Solà et al., 2013). Localities with an asterisk have not been used in the molecular analyses of the present study. Populations belonging to different molecular clades are surrounded with different colors, these are in correpondance with the inset showing the relationships among the entities obtained with the GMYC analysis. Neotype locality indicated by an open star.



Appendix S1: Supplementary Table 1. Sampling localities and their various codes.

Code Solà et al., 2013	Sampling locality	No. of individuals per locality (total: 155)	Field code	Museum catalogue N°	GenBank Acc. N°	Lat./Long.
27	Roda, Corfu, GR	5	RS 158	ZMA V.Pl. 7124	KF308786 - KF308790	39°47'23.94"N/19°47'29.46"E
29	Kato vrisi spring, Klimatia, Corfu, GR	5	RS 160	ZMA V.Pl. 7126	KC006996, KF308759 - KF308762	39°44'30.48"N/19°46'49.20"E
33	North of Vouniatades, Corfu, GR	5	RS 162	ZMA V.Pl. 7120	KF308814 - KF308818	39°31'16.33"N/19°52'38.12"E
34	Benitses, Corfu, GR	5	RS 163	ZMA V.Pl. 7121	KF308715 - KF308719	39°32'44.39"N/19°54'35.35"E
31	Ermones, Corfu, GR	4	RS 156	ZMA V.Pl. 7119	KC006995, KC006994, KF308738, KF308741	39°36'37.98"N/19°46'41.64"E
32	Ermones, slightly higher than 31, Corfu, GR	5	RS 157	ZMA V.Pl. 7118	KC006992, KF308736, KF308737, KF308739, KF308740	39°36'41.93"N/19°47'1.40"E
35	Near Agia Eirini, Cephalonia	5	RS 166	ZMA V.Pl. 7142	KC006963, KF308699 - KF308702	38° 7'34.92"N/20°44'31.62"E
36	Digaleto, Cephalonia	5	RS 170	ZMA V.Pl. 7143	KF308724 - KF308728	38°10'46.99"N/20°40'46.80"E
17	Potamia, Preveza, GR	5	-		KC007036, KF308781 - KF308784	39°22'37.42"N/20°52'38.41"E
14	Vafkeri, Lefkada, GR	5	-		KC007009, KF308806 - KF308809	38°43'31.41"N/20°39'46.59"E
13	Varia, Aetolia-Acarnania, GR	5	-		KC007011, KF308810 - KF308813	38°35'34.87"N/21°35'11.02"E
12	Eleonas-Gravia, Phocis, GR	4	RS 260	ZMA V.Pl. 7151	KC007018, KC007014, KF308734, KF308735	38°34'29.21"N/22°23'38.50"E
20	Polidrosos, Phoci, GR	2	RS 261	ZMA V.Pl. 7130	KC007022, KC007023,	38°38'4.43"N/22°30'49.69"E
19	Mexiates, Phthiotis, GR	5	RS 263	ZMA V.Pl. 7132	KC006989, KF308775 - KF308778	38°53'4.09"N/22°18'53.16"E
18	Gorgopotamos, Phthiotis, GR	5	RS 262	ZMA V.Pl. 7131	KC006990, KC006991, KF308748 - KF308750	38°49'46.06"N/22°22'53.37"E
16	Filiates, Thesprotia, GR	3	-		KC007028, KC007035, KF308742	39°38'16.09"N/20°23'41.48"E

23	Dorio-Psari, Peloponnese, GR	5	-		KF308729 - KF308733	37°18'29.61"N/21°51'55.96"E
21	Tripi, Peloponnese, GR	5	RS 264	ZMA V.Pl. 7133	KF308801- KF308805	37° 5'38.47"N/22°20'46.29"E
26	Sella, Peloponnese, GR	5	RS 258	ZMA V.Pl. 7129	JN376140, KC006970, KF308795	38°17'3.02"N/21°52'45.80"E
25	Chalandritsa, Peloponnese, GR	3	RS 259	ZMA V.Pl. 7128	KC006971, KF308720 - KF308723	38° 6'31.85"N/21°47'13.73"E
24	Theisoa-Andritsaina, Peloponnese, GR	5	-		KF308707 - KF308711	37°29'13.97"N/21°55'4.88"E
22	Agios Floros, Peloponnese, GR	5	-		KC007029, KF308703 - KF308706	37°10'8.94"N/22° 1'33.92"E
6	Apollonas, Naxos, GR	5	RS 201, RS 202	ZMA V.Pl. 7146, V.Pl. 7147	JN376142, KC006972 KF308712 - KF308714	37° 9'53.96"N/25°32'42.94"E
7	Melanes, Naxos, GR	5	RS 203, RS 204	ZMA V.Pl. 7116	KC006987, KF308771 - KF308774	37° 5'3.38"N/25°26'59.40"E
9	Nagos, Chios, GR	3	RS220	ZMA V.Pl. 7115	KC006981, KF308779, KF308780	38°33'27.73"N/26°4'28.26"E
10	Nagos, before the opening to the sea, Chios, GR	2	RS 221	ZMA V.Pl. 7114	KC006980, KC006982	38°33'34.73"N/26°4'56.86"E
8	Manolates, Samos, GR	5	RS 242	ZMA V.Pl. 7135	KC006978, KF308767 - KF308770	37°47'21.26"N/26°49'17.80"E
11	Kalamoudi, Euboea, GR	3	-		KC007026, KC007010, KF308785	38°49'45.72"N/23°16'53.48"E
5	Petaloudes Valley, Rhodes, GR	2	RS 183	ZMA V.Pl. 7145	KC006985, KC006984	36°20'13.51"N/28° 3'44.90"E
1	Georgioupoli, Crete, GR	5	RS 193, RS 194	ZMA V.Pl. 7136, V.Pl. 7137	KF308743 - KF308747	35°21'37.94"N/24°15'6.51"E
3	Sasalos, Crete, GR	5	RS199, RS 200	ZMA V.Pl. 7140, V.Pl. 7141	KC006977, KF308791 - KF308794	35°24'9.86"N/23°42'42.39"E
2	Kakopetros, Crete, GR	5	RS 197, RS 198	ZMA V.Pl. 7138, V.Pl. 7139	KC006973, KF308751 - KF308754	35°24'29.34"N/23°45'19.23"E
-	Vernár, Slovak Republic	1	-		KC007033	48°55'21.06"N/20°18'34.45"E
-	Ludrová, Slovak Republic	1	-		KC007013	49°1'46.18"N/19°19'49.07"E
-	Prosiek, Slovak Republic	1			KC007030	49°9'15.18"N/19°29'53.64"E
-	Limburg, Netherlands	2	-		FJ646941+FJ646986, FJ646942+FJ646987	
-	Fita-Kimpouries, Chios, GR	4	RS 227	ZMA V.Pl. 7117	KF308755 - KF308758	38°30'43.31"N/25°59'55.06"E
-	Tripes-Parparia, Chios, GR	5	RS 233	ZMA V.Pl. 7152	KF308796 - KF308800	38°31'39.85"N/25°52'6.42"E

Appendix S2: Supplementary Table 2. Sampling localities of *Recurva* specimens.

Species	Sampling locality	Lat./Long.
Recurva postrema	North East of Laerma, Rhodes, Greece	36º10'06.76"N/27º57'34.55"E
Recurva conjuncta	North of Pastra, Cephalonia, Greece	38º06'00.72"N/20º44'55.50"E
Recurva sp.	South West of Lefkes, Paros, Greece	37º02'46.58"N/25º11'59.85"E
Recurva sp.	South West of Lefkes, Paros, Greece	37º02'32.19"N/25º11'58.72"E
Recurva sp.	South West of Lefkes, Paros, Greece	37º02'05.19"N/25º12'01.38"E

Appendix S3: Supplementary Table 3. Forward (F) and Reverse (R) primers used in amplification and sequencing.

Name	Sequence 5'-3'	Annealing Temperature (°C)	Source
18S			
18S_1F(F)	TACCTGGTTGATCCTGCCAGTAG	45	Carranza et al., 1996
18S_5R(R)	CTTGGCAAATGCTTTCGC	45	Carranza et al., 1996
18S_4F(F)	CCAGCAGCCGCGCTAATTC	45	Carranza et al., 1996
18S_9R(R)	GATCCTTCCGCAGGTTCACCTAC	45	Carranza et al., 1996
COI			
BarT (F)	ATGACDGCSCATGGTTTAATAATGAT	43	Álvarez-Presas et al., 2011
COIEF3 (F)	CCWCGTGCWAATAATTTRAG	48	Solà et al., 2013.
COIR (R)	CCWGTYARMCCHCCWAYAGTAAA	43	Lázaro et al., 2009

Supplementary Table 4. Species and genes used in the phylogenetic analysis.

Species	18S type II	COI
Outgroup		
Bipalium adventium	DQ666000	AF178306
Ingroup		
Dugesia		
D. aenigma	KF308698	KF308698
D. aethiopica	-	FJ646932+FJ646976
D. arcadia	KF308694	KF308694
D. cretica	KF308697	JN376141
D. naiadis	-	KF308755
D. elegans	KF308695	KC006985
D. gonocephala	DQ666002	FJ646942+FJ646987
D. improvisa	KF308696	KC006987
D. japonica	D83382	DQ666034
D. sicula	KF308693	KF308797
D. subtentaculata	AF013155	FJ646950+FJ646996
D. ryukyuensis	AF050433	AF178311
Girardia		
G. schubarti	DQ666015	DQ666041
G. tigrina	AF013156	DQ666042
Recurva		
R. conjuncta	KF308692	-
R. postrema	KF308691	KF308763
R. sp.	KF308690	KF308765
R. sp.	KF308689	KF308766
R. sp.	KF308688	KF308764
Schmidtea		
S. mediterranea	U31085	JF837061
S. polychroa	AF013154	FJ647021+FJ646975

Appendix S6: Supplementary Table 5. Species status, according to COI gene identity, of asexual *Dugesia* populations from Chios.

Putative species	Sampling locality	Field code	Lat./Long.
D. naiadis	Spring just at the entrance of Kampia	RS225	38°34'22.35"N /25°59'8.47"E
	Near Spartounda	RS226	38°32'51.26"N /25°59'27.95"E
	600 m before Nenitouria	RS232	38°34'9.68"N /25°52'29.81"E
D. sicula	Yosona (Giosonas), spring alongside the road	RS222	38°33'54.21"N /26° 4'36.74"E
	Chori	RS229	38°28'17.05"N /25°56'12.99"E
	Just at the end of Chalandra	RS231	38°32'52.73"N /25°55'51.13"E
	1.8 km after Parparia, in the direction of Volissos	RS234	38°30'58.15"N /25°53'38.46"E

3

Mitogenomes and molecular markers

Chapter 4

Evolutionary analysis of mitogenomes from parasitic and free-living flatworms

Reference

Solà E, Álvarez-Presas M, Frías-López C, Littlewood DTJ, Rozas J, Riutort M. **Evolutionary analysis of mitogenomes from parasitic and free-living flatworms.** *In preparation.*

Summary

The original aim of this work was to sequence the mitogenomes of five different species of Tricladida, including the species *Dugesia subtentaculata*, in order to take profit of the results for developing more markers for our posterior studies. However, the development of this work delayed for a long time and, moreover, we did not succeed in the obtention of the *Dugesia* species mitogenome, but only of the species *Crenobia alpina* (Planariidae) and *Obama* sp. (Geoplanidae). Not allowing the use of the new obtained sequences to develop new markers for this thesis.

We successfully used the next-generation sequence method 454 (Roche) pyrosequencing to obtain the complete mitogenomes of the two mentioned triclad species, thus increasing the total number of available free-living platyhelminthes complete mitogenomes from three to five. The original 454 results were double-checked by traditional PCR after designing primers from the 454 output. Once the new mitochondrial genomes were obtained, we took profit of the new available material and we carried out an analysis on the selective pressure on the nucleotide composition of the Platyhelminthes mitogenomes, including four triclads (*C. alpina*, *Obama* sp., *D. japonica* and *S. mediterranea*) and six neodermatans in order to search for differences in their evolutive selection pressure considering their parasitic or free living life cycle. We expected to find such differences in pressure according to the organisms life cycle since they are related to their putative population sices. Thus, neodermatans or parasitic platyhelminths suffering bottlenecks would present a relaxed selection, while free-living flatworms (triclads) would be under a higher selective pressure.

Surprisingly, the results showed no differences between parasites and free living platyhelminthes. In fact, representatives of the Geoplanoidea (*Obama*, *Dugesia* and *Schmidtea*) presented a higher relaxation in comparison with the parasites (and *Crenobia*). We found mitogenomes to be of great potential for better understanding flatworm evolutionary history whilst this new information will be useful for future phylogenetic, biogeographic and phylogeographic studies by providing new and valuable molecular markers.

Evolutionary analysis of mitogenomes from parasitic and free-living flatworms

Eduard Solà[1], Marta Álvarez-Presas[1], Cristina Frías-López[1], D. Timothy J. Littlewood[2], Julio Rozas[1], Marta Riutort[1*]

[1]Institut de Recerca de la Biodiversitat and Dept. Genètica, Facultat de Biologia, Universitat de Barcelona, Av Diagonal, 643, Barcelona 08028, Spain
[2]Department of Life Sciences, Natural History Museum, Cromwell Road, London, UK

Keywords: Codon bias, *Crenobia*, Geoplanidae, Mitogenome Skew, Parasitic Flatworms, Planariidae, Platyhelminthes, Tricladida, *Obama*.

* Correspondence: mriutort@ub.edu

Abstract

Mitochondrial genomes (mitogenomes) are useful and relatively accessible sources of molecular data to explore and understand the evolutionary history and relationships among different organisms at different levels. The availability of complete mitogenomes from Platyhelminthes is scarce; of the 40 or so published most are from parasitic flatworms (Neodermata). Here, we present the mitogenomes of two new free-living flatworm (Tricladida), the freshwater species *Crenobia alpina* (Planariidae) and the land planarian *Obama* sp. (Geoplanidae). This contribution doubles the total number of Tricladida mtDNAs published. We took the opportunity to conduct comparative mitogenomic analyses between free-living and parasitic flatworms in order to find out whether nucleotide composition and selection between these two groups reflects their life cycle. Unexpectedly we did not found the selective relaxation expected in parasitic species; on the contrary, triclad mitogenomes, exhibit higher A+T content and selective relaxation levels. We show that mitogenomes have great potential for better understanding flatworm evolutionary history whilst providing new and valuable molecular markers for phylogenetic studies on planariids and geoplanids.

Introduction

Mitochondrial genetic sequences are generally considered appropriate markers to reconstruct phylogenetic relationships at low taxonomic levels because they usually have higher substitution rates than nuclear loci (Brown *et al*., 1979). Additionally, mitochondrial genes do not usually recombine, commonly exhibit a neutral evolution, and have small effective population size than nuclear counterpart which result in shorter coalescent times (Ballard and Whitlock, 2004; Barr *et al.*, 2005). These features make these sequences appropriate for either phylogeographical analyses of closely related species or even within-species population genetics studies (e.g. Stöck *et al*., 2006; Tryfonopoulos *et al*., 2010). Moreover, complete mitochondrial genomes offer the opportunity to study relevant biological aspects such as the effects of different life habits (e.g. Ballard and Melvin, 2010; Ballard and Pichaud, 2014).

Amongst the Platyhelminthes (Lophotrochozoa), the free-living triclads (Tricladida) have been recently included in biogeographical, phylogeographical and conservation studies (Lázaro *et al*., 2009; Álvarez-Presas *et al*., 2014; Solà *et al*., 2013). In particular the land planarians have become convenient models for understanding the origins and maintenance of biological diversity because of their low vagility and extreme dependence on the continuity and stability of their habitats. To date, all these studies have been based mainly on partial fragments of the mitochondrial gene cox1, due to limitations in amplifying other mitochondrial genes/regions.

Hitherto, currently there are scarce data of metagenome information from free-living flatworms, only one complete mitogenome (*Dugesia japonica*; ~18 kb), other almost complete (*Dugesia ryukyuensis*; ~17 kb) and a fragment of 6.8 kb (*Microstomum lineare*) (Ruiz-Trillo *et al*., 2004; Sakai and Sakaizumi, 2012). In contrast, there is considerably much more information for mitogenomes of parasitic platyhelminths (Neodermata); up-to 40 (Wey-Fabrizius *et al*., 2013). The neodermatan clade (includes Trematoda, Cestoda and Monogenea) is not far related to the Tricladida, which form part of the neodermatan sister group together with other flatworm lineages (Figure 1) (Riutort *et al*., 2012). However, the mitogenomes of neodermatans are highly divergent (e.g. see Wey-Fabrizius *et al.*, 2013), and found inappropriate for the design of specific primers to amplify triclad mitochondrial genes.

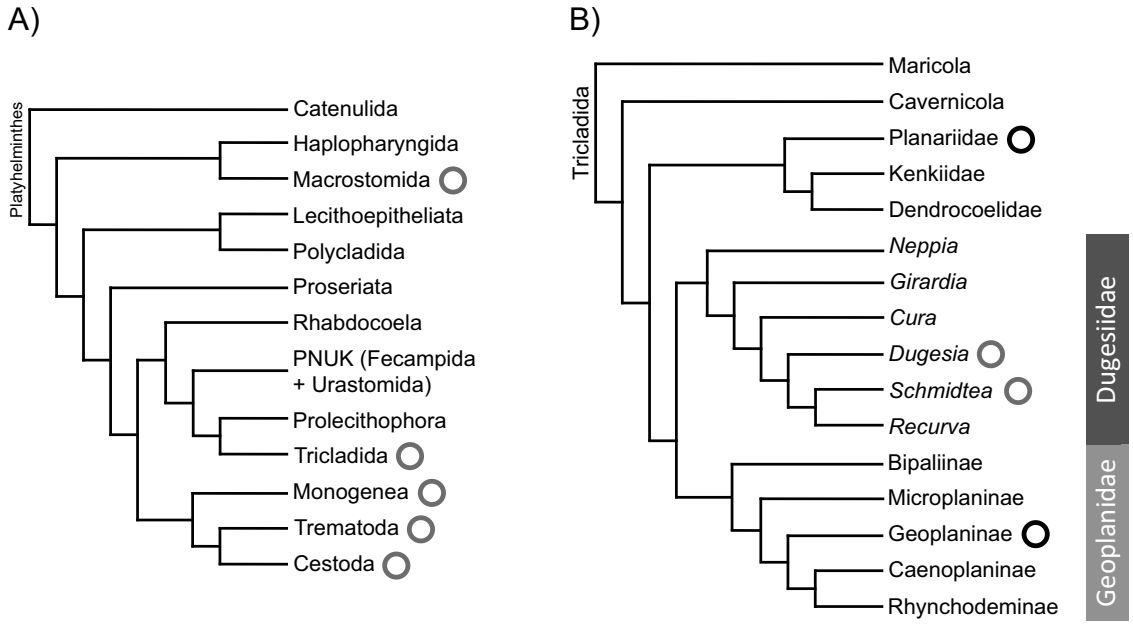


Figure 1. A) Phylogeny of the Platyhelminthes according to Riutort *et al.*, 2012 and B) phylogeny of the Tricladida according to Riutort *et al.*, 2012 and Sluys *et al.*, 2013. Monogenea, Trematoda and Cestoda constitute the Neodermata (parasitic flatworms) group. Grey circles indicate those groups for which mitogenomes are already available. Black circles indicate new obtained mitogenomes.

Through denser taxon sampling the development of universal and specific primers within this group should be achievable. Additionally, this will provide gene order, nucleotide and amino acid data for phylogenetic studies across the phylum, confirming for example the use of the rhabditophoran mitochondrial genetic code for the whole group (Telford *et al*., 2000), the identity of initiation and stop codons, and composition skews (Le *et al*. 2004). Finally, it will also allow the comparison of free-living to parasitic genomes to find out whether the different lifestyles result in differences on their genomes evolution.

Here we have determined the mitochondrial genomes of two Tricladida species belonging to two different superfamilies (*Crenobia alpina*, Planarioidea; *Obama* sp., Geoplanoidea) with two major aims, (i) to study the molecular evolution of mitochondrial molecules in the platyhelminths and (ii) to determine putative evolutionary selective differences between free-living and parasitic species according to their lifestyles. In order to achieve the first objective we have compared the sequence and gene annotations of the new mitogenomes together with those of available free-living species (*Dugesia*, Sakai and Sakaizumi, 2012; *Schmidtea mediterranea*. Ross *et al*., unpublished). To achieve the second objective we have used complete mitogenomic data to test whether, as previously proposed, parasitic species exhibit a relaxation of

natural selection, as compared with free-living organisms, caused by a putative reduction on their effective population sizes (Dowton and Austin, 1995; Castro *et al*., 2002; Bromham *et al*., 2013). For the study, we analyzed the impact of mutational and selective strengths on codon bias, coding regions, and on the whole mtDNA molecule. Besides, our new mitogenomic data will be useful to further conduct phylogenetic and phylogeographic-based analyses in triclads.

Material and methods

Samples

Five species of Tricladida from three different families (Dugesiidae, Geoplanidae, Planariidae) were selected to obtain the complete mitochondrial genome sequence (Table 1). Live specimens of *Crenobia alpina* (Dana, 1766), *Polycelis felina* (Dalyell, 1844), *Dugesia subtentaculata* (Draparnaud, 1801) and *Obama* sp. (*Obama* sp. 6: Carbayo *et al*. 2013) were collected in different localities from Catalonia, and *Microplana terrestris* (Müller, 1774) specimens were kindly provided by Mrs. Jill McDonald (UK). Information on the sampling localities is shown in Supplementary Table 1. The complete mitochondrial genomes of eight neodermatans were also retrieved from GenBank (Table 1) to carry out a preliminary gene checking of the mitogenomes obtained in the present work by 454 (Roche) pyrosequencing and to perform analytical comparisons between triclads and parasitic flatworms.

Mitochondrial DNA extraction

We isolated mitochondrial DNA from about 100 animals for each species based on a modification of the protocol described in Bessho and collaborators (1997). We first removed the mucus from the planarians with a diluted Cysteine-Chloride solution (pH 7.0) obtained from effervescent tablets (CINFA) and then dipped the animals in buffer 1 (0.1 M Sucrose, 10 mMTrisHCl, pH 7.4) overnight at −80°C. Animals were next homogenized, transferred to two PPCO tubes and centrifuged at 600 g (Beckman JA-20 rotor) at 2°C during 10 minutes in order to remove nuclei. The supernatant was centrifuged in FEP tubes at 15,000 g at 2°C for 10 minutes in a Sorvall^{TM} centrifuge (SS-34 rotor). The pellet was dissolved in 40 mL (20 mL in each tube) of 0.1 M Sucrose

solution containing 50 mM $MgCl_2$ (buffer 2). To remove any contamination of nuclear DNA from mitochondrial membranes, the solution was treated with 10 µl of 70 units/mL DNase. After inactivating the DNase (80ºC for 10 minutes), 200 mL (100 mL per tube) of 0.6% SDS, 10 mM EDTA, 10 mM Tris-HCl (pH 8.0) (buffer 3) were added and incubated at 60°C for 10 minutes to disrupt mitochondrial membranes. Finally, an ordinary phenol chloroform extraction was applied to isolate mitochondrial DNA (Chomczynski *et al*., 1987).

Table 1. List of all Platyhelminthes species included in the present work. Acronyms indicating the different analyses: **CG**, Comparative genomics; **PGS**, Preliminary gene screening; **SQ**, Sequencing. *, Species attempted to be sequenced but failed.

Life cycle	Species	Classification	Acc. Number	Analysis			References
				CG	PGS	SQ	
Free-living	*Crenobia alpina*	Tricladida, Planariidae	Pending	X		X	This work
	Dugesia japonica	Tricladida, Dugesiidae	AB618487.1	X			Sakai & Sakaizumi, 2012
	Dugesia subtentaculata	Tricladida, Dugesiidae	None			X*	This work
	Microplana terrestris	Tricladida, Geoplanidae	None			X*	This work
	Obama sp.	Tricladida, Geoplanidae	Pending	X		X	This work
	Polycelis felina	Tricladida, Planariidae	None			X*	This work
	Schmidtea mediterranea	Tricladida, Dugesiidae	NC_022448.1	X			Not published
Parasitic	*Benedenia hoshinai*	Monogenea, Capsalidae	NC_014591.1	X			Kang *et al*., 2012
	Diplogonoporus balaenopterae	Cestoda, Diphyllobothriidae	NC_017613.1	X			Yamasaki *et al*., 2012
	Fasciola hepatica	Trematoda, Fasciolidae	NC_002546.1	X	X		Le *et al*., 2000
	Schistosoma japonicum	Trematoda, Schistosomatidae	NC_002544.1	X			Le *et al*., 2000
	Taenia saginata	Cestoda, Taeniidae	NC_009938.1	X			Jeon *et al*., 2007
	Taenia solium	Cestoda, Taeniidae	AB086256.1		X		Nakao *et al*., 2003
	Tetrancistrum sigani	Monogenea, Ancyrocephalidae s.l.	NC_018031.1	X			Zhang *et al*., 2014
	Gyrodactylus derjavinoides	Monogenea, Gyrodactylidae	NC_010976.1		X		Huyse *et al*., 2008

Mitochondrial DNA quantification and 454 sequencing

We quantified the DNA amount by a Qubit 2.0 fluorometer (Invitrogen) following manufacturer's instructions, obtaining very dissimilar amounts across the 5 species, from around 2 µg in *Polycelis felina* to 500 ng from *Dugesia subtentaculata*, and a nearly undetectable amount for *Microplana terrestris*. After precipitating the DNAs it was resuspended in TE to a final concentration of 20 ng/µL. The five DNA samples were multiplexed identifier (MID) tagged, and the 454 libraries prepared at the Centres Científics i Tecnològics de la Universitat de Barcelona (CCiTUB). The samples were run into a ¼ 454 plate of the GS FLX titanium platform.

Sequencing reads processing

DNA sequences (reads) and quality information were extracted independently of each MID's in fasta format from the Standard Flowgram Format file (SFF) using the sffinfo script from Roche's Newbler package (454 SFF Tools). We removed adapters, putative contaminant sequences (upon the UniVecdatabase and the *E. coli* genome sequence) and reads shorter than 50 bp were removed using the SeqClean (http://compbio.dfci.harvard.edu/tgi/software/) script. All reads with a mean quality score below 20, and trimmed the low-quality bases at the ends of the reads were also removed using PRINSEQ (Schmieder and Edwards, 2011).

Sequencing reads post-processing

We determined whether the mitochondrial genes were present in sequencing reads by a BLAST analysis (v. 2.2.24) using available mitochondrial genome data (downloaded from NCBI) of parasitic flatworms (Table 1) as query. In particular we used the protein information of *Taenia solium* (Nakao *et al*., 2003), *Gyrodactylus derjavinoides* (Huyse *et al*., 2008) and *Fasciola hepatica* (Le *et al*., 2000) (Supplementary Table 2). For the analyses we applied the tBLASTn algorithm (e-value cut-off: 10^{-3}), using translation table 9 (echinoderm and flatworm mitochondrial code) to translate DNA information of the 454 reads in all six reading frames.

Mitochondrial genomes assembling, annotation, PCR amplification and re-sequencing

We first tried to assemble the DNA genome sequence using Newbler 2.6 (454 life Sciences, with settings: -urt -ml 40 -mi 85 -minlen 50), but with little success. Actually, we only obtained several short contigs, with a N50 length of about 400 nucleotides. However, we got a large, nearly complete mtDNA sequences including all filtered 454 reads using the SeqMan software (DNASTAR, http://www.DNASTAR.com). The assembled mitogenomes were annotated with Geneious Pro 6.1.7 (Biomatters, 2014). Later, we validated the genome assemblies by further Sanger DNA sequencing; this experimental approach allowed us to determine the existence of, and thereby correct, some 454-induced sequence errors (e.g. frameshifts; Huse *et al*., 2007), to complete the molecules, and to confirm the gene order resulting from the assembled genomes. For such analysis, we designed 34 primers for PCR amplification in *C. alpina* and 20 primers for *Obama* sp. (Supplementary Tables 3 and 4) covering the whole length of the

genomes. PCR reactions initially included: 1 µl of DNA, 5 µl of Promega 5X Buffer, 1 µl of dNTPs (10 mM), 0.5 µl of each primer (25 µM), 2 µl of $MgCl_2$ (25 mM), 0.15 µl of *Taq* polymerase (GoTaq® Flexi DNA Polymerase, Promega). Double-distilled and autoclaved water was added to obtain a final 25 µl PCR volume for all molecules. In many cases it was necessary to vary the annealing temperatures or the amount of $MgCl_2$ or DNA to obtain amplification products. PCR products that yielded direct sequences of not enough quality were cloned. Cloning was carried out with the TOPO TA Cloning® Kit of Invitrogen™ following the manufacturers' instructions. For every PCR product cloned, five bacterial colonies in average were picked and sequenced in order to obtain representation of the different haplotypes. The cloned fragment was amplified using universal vector primers T3 and T7. All PCR products were purified before sequencing using the purification kit illustra™ (GFX™ PCR DNA and Gel Band of GE Healthcare) or by using a vacuum system (MultiScreen™$_{HTS}$ Vacuum Manifold, Millipore). Sequencing reactions were performed by using Big-Dye (3.1, Applied Biosystems) with the same primers used to amplify the fragment. Reactions were run on an automated sequencer ABI Prism 3730 (Unitat de Genòmica of Centres Científics i Tecnològics de la Universitat de Barcelona − CCiTUB) or at Macrogen Corporation (Amsterdam, the Netherlands). The chromatograms were visually checked. These additional DNA sequences were aligned and compared with the 454-based assemblies using the software Geneious 6.1.7, which was also used to obtain the final assemblies.

Prediction of protein-coding genes

We determined the location of the protein-coding genes by using a combination of BLAST searches, ORF finder and Glimmer plug-in in Geneious 6.1.7, MITOS online software (Bernt *et al*., 2013), and using information from published Platyhelminthes sequences.

Prediction of tRNAs and genes for rrnL and rrnS

Putative tRNA genes were identified using a combination of the following software: ARWEN (http://130.235.46.10/ARWEN) (Laslett and Canbäck, 2008), tRNAscan-SE 1.21 (Schattner *et al*., 2005), MITOS (Bernt *et al.*, 2013) and DOGMA (Wyman *et al*., 2004). The tRNAs not found with these programs were annotated by eye with reference to known platyhelminth sequences. Repetitive regions were searched with the online software Tandem Repeats Finder (Benson, 1999). In addition to our mtDNA molecules,

we included the already published *D. japonica* mitochondrial genome (Sakai and Sakaizumi, 2012) to double-check the annotation of the molecule.

Nucleotide composition analyses

In addition to the A+T (or G+C) content, we also estimated the putative nucleotide frequencies bias (NB statistic) at a given strand. Similarly to Shields and collaborators (1988), we defined the NB statistic as:

$$NB = \left[\sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i} \right] / n$$

Where O_i and E_i are the observed and the expected (under equifrequency) numbers of nucleotide variant i (i = 1, 2, 3, and 4 correspond to A, C, G, and T), and n is the total number of positions analyzed. We applied the NB statistic in different portions of the mitochondrial molecule: NBp, NB at the protein coding regions; NB2, NB at the second position of codons; NB3, NB at the third of position of four-fold degenerate codons; NBr and NBt, NB at the ribosomal and tRNA genes, respectively.

We also estimated the particular AT and GC strand skews, using the Perna and Kocher (1995) indices, where the AT skew (sAT) is computed as (A−T)/(A+T) and the GC skew (sGC) = (G−C)/(G+C); in both cases the nucleotide frequencies are those of the focal strand (in all cases the coding strand). These values range from −1 to +1, where a value of zero indicates that the frequency of A is equal to T (AT skew), or G equal to C (GC skew). We calculated these indices for each gene and for the whole mitochondrial genome of *C. alpina* and *Obama* sp., but also for other free-living flatworms with available mitochondrial genome sequence data, and for six selected parasitic species (Table 1). We also computed the sAT (and sGC index) in different functional regions of the mitochondrial molecule, being sATp, the sAT at the protein coding regions; sAT2, sAT at the second position of codons; sAT3, the sAT at the third of position of four-fold degenerate codons; sATr and sATt, sAT at the ribosomal and tRNA genes, respectively.

Codon Bias analyses

We estimated the codon usage bias applying the scaled chi squared (SC) (Shields *et al*. 1988), which is a measure based on the chi square statistic normalized by the number of

codons, and Effective Number of Codons statistics (ENC) (Wright, 1990). For the SC calculation we conducted two types of analyses: for one we used as the expected values those values assuming codon equifrequency (the standard way to compute SC), for the other, we used the observed nucleotide frequencies to determine the expected codon frequency values. For the latter we conducted the analysis separately for each species, and using 4 different types of observed nucleotide frequencies: the SC statistic computed (SCp) using as the expected number of codons (at each codon class) those values based on the observed nucleotide frequency at the protein coding region (the average for all genes within a species); SC2, the SC using information of the observed nucleotide frequencies at the second position of codons; SC3, SC using information at the third position of four-fold degenerate codons; and SCr and SCt, those SC values using the observed nucleotide frequencies at the ribosomal and tRNA genes, respectively.

Results

454 raw data processing, assembling and gene annotation

The statistics for the 454 sequencing are shown in the Supplementary Table 5. After quality pre-processing, *Microplana terrestris* was excluded since there were practically no reads. Surprisingly, for the rest of samples the better results obtained in the 454 sequences did not coincide with the species that showed a higher amount in the fluorimetric measures for the DNA quantification. The length of the reads of *P. felina* and *D. subtentaculata* were short (N50 of 246 and 146, respectively), and the prospective tBLASTn analyses showed that only a few reads included protein coding gene (PCG) information (of only 4 or 6 protein coding genes, respectively) (Supplementary Table 6). Although it was not possible to assemble the whole mitogenome for *D. subtentaculata* we have been able to map the reads obtained on the mitogenome of *D. japonica*, and used that information to develop some specific primers for future studies (data not shown).

The 454 reads of *C. alpina* and *Obama* sp. provided sufficient information to assemble the mitogenomes successfully (Figure 2). The SeqMan assembly of *C. alpina* generated a single contig of 17,079 bp, including ambiguous positions that were automatically excluded when the contig was saved. The average coverage of the

assembly was of 29.07. *Obama* sp. 454 output generated a 14,893 bp contig with an average coverage of 24.28. The quality of the sequence for this assembly was poorer than that obtained for *C. alpina*. This is probably due to an increased 454 error rate in *Obama* sp. as a consequence of its higher frequency of homo-polymer sequences. The final assemblies for both species contained all mitochondrial genes but lacked a large portion of the main non-coding region.

These preliminary assemblies were improved and completed by further Sanger DNA sequencing. We carried out additional PCR partial amplifications on the basis of the first assembly, and identified missing and/or extra bases. For instance, in the putative sequence of *C. alpina* for nad4 and nad5 a nucleotide was lacking from the 454 reads leading to a frameshift, making the recovery of an appropriate ORF impossible without adding a nucleotide. This situation was the same in several genes in *Obama* sp. assembly. Comparison between the two sets of sequences allowed complete annotation of the genes for both species. Sequences obtained from the cloned PCR products showed the presence of 3 polymorphic sites (in one case including a nucleotide indel) in intergenic regions in *C. alpina*.

It was not possible to re-sequence the complete mitogenome of *C. alpina*. Designed primers failed to PCR amplify a fragment of the genome containing the putative repetitive region (Figure 2A). The 454 assembly of this region by SeqMan software recovered only two copies of the 186 bp repetitive sequence (consensus size) due to the limitation in 454 read lengths. However, when the 454 reads were aligned with the whole mitochondrial molecule, this repetitive region showed a much higher read coverage than the rest of the molecule. A comparison with the general coverage suggests that the repeat unit must be repeated around 4 times.

Hence we do not know the exact length of repeat region, and thus the full mitogenome. For *Obama* sp. we PCR amplified a band of around 2,000 bp from the 3' end of *rrnL* to the 5' end of *cob* gene. However, it was not possible to obtain clean sequences, probably because the presence of a repetitive region within this fragment (Figure 2B), hence the complete mitogenome length is also unknown for this species.

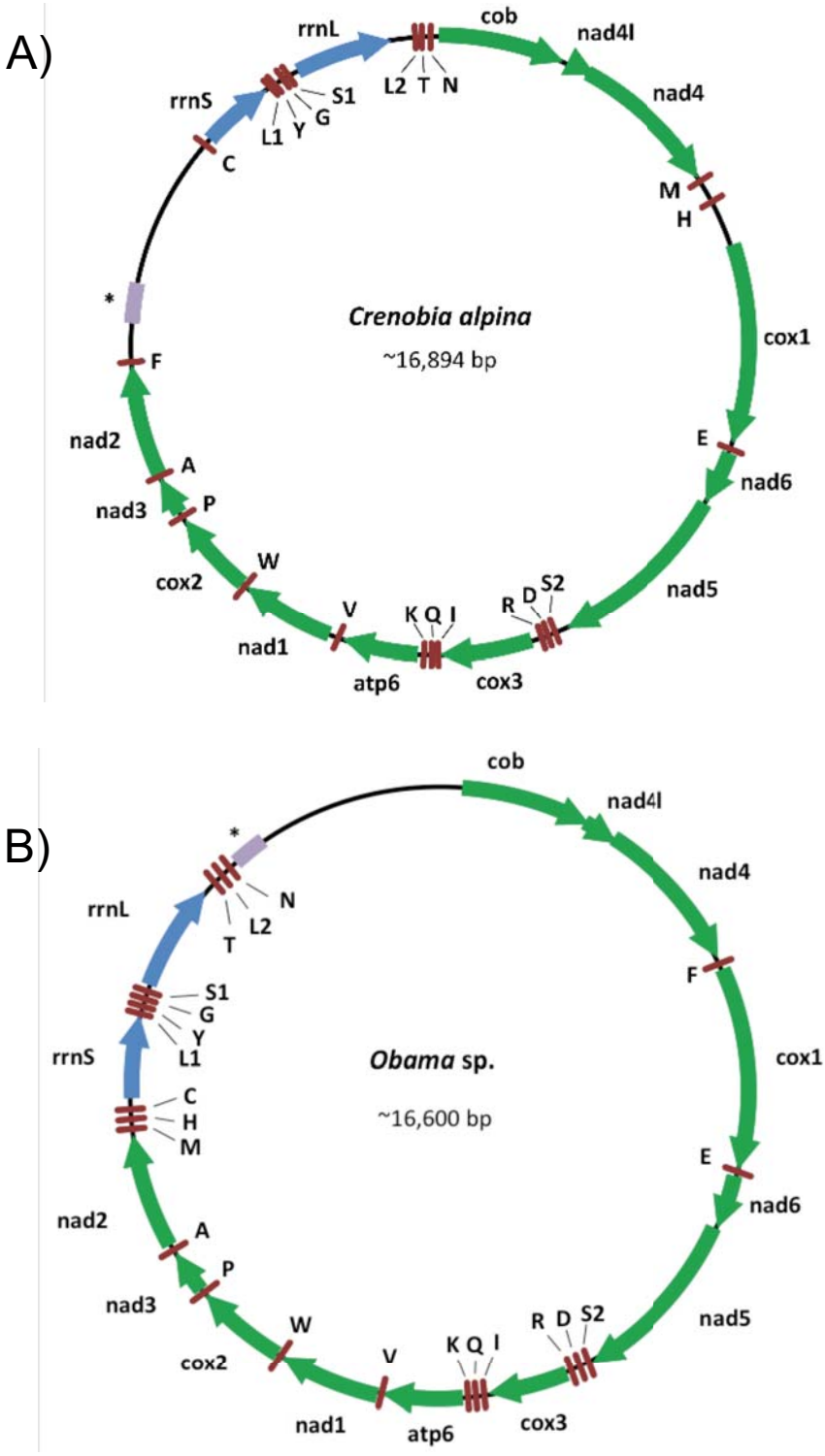


Figure 2. Arrangement of the mitogenomes of *Crenobia alpina* (A) and *Obama* sp. (B). Green arrows correspond to the protein coding genes; blue arrows ribosomal genes; brown rods tRNAs; Purple bar indicates the putative repetitive region.

The mitochondrial genome of *C. alpina* (estimated size >16,894 bp; GenBank ID: pending submission) and *Obama* sp. (estimated size ~16,600 bp; GenBank ID: pending submission) encode 12 protein-coding genes (lacking *atp8*, absent in all the characterized platyhelminth mitochondrial genomes; Wey-Fabrizius *et al*., 2013), 22 tRNA genes and 2 ribosomal genes (Figure 2 and Supplementary Tables 7 and 8). Consistent with other platyhelminth mitogenomes all the genes are transcribed from the same strand. *Nad4l* gene was the single case of a gene overlapping with other genes; while in *Obama* sp. it overlaps at both ends (17 bp with *cob* and 32 bp with *nad4*), in *C. alpina* it only overlaps with *nad4* (32 bp).

Gene order

The PCG order is conserved across Tricladida, but it is radically different from that found in *Microstomum* (the only but partial available genome from a non-triclad free-living platyhelminth), and all the parasitic species (Figure 3). Only three blocks of genes are conserved between parasites and triclads (Supplementary Figure 1). Our re-annotation of the *D. japonica* mitogenome implied the change of three tRNAs to positions more similar, or identical, to those found in the other triclads: *trnC* is on the same strand as the rest of genes and *trnA* and *trnL1* are in the same relative position than in the other triclads (Supplementary Figure 2). In spite of these changes all four triclad species (*C. alpina*, *Obama* sp., *S. mediterranea* and *D. japonica*) differ in the location of some tRNAs (Supplementary Figure 3).

The ribosomal genes are situated close to the long non-coding region in the four Tricladida species, although in a different position. For *C. alpina* and *S. mediterranea* the long non-coding region is situated 5' upstream of the ribosomal genes while for *Obama* sp., and *D. japonica* it is situated at its 3' end. Moreover, at difference to the rest of platyhelminths, for triclads *rrnS* is situated upstream of *rrnL* (Figure 3).

Start and terminal codons

We infer that four start codons are used in the two species analyzed. TTG and ATG are used at equivalent frequencies in *Obama* sp. while ATG is more frequent than TTG in *C. alpina*, TTA is also used in both species and GTG only in *Obama* sp. (Supplementary Tables 7 and 8). Stop codons are TAG and TAA. In *C. alpina*, *cox2* gene has a TAR stop codon, showing the presence of the two possible stop codons

within the population (Heterozygosity). Alternatively this could be a case of a truncated TA stop codon.

The length of the genes is very similar between the two species. However, in general the predictions for *Obama* sp. are slightly longer resulting in a more compact genome (shorter intergenic regions).

Transfer RNAs and ribosomal genes

Both *Crenobia alpina* and *Obama* sp. present 22 tRNA genes (Supplementary Figures 4 and 5). The tRNAs *trnS2* and *trnT* lack the DHU arm in both species, while in *C. alpina* the *trnQ* could have two alternative structures: either lacking the TΨC arm or the DHU arm.

In *C. alpina*, four tRNAs overlap (*trnI*, *trnW*, *trnA*, *trnF*) with the last two bases of four genes (*cox3*, *nad1*, *nad3*, nad2 respectively). Moreover, *trnL1* overlaps with *trnaY*. In *Obama* sp., *trnF* and *trnV* overlap 1 nucleotide with genes *nad4* and *atp6* respectively. On the other hand, there are 3 cases of overlap between tRNAs (*trnD* and *trnR*, 5 bp; *trnQ* and *trnK*, 8 bp; *trnY* and *trnG*, 4 bp).

The new annotation of *D. japonica* mitogenome implies the relocation of three tRNAs. Considering their secondary structure, the *trnA* and *trnL1* preserve the four-arms while *trnC* lacks TΨC arm (Supplementary Figure 6).

Non-coding regions

For both species the initial assembly by SeqMan generated a linear contig containing all mitochondrial genes flanked by non-coding regions. For *C. alpina*, we obtained the final assembly (a circular genome) after some additional resequencing work based on the design of PCR primers at the two ends of the initial assembly. In *Obama* sp., although we have been able to close the circle by PCR amplification (as explained above) we have not been successful in the sequencing of the amplified fragment, thus we are not confident about the real length of the molecule or the number of repeat elements within the long non-coding region.

Figure 3. Linearized schemes of gene orders in Platyhelminthes, based on Wey-Fabrizius *et al.*, 2013. Those genes that are variable within each of the three parasitic groups (Cestoda, Monogenea and Trematoda) are in bold. Multiple genes in the same box indicate variable gene orders within the specific group. Gene identifier: *rrnS*/*rrnL* = *small and large subunit rRNA*, *nad1−6,4L* = *NADH dehydrogenase subunits 1−6 + 4L*, *cox1−3* = *cytochrome c oxidase subunits* 1−3, *atp6* = *ATP synthase subunit 6*, *cob* = *cytochrome b*. The tRNAs are shown according to the amino acid code letter.

The comparison of the final assembly of *C. alpina* with the 454 reads showed that the long non-coding region probably contains four repeats or more of 186 bp (consensus size), preceded by a non-repetitive region of 309−311 bp and followed by another non-repetitive region of 1,363 bp. The total length of this large non-coding region is, at least, 2,028 bp. In the case of *Obama* sp. we only have the information of the length of the amplified fragment, around 2,000 bp, resulting in a full approximate length of 16,600 bp.

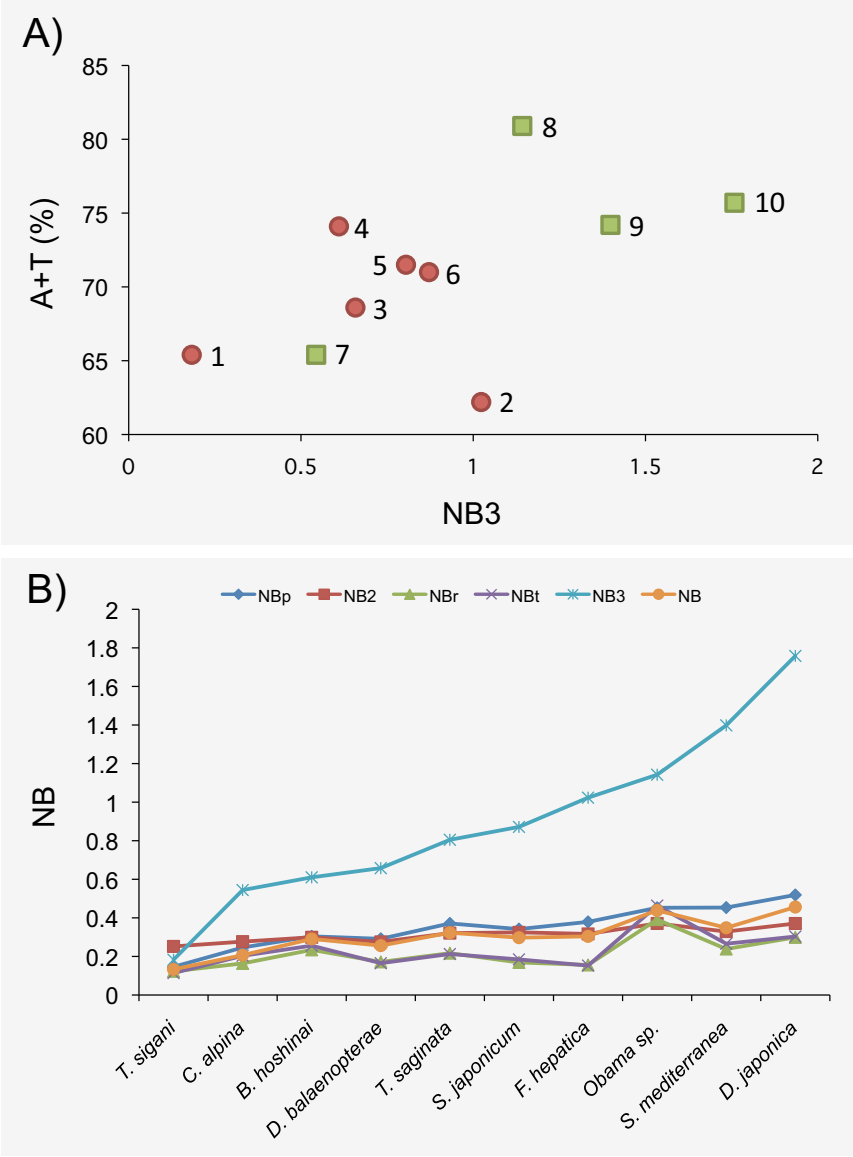


Figure 4. A) Relationship between A+T content and NB3 (NB at the third position of four-fold degenerate codons) values. Green squares and red circles indicate free-living and parasitic platyhelminths, respectively. The surveyed species are shown in numbers: 1, *T. sigani*; 2, *F. hepatica; C. alpina*; 3, *D. balaenopterae*; 4, *B. hoshinai*; 5, *T. saginata*; 6, *S. japonicum*; 7, *C. alpina*; 8, *Obama* sp.; 9, *S. mediterranea*; 10, *D. japonica*. B) Values of the different NB-based statistic across species.

Nucleotide composition and strand skew bias

We found that triclad mitogenomes present high A+T content values (over 60%), being *Obama* sp. the taxon with a higher bias (mean: 81.2%) (Figure 4A). As might be expected from this result, there is also substantial per strand nucleotide frequency bias, both in free-living and parasitic species (Figure 4B; Supplementary Figure 7). We found such bias both at the whole molecule (NB statistic) and in different portions of the same (NBp, NB2, NB3, NBr and NBt). Interestingly, the highest values correspond to the NB3 statistic (Figure 4B), and clearly overlap with species exhibiting the higher A+T content values (Figure 4A). This result points to mutation, and not natural selection, as the major evolutionary force responsible for the bias in the nucleotide frequencies. Remarkably, the free-living and parasitic species differ considerably in their nucleotide frequency bias, with free-living species having higher values (with the exception of *C. alpina*). Moreover, this pattern is consistent across the different NB measures (Supplementary Figure 7).

In contrast to the A+T and NB values, free-living and parasitic species do not form separate clusters with respect to sAT or sGT values, neither for the total data nor for the values estimated at positions with different functional behavior (Figure 5; Supplementary Figure 8). All sAT values are negative (in all genes and in all species), with the exception of the *rrnS* gene of *Obama* and *T. sigani* that are slightly positive (Figure 6A and B). Thus, there is a clear prevalence of T over A in the coding strand. Moreover, the general sAT skew varies considerably among species (−0.187 to −0.4 Tricladida; −0.168 to −0.483 Neodermata), but it is consistent across genes; for instance *F. hepatica* has the highest overall sAT values, a feature exhibited in all of its genes (Figure 6B). The sAT and A+T content, however, are uncoupled; for instance, in *Obama* sp., the species with highest A+T content, exhibits nearly the lower sAT values. The general sGC estimates also show important strand skews, ranging from 0.246 to 0.283 in triclads and 0.148 to 0.475 in parasites, which indicate a higher frequency of G than C. Although the sGC values also show some species-specific pattern is much less consistent across genes. Overall, the analyses uncover a species-specific pattern that (i) is not correlated with the actual A+T content (Figure 5), (ii) differs between sGC and sAT estimates, and (iii) does not cluster separately free-living or parasitic species separately.

To gain insights into the relevance of the variation of skew levels along the mitochondrial sequence, we analyzed the sAT and sGC levels in the different genes as a function of their relative physical order (Figure 6). We found no clear polarity, either in sGC or in sAT levels. The analyses conducted separately in different functional positions (such as sGC2 or sAT2) do also not show any polarity (Data not shown).

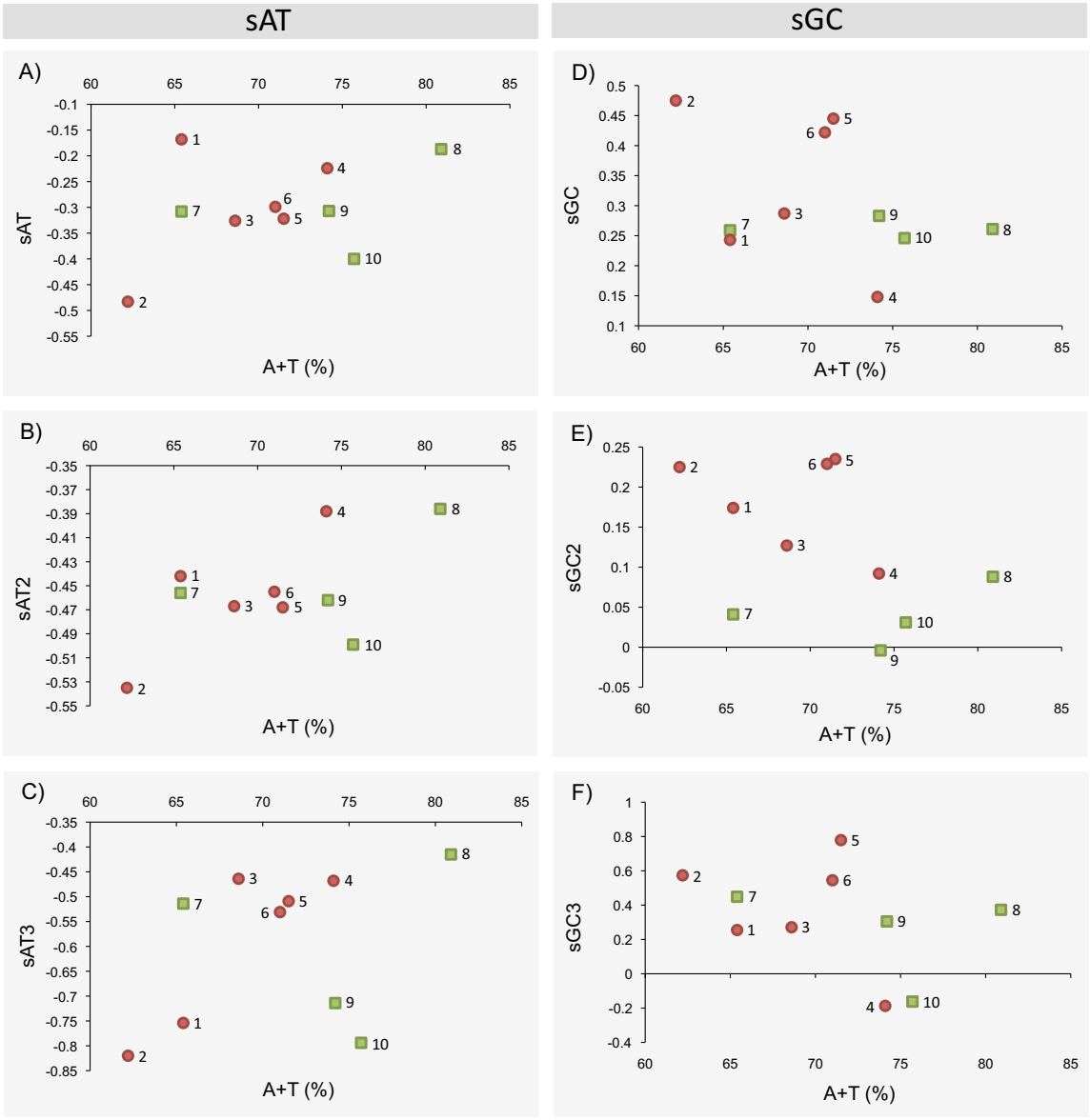


Figure 5. Relationship between sAT and sGC values and A+T content. sAT general skew; sAT2, sAT skew at the second positions; sAT3, sAT at the third positions. sGC, general skew; sGC2, sGC skew at the second positions; sGC3, sGC at the third positions. Green squares and red circles indicate free-living and parasitic platyhelminths, respectively. The surveyed species are shown in numbers: 1, *T. sigani*; 2, *F. hepatica*; 3, *D. balaenopterae*; 4, *B. hoshinai*; 5, *T. saginata*; 6, *S. japonicum*; 7, *C. alpina*; 8, *Obama* sp.; 9, *S. mediterranea*; 10, *D. japonica*.

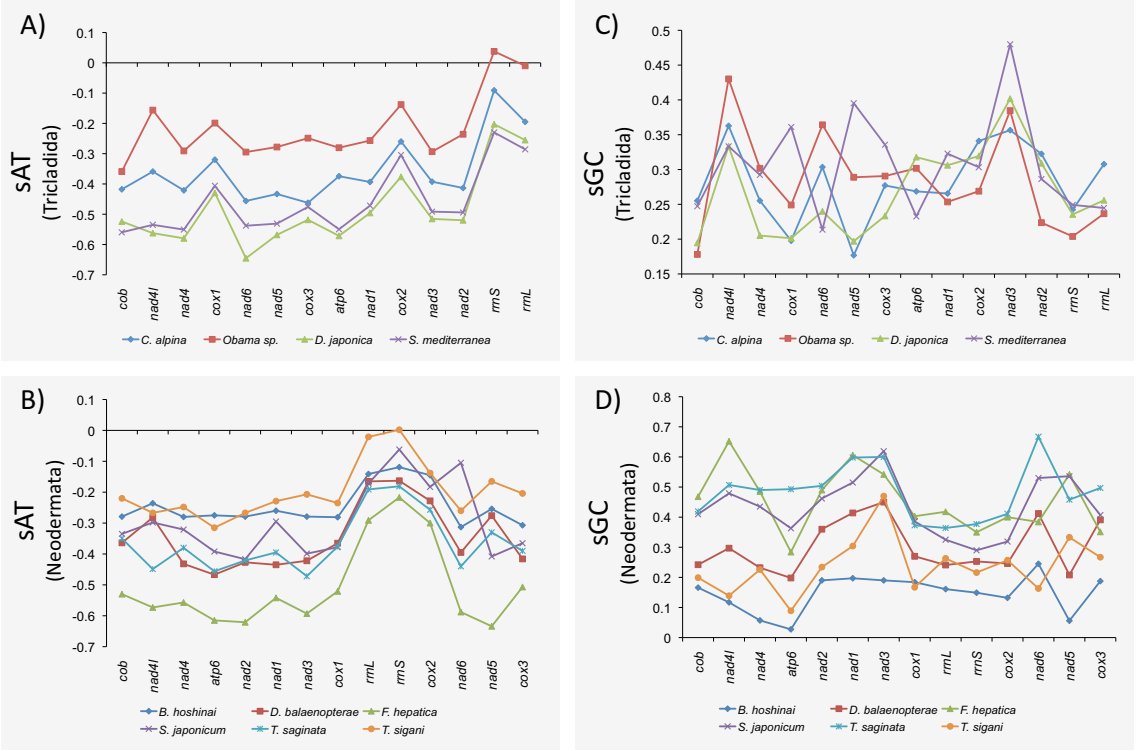


Figure 6. sAT and sGC values of the protein coding genes (PCG) along the mtDNA molecule. A) sAT of Tricladida; B) sAT of Neodermata; C) sGC of Tricladida; D) sGC of Neodermata.

Codon composition bias

The results of the codon usage also show high levels of bias across those species surveyed (Figure 7), both using the SC or ENC estimators. Interestingly, and in agreement with the nucleotide bias analyses, the free-living species again show high levels of codon bias (excepting *C. alpina*). The codon bias might be a by-product of the mutational input or might result from the action of natural selection for increased translational efficiency or accuracy (Bernardi and Bernardi 1989; Poh *et al*., 2012; Lawrie *et al*., 2013). To disentangle both effects we studied the level of codon bias adjusting for the observed mutation bias (Figure 7C; Supplementary Figure 9). As expected, the SC values drop dramatically, and especially for SC3 values. However, we do not observe any clear pattern that differentiates free-living and parasitic species. Moreover, using different SC-mutational adjusting estimators yields different species-rank orders, meaning that the codon bias feature disappears.

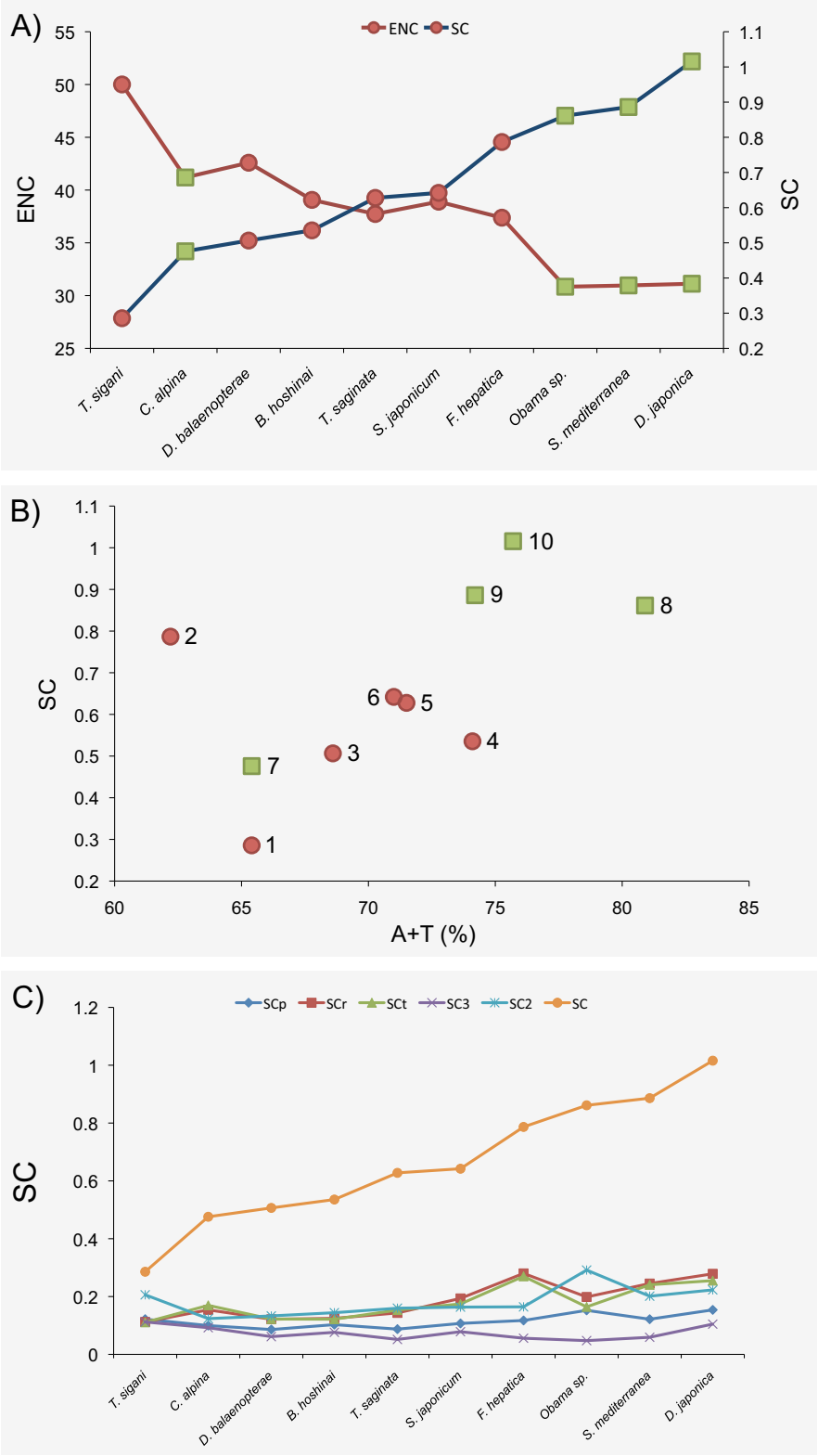


Figure 7. Relationship between different codon bias measures. A) Relationship between ENC and SC values. B) Relationship between SC and A+T% values. C) SC values across species. Green squares and red circles indicate free-living and parasitic platyhelminthes, respectively. The surveyed species are shown in numbers: 1, *T. sigani*; 2, *F. hepatica*; 3, *D. balaenopterae*; 4, *B. hoshinai*; 5, *T. saginata*; 6, *S. japonicum*; 7, *C. alpina*; 8, *Obama* sp.; 9, *S. mediterranea*; 10, *D. japonica*.

Discussion

Mitochondrial genomes of tricladida: general features

The mitochondrial genomes of the new triclad species characterized in the present work share the same gene composition with the majority of the Platyhelminthes sequenced so far. *C. alpina* and *Obama* sp. contain 12 PCG, lacking atp8, a gene absent in the mitochondrial genomes of Chaetognatha, Rotifera and Bivalvia among lophotrochozoans as well as in most Nematoda (Wey-Fabrizius *et al*., 2013, Gissi *et al*., 2008). The tRNAs number is 22, as found in almost all other platyhelminth genomes, except two species of the digenean genus *Schistosoma*, having 23 in *S. japonicum* and *S. mansoni* due to a duplication of trnC (Le *et al*., 2000; Zhao *et al*., 2012), and within the cestode genus *Echinococcus* (Le *et al*., 2002; Thompson *et al*., 2006). All genes are transcribed from the same strand, a situation also found in Cnidaria, Porifera, Tunicata and many other lophotrochozoan phyla (Gissi *et al*., 2008; Wey-Fabrizius *et al*., 2013). Although in the published *D. japonica* mitogenome the *trnC* has a reverse orientation, our re-annotation shows that in fact all genes are transcribed from the same strand in this taxon also.

The genetic code used by all triclad species is consistent with that used for the majority of Platyhelminthes, EMBL-NCBI genetic code 9: Echinoderm and Flatworm. We have found no evidence of TAA coding for Tyr (as proposed by Bessho *et al*. 1992a,b); on the contrary TAA is shown to be the stop codon for most of our predicted genes (a situation also found for some genes in *D. japonica*, Sakai and Sakaizumi, 2012). Hence the proposed alternative code for Platyhelminthes, code 14 from EMBL-NCBI, is most likely a feature exclusive to nematodes (Jacob *et al*., 2009).

Gene order

The PCG order is identical in *C. alpina* and *Obama* sp. (Figure 2 and 3), and also with the mitochondrial genomes of *D. japonica*, *D. ryukyuensis* and *S. mediterranea*. The only differences include the identity and arrangement of the tRNAs and the relative position of the long non-coding regions. In *C. alpina* and *S. mediterranea* a large non-coding region and the repetitive region (RR) is situated between *nad2* and the ribosomal genes, while in *Obama* sp. the long non-coding region including a RR is situated just after the ribosomal genes before *cob*, as in *D. japonica*

and *D. ryukyuensis*. This is surprising considering the closer phylogenetic relationships between *S. mediterranea* and *Dugesia* and *Obama*, all belonging to the superfamily Geoplanoidea, sister to the Planarioidea to which *Crenobia* belongs (Figure 1). However, the length of the main non-coding region in *S. mediterranea* is extremely long (nearly as long as the whole coding region), which invites to be cautious on its validity. On the other hand, the small number of changes in tRNAs order (Supplementary Figure 3) among all Tricladida is a notable feature given the very likely antiquity of the lineage (Solà *et al.*, in prep).

The gene order among Tricladida is considerably different from that found in the parasitic platyhelminths and in *Microstomum*. One special feature for Tricladida is the relative order of the two ribosomal genes; *rrnS* is located at 5' from *rrnL*, being the other way around in all the other Platyhelminthes sequenced until now. Futhermore, in neodermatans *rrnL* and *rrnS* are flanked by *cox1* and *cox2*, whereas in triclads *rrnS* and *rrnL* are flanked by *nad2* and *cob*.

Start and terminal codon usage

While parasitic flatworms use only ATG and GTG as start codons, with the exception of a GTT used in *Hymenolepis diminuta* (Le *et al*. 2002; Wey-Fabrizius *et al*., 2013), there seems to be a much higher versatility in Tricladida (Supplementary Tables 7, 8 and 9; Sakai and Sakamuzi, 2012). In this group apart from ATG and GTG, start codon TTG seems to be commonly used. Additionally, TTA and TAT putative start codons have also been found. There is no conservation on the start codon used for each gene through the Tricladida; in fact, only the start codon of *atp6* (TTG), is shared between all triclads. The diversity should therefore most probably have arisen independently in the different species. Although abbreviated stop codons (TA or T) are common in animal mitogenomes (Boore and Brown, 1995 and references therein), we found that triclads have standard trinucleotide stop codons. In *Obama* sp., 10 out of the 12 PCG terminate in TAA, while *D. japonica* has the reverse situation 10 out of 12 PCG have TAG as stop codon. In *C. alpina* and *S. mediterranea* the usage of both stop codons is almost the same. The preference of the TAA stop codons in *Obama* sp. could be explained by the high frequency of A over G along its genome. The situation in the other three species with a similar proportion of A and G can explain the proportions of stop codons found in *S. mediterranea* and *C. alpina*, but not in *D. japonica*.

Although we used different methods to infer the start and stop codons, it must be taken into account that the genes could be not perfectly delimited because of the lack of transcriptional information. Future studies involving transcriptomic analyses will help for a more accurate annotation of these species' genes.

A+T content and asymmetric strand bias

We have found that triclads have high A+T content values, feature already detected in parasitic flatworms. Nevertheless, while some parasitic species have A+T content values around 70%, *Obama* sp. exhibit a much more extreme bias (over 80%), close to the highest described cases (Hymenoptera; Wei *et al*., 2009).

The surveyed triclad species exhibit negative sAT and positive sGC skew values in the coding strand, a typical feature also reported in Platyhelminthes (Castellana *et al*., 2011; Weber *et al*., 2013; Wey-Fabrizius *et al*., 2013). It has been proposed that this feature would be linked to the replication process (Tillier and Collins, 2000; Necsulea and Lobry, 2007; Marin and Xia, 2008). That is, the longer strands are kept single during replication, the higher the likelihood of depurination mutations resulting in substitutions from A to G and from C to T (100 times more frequent). However, our results do not show any polarity in the skew values across genes along the mtDNA molecule as would be expected; but there is a clear species-specific pattern with contrasting values across species (Figure 6). The fact that the A+T content (or the NB3 value) and skew values do not correlate across species (Figure 5 and Supplementary Figure 8) does not support the mutational input as a major source for the skew. The situation is the same when we consider the skews for only second or third sites within the coding regions (Figure 5B and C and Supplementary Figure 8B and C). In contrast, species exhibiting high AT levels (such as *Obama* sp.) have indeed the lowest sAT values. These results suggest that the asymmetric nucleotide composition strand bias has some significance. This could be related to the fact that all genes are situated on the same strand. For example, in bacteria it has been proposed that as a consequence of the excess of genes situated on the same strand, biases in transcription-coupled repair could lead to a skew between the strands in nucleotide composition (Francino *et al.*, 1996).

Effect of natural selection on free-living and parasitic species

It has been proposed that parasitic species might exhibit a relaxation of natural selection, as compared with free-living organisms, because of a putative reduction in

their effective population sizes (Huyse *et al*., 2005; Woolfit and Bomham, 2003). Eventually, this can be detected since changes in the natural selection strength may imprint a plethora of characteristic molecular hallmarks on DNA and protein sequences. For instance, the relaxation of the intensity of natural selection can cause an increase of the nucleotide and amino acid substitution rates, a decrease in the selective constraint levels (increased values of $\omega = d_N/d_S$ parameter), and an increase in the mutational bias. The effect of such relaxation on the codon usage bias, however, is likely to be more complex: a reduction of codon bias if the bias is actively maintained by the action of natural selection, but an increase in case that the main responsible was the mutational force (Sharp *et al*., 2010). Here we have taken advantage of the availability of complete mtDNA data for a number of flatworm species to check this hypothesis. Nevertheless, we cannot analyze either the putative different patterns left on the evolutionary rates (there is no reliable data of divergence times) or its impact of selective constraint levels because of the high saturation of dS values.

Focused on the impact of nucleotide and codon bias, our results show a clear pattern; the parasitic platyhelminth species do not exhibit a higher relaxation of nucleotide selection than free-living species. On the contrary, three out of the four free-living species (Geoplanoidea representatives) exhibit higher mutational bias, at the A+T content, nucleotide frequency and codon usage levels. Moreover, our results also reveal that the observed codon bias is primarily caused by mutation and not by natural selection mechanisms. First, the species with higher mutational bias also exhibit greater codon bias (Figure 7B). Second, once adjusted the codon bias for the mutational input (SC values against different types of sites), the codon bias effect disappears (Figure 7C) and the clustering pattern separating free-living and parasitic species also vanishes. These results agree with data for bacteria (Sharp *et al*., 2010) but differ from plants, where a higher mutation rate for parasitic over non-parasitic groups has been observed (Bromham *et al.*, 2013) although the connection with a selective pressure is not clear.

In summary, despite parasites life cycles make them prone to suffer genetic bottlenecks leading to putative reductions on the effective population size, we did not find the molecular hallmark of the relaxed selection process. On the contrary, free-living triclads appears to exhibit higher levels of relaxed selection.

Acknowledgements

We want to thanks to Mrs. Jill McDonald who kindly contributed with samples of *M. terrestris*, to M. Gorchs that helped in the collection of *C. alpina*, and to Jitka Aldhoun. This research was supported by CGL 2008-00378 and CGL 2011-23466 Grants to M.R. by the doctoral fellowship BES−2009−022530 from the Ministerio de Ciencia e Innovacion from Spain (to E.S.), and by the fellowships SEG and REDES CGL2009/06185-E (to M.A.P).

References

Álvarez-Presas M, Sánchez-Gracia A, Carbayo F, Rozas J, Riutort M (2014). Insights into the origin and distribution of biodiversity in the Brazilian Atlantic forest hot spot: a statistical phylogeographic study using a low-dispersal organism. *Heredity* In press.

Ballard JWO, Melvin RG (2010). Linking the mitochondrial genotype to the organismal phenotype. *Mol Ecol* **19**:1523–1539.

Ballard JWO, Pitchaud N (2013). Mitochondrial DNA: more than an evolutionary bystander. *Funct Ecol* **28**:218–231.

Ballard JWO, Whitlock MC (2004). The incomplete natural history of mitochondria. *Mol Ecol* **13**:729–744.

Barr CM, Nelman M, Tayler DR (2005). Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. *New Phytol* **168**:39–50.

Benson G (1999). Tandem Repeats Finder: a program to analyze DNA sequences. *Nucleic Acids Res* **27**:573–580.

Bernardi G, Bernardi G (1986). Compositional constraints and genome evolution. *J Mol Evol* **24**:1–11.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, *et al.* (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Mol Phylogenet Evolution* **69**:313–319.

Bessho Y, Ohama T, Osawa S (1992a). Planarian mitochondria. I. Heterogeneity of cytochrome c oxidase subunit I gene sequences in the freshwater planarian, *Dugesia japonica*. *J Mol Evol* **34**:324–330.

Bessho Y, Ohama T, Osawa S (1992b). Planarian mitochondria. II. The unique genetic code as deduced from cytochrome c oxydase subunit I gene sequence. *J Mol Evol* **34**:331–335.

Bessho Y, Tamura S, Hori H, Tanaka H, Ohama T, Osawa S (1997). Planarian mitochondria sequence heterogeneity: relationships between the type of cytochrome c oxidase subunit I gene sequence, karyotype and genital organ. *Mol Ecol* **6**:129–136.

Biomatters (2014). *Geneious version 6.1.7*. Available at http://www.geneious.com/. Biomatters Limited.

Boore JL, Brown WM (1995). Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris*. *Genetics* **141**:305–319.

Bromham L, Cowman PF, Lanfear R (2013). Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evolutionary Biology* **13**:126.

Brown WM, George M, Wilson AC (1979). Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA* **76**:1967–1971.

Carbayo F, Álvarez-Presas M, Olivares CT, Marques FPL, Froehlich EM, Riutort M (2013). Molecular phylogeny of Geoplaninae (Platyhelminthes) challenges current classification: proposal of taxonomic actions. *Zool Scr* **42**:508–528.

Castellana S, Vicario S, Saccone C (2011). Evolutionary patterns of the mitochondrial genome in Metazoa: exploring the role of mutation and selection in mitochondrial protein coding gene. *Genome Biol Evol* **3**:1067–1079.

Castro LR, Austin AD, Dowton M (2002). Contrasting rates of mitochondrial molecular evolution in parasitic diptera and hymenoptera. *Mol Biol Evol* **19**: 1100–1113.

Chomczynski P, Sacchi N (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**:156–159.

Dowton M, Austin AD (1995). Increased genetic diversity in mitochondrial genes is correlated with the evolution of parasitism in the hymenoptera. *J Mol Evol* **41**:958–965.

Francino MP, Chao L, Riley MA, Ochman H (1996). Asymmetries generated by transcription-coupled repair in Enterobacterial genes. *Science* **272**:107–109.

Gissi C, Iannelli F, Pesole G (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity* **101**:301–320.

Huse SM, Huber JA, Morrison HG, Sogin ML, Welch M (2007). Accuracy and quality

of massively parallel DNA pyrosequencing. *Genome Biol* **8**:R143.

Huyse T, Buchmann K, Littlewood DTJ (2008). The mitochondrial genome of *Gyrodactylus derjavinoides* (Platyhelminthes: Monogenea) -- a mitogenomic approach for *Gyrodactylus* species and strain identification. *Gene* **417:**27–34.

Huyse T, Poulin R, Théron A (2005). Speciation in parasites: a population genetics approach. *Trends Parasitol* **10**:469–475.

Jacob JEM, Vanholme B, Leeuwen TV, Gheysen G (2009). A unique genetic code change in the mitochondrial genome of the parasitic nematode *Radopholus similis*. *BMC Res Notes* **2**:192.

Jeon HK, Kinm KH, Eom KS (2007). Complete sequence of the mitochondrial genome of *Taenia saginata*: comparison with *T. solium* and *T. asiatica*. *Parasitol Int* **56**:243–246.

Kang S, Kim J, Lee J, Min GS, Park JS (2012). The complete mitochondrial genome of an ectoparasitic monopisthocotylean fluke *Benedenia hoshinai* (Monogenea: Platyhelminthes). *Mitochondrial DNA* **23**:176**–**178.

Laslett D, Canbäck B (2008). ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* **24**:172–175.

Lawrie DS, Messer PW, Hershberg R, Petrov DA (2013). Strong purifying selection at synonymous sites in *D. melanogaster*. *PLoS Genet* **9**:e1003527.

Lázaro EM, Sluys R, Pala M, Stocchino GA, Baguñà J, Riutort M (2009). Molecular barcoding and phylogeography of sexual and asexual freshwater planarians of the genus *Dugesia* in the Western Mediterranean (Platyhelminthes, Tricladida, Dugesiidae). *Mol Phylogenet Evol* **52**:835–845.

Le TH, Blair D, Agatsuma T, Humair PF, Campbell NJ, Iwagami M, Littlewood DT, Peacock B, *et al*. (2000). Phylogenies inferred from mitochondrial gene orders **–** a cautionary tale from the parasitic flatworms. *Mol Biol Evol* **17**:1123**–**1125.

Le TH, Blair D, McManus DP (2001). Complete DNA sequence and gene organization of the mitochondrial genome of the liverfluke, *Fasciola hepatica* L. (Platyhelminthes; Trematoda). *Parasitology* **123**:609–621.

Le TH, Pearson MS, Blair D, Dai N, Zhang LH, McManus DP (2002). Complete mitochondrial genomes confirm the distinctiveness of the horse-dog and sheep-dog strains of *Echinococcus granulosus*. *Parasitology* **124**:97–112.

Marin A, Xia X (2008). GC skew in protein-coding genes between the leading and lagging strands in bacterial genomes: new substitution models in bacterial

genomes: new substitution models in corporating strand bias. *J Theor Biol* **253**:508–513.

Nakao M, Sako Y, Ito A (2003). The mitochondrial genome of the tapeworm *Taenia solium*: A finding of the abbreviated stop codon U. *J Parasitol* **89**:633–635.

Necsulea A, Lobry JR (2007). A new method for assessing the effect of replication on DNA base composition asymmetry. *Mol Ecol Biol* **24**:2169–2179.

Perna NT, Kocher TD (1995). Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J Mol Evol* **41**:353–358.

Poh YP, Ting CT, Fu HW, Langley CH, Begun DJ (2012). Population geneomic analysis of base composition evolution in *Drosophila melanogaster*. *Genome Biol Evol* **4**:1245–1255.

Riutort M, Lázaro EM, Álvarez−Presas M, Solà E, Paps, J (2012). Evolutionary history of the Tricladida and the Platyhelminthes: an up−to−date phylogenetic and systematic account. *Int J Dev Biol* **56**:5−17.

Ruiz-Trillo I, Riutort M, Fourcade HM, Baguñà J, Boore JL (2004). Mitochondrial genome data support the basal position of Acoelomorpha and the polyphyly of the Platyhelminthes. *Mol Phylogenet Evol* **33**:321−332.

Sakai M, Sakaizumi M (2012). The Complete Mitochondrial Genome of *Dugesia japonica* (Platyhelminthes; Order Tricladida). *Zoolog Sci* **29**:672−680.

Schattner P, Brooks AN, Lowe TM (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* **33**:W686−689.

Schmieder R, Edwards R (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**:863−864.

Sharp PM, Emery LR, Zeng K (2010). Forces that influence the evolution of codon vias. *Phil Trans R Soc B* **365**:1203−1212.

Shields DC, Sharp PM, Higgins DG, Wright F (1988). "Silent" sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. *Mol Biol Evol* **5**:704–716.

Sluys R, Solà E, Gritzalis K, Vila−Farré M, Mateos E, Riutort M (2013). Integrative delineation of species of Mediterranean freshwater planarians (Platyhelminthes: Tricladida: Dugesiidae). *Zool J Linn Soc* 169:523−547.

Solà E, Sluys R, Gritzalis K, Riutort M (2013). Fluvial basin history in the northeastern Mediterranean underlies dispersal and speciation patterns in the genus *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae). *Mol Phylogenet Evol* **66**:877−888.

Solà E, Stocchino GA, Manconi R, Leria L, Riutort M. *Dugesia* (Platyhelminthes, Continenticola), a widespread and morphologically homogeneous living genus from the Mesozoic. In preparation.

Stöck M, Moritz C, Hickerson M, Frynta D, Dujsebayeva T, Eremchenko V, *et al.* (2006). Evolution of mitochondrial relationships and biogeography of Palearctic green toads (*Bufo viridis* subgroup) with insights in their genomic plasticity. *Mol Phylogenet Evol* **4**:663–689.

Telford MJ, Herniou EA, Russell RB. Littlewood DTJ (2000). Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. *Proc Natl Acad Sci U S A* **97**:11359−11364.

Thompson RC, Boxell AC, Ralston BJ, Constantine CC, Hobbs RP, Shury T, *et al.* (2006). Molecular and morphological characterization of *Echinococcus* in cervids from North America. *Parasitology* **132**:439−447.

Tillier ER, Collins RA (2000). The contributions of replication orientation, gene direction, and signal sequences to base-composition asymmetries in bacterial genomes. *J Mol Evol* **50**:249−257.

Tryfonopoulos GA, Thanou EG, Fraguedakis-Tsolis SE, Chondropoulos BP (2010). New data on the distribution and genetic structure of Greek moles of the genus *Talpa* (Mammalia, Talpidae). *J Zoolog Syst Evol Res* **48**:188–193.

Weber M, Wey AR, Podsiadlowski L, Witek A, Schill RO, Sugár L, *et al.* (2013). Phylogenetic analyses of endoparasitic Acanthocephala based on mitochondrial genomes suggest secondary loss of sensory organs. *Mol Phylogenet Evol* **66**:182–189.

Wei SJ, Shi M, He JH, Sharkey M, Chen XX (2009). The complete mitochondrial genome of *Diadegma semiclausum* (hymenoptera: ichneumonidae) indicates extensive independent evolutionary events. *Genome* **52**:308−319.

Wey-Fabrizius AR, Podsiadlowski L, Herlyn H, Hankeln T (2013). Platyzoan mitochondrial genomes. *Mol Phylogenet Evol* **69**:365−375.

Woolfit M, Bromham L (2003). Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population size. *Mol Biol Evol* **20**:1545−1555.

Wright F (1990). The 'effective number of codons' used in gene. *Gene* **87**:23−29.

Wyman SK, Jansen RK, Boore JL (2004). Automatic Annotation of Organellar Genomes with DOGMA. *Bioinformatics* **20**:3252−3255.

Yamasaki H, Ohmae H, Kuramochi T (2012). Complete mitochondrial genomes of *Diplogonoporus balaenopterae* and *Diplogonoporus grandis* (Cestoda: Diphyllobothriidae) and clarification of their taxonomic relationships. *Parasitology Int* **61**:260−266.

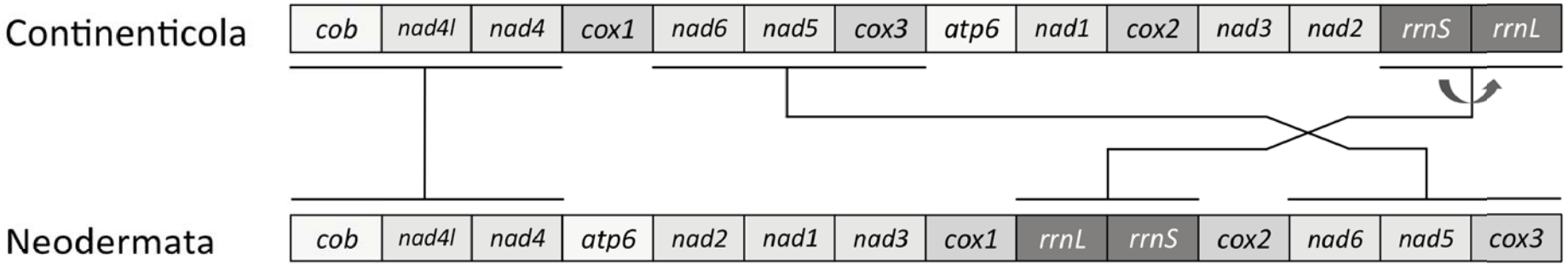
Zhang J, Wu X, Li Y, Xie M, Li A (2014). The compelete mitochondrial genome of *Tetrancistrum nebulosi* (Monogenea: Ancyrocephalidae). *Mitochondrial DNA* In press.

Zhao GH, Li J, Song HQ, Li XY, Chen F, Lin RQ, *et al.* (2012). A specific PCR assay for the identification and differentiation of *Schistosoma japonicum* geographical isolates in mainland China based on analysis of mitochondrial genome sequences. *Infect Genet Evol* **12**:1027−1036.

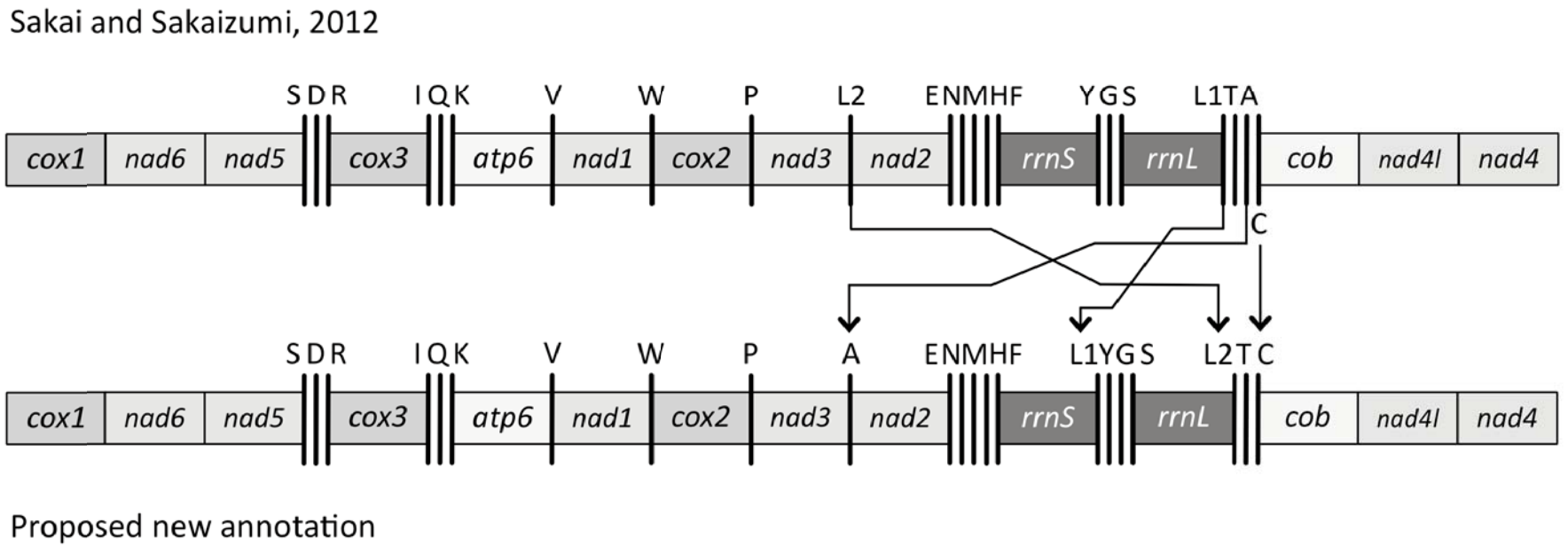
Supplementary Information

Supplementary Figures & Supplementary Tables

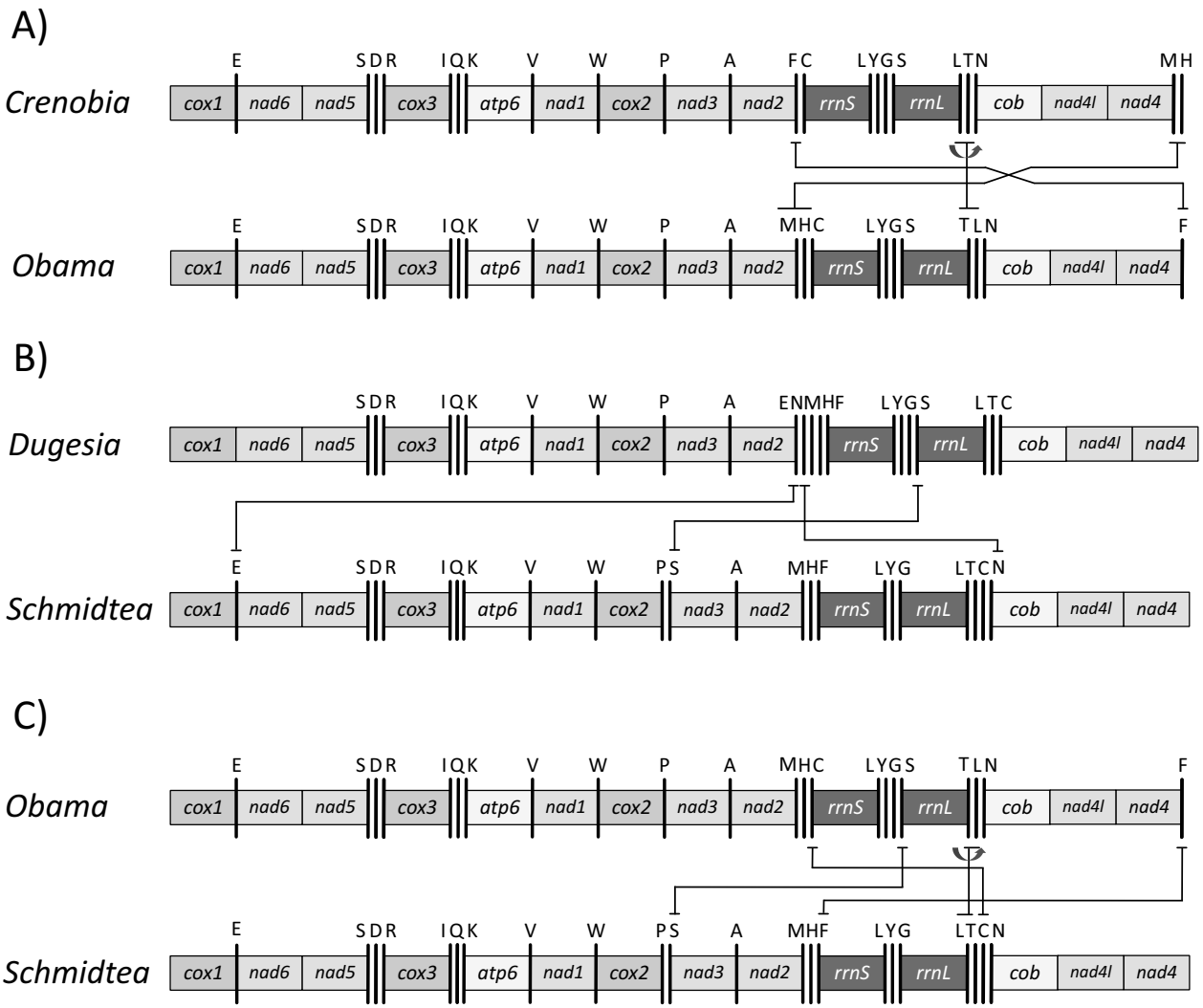
Supplementary Figure 1. Linearized scheme showing a comparison of the general protein coding and ribosomal genes between generalized mitogenomes of Neodermata (parasitic platyhelminths) and Continenticola (free-living platyhelminths).



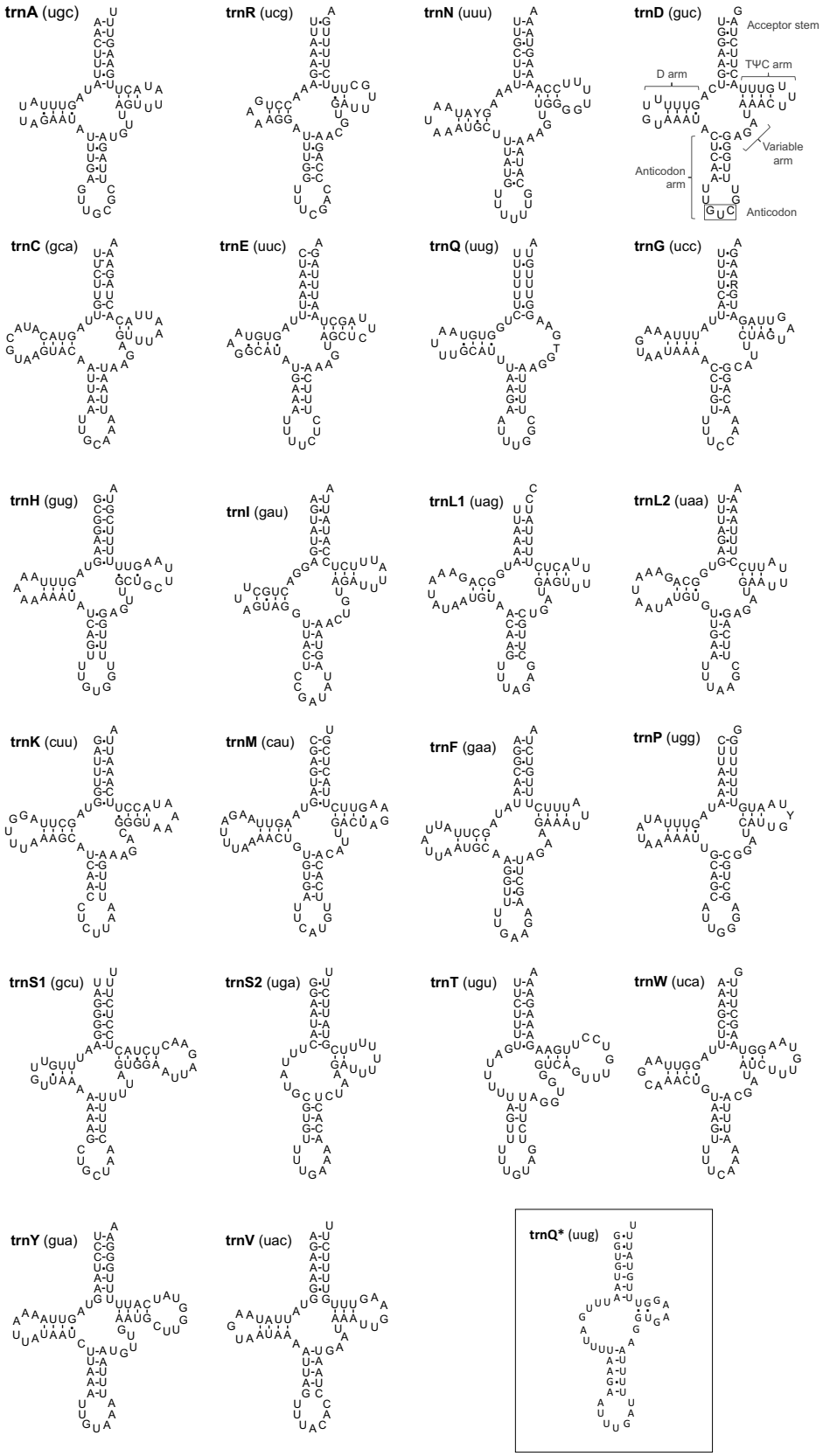
Supplementary Figure 2. Comparison between the annotation proposed by Sakai and Sakaizumi for *Dugesia japonica* and the new annotation proposed in the present study.



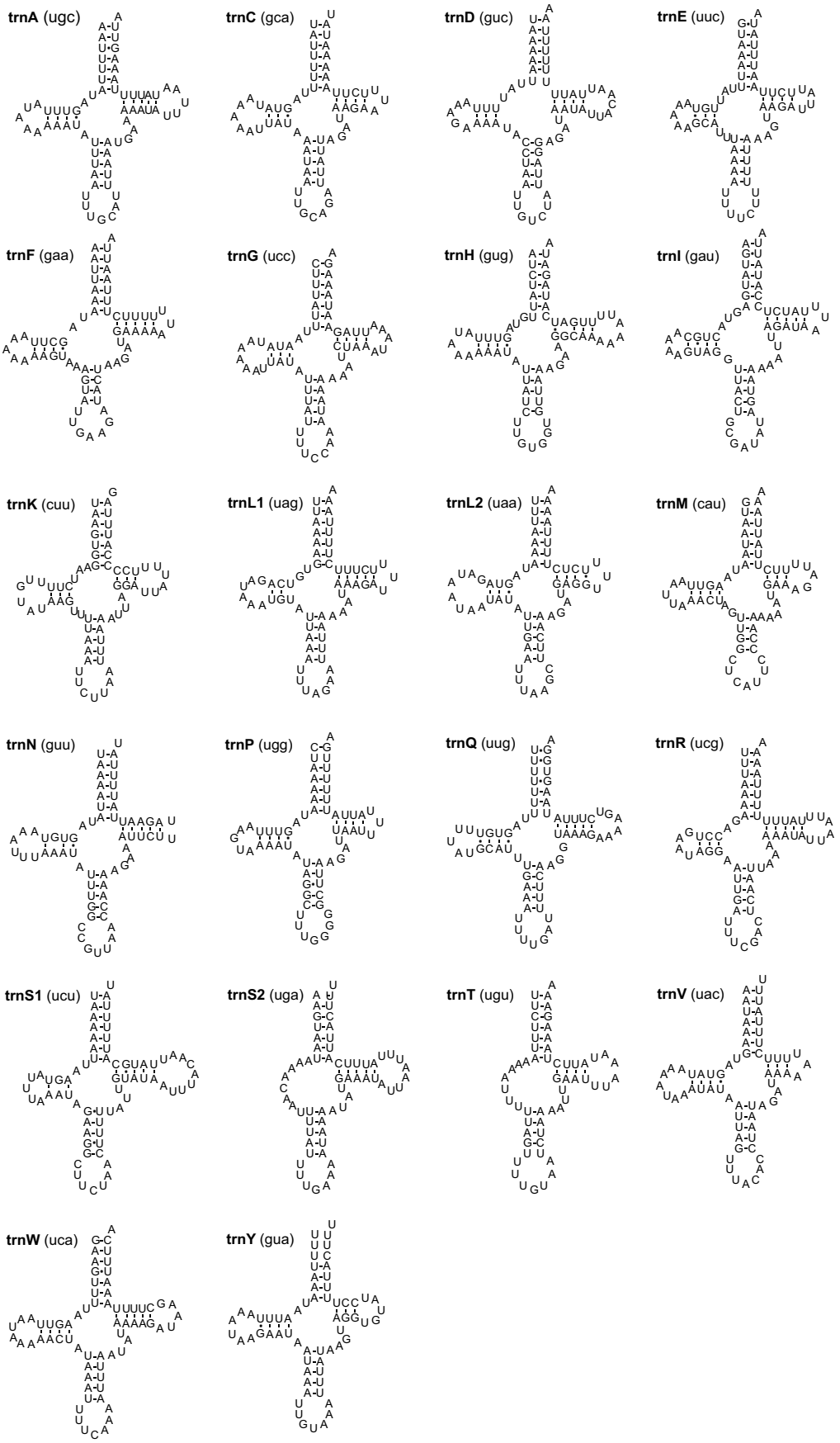
Supplementary Figure 3. Comparison by pairs of the tRNA order of the different Tricladida species included in this work.



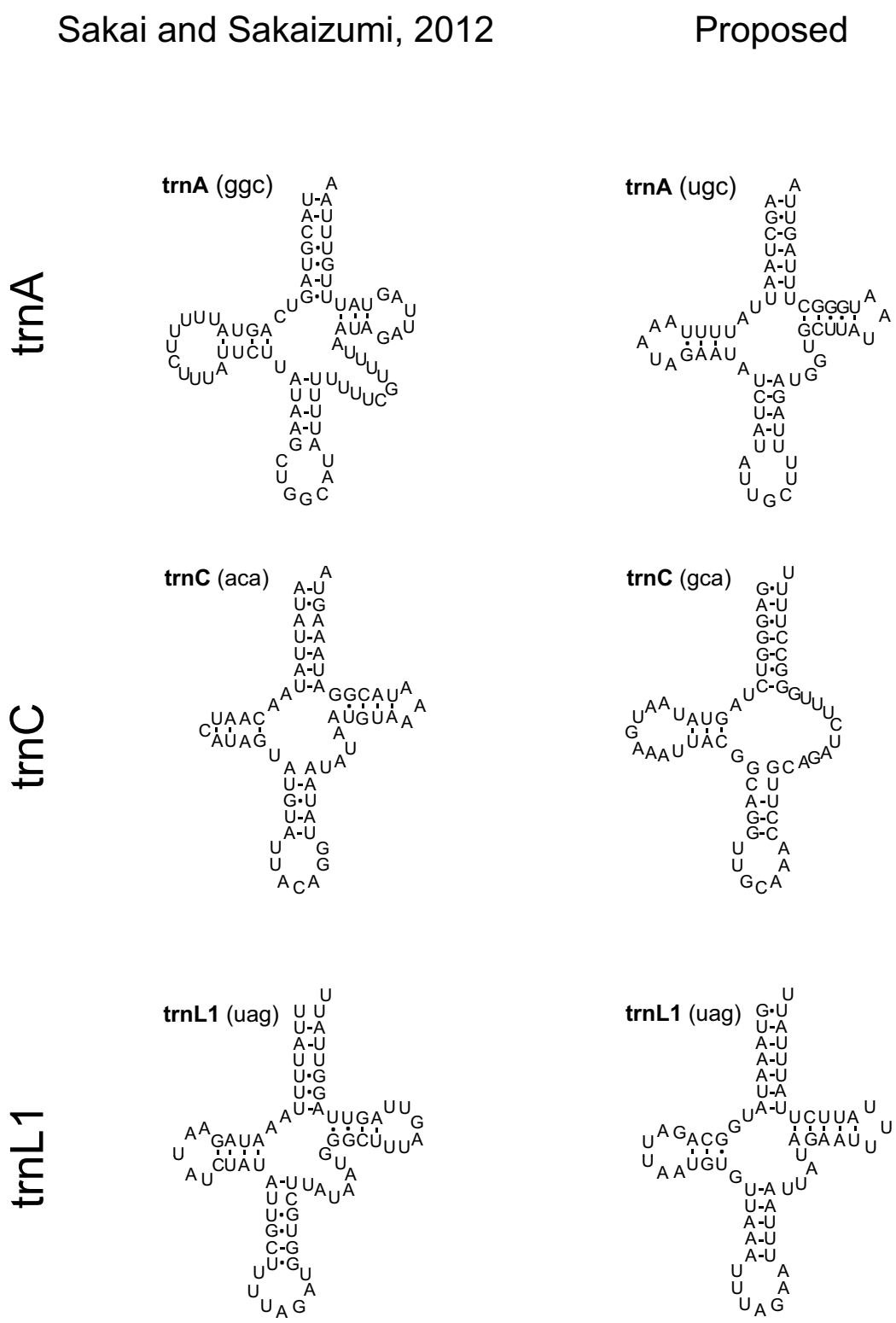
Supplementary Figure 4. Secondary structure of the 22 tRNA of *Crenobia alpina*. *trnQ** in a box shows the alternative structure proposed for this tRNA. The different tRNA parts are showed on *trnD*.



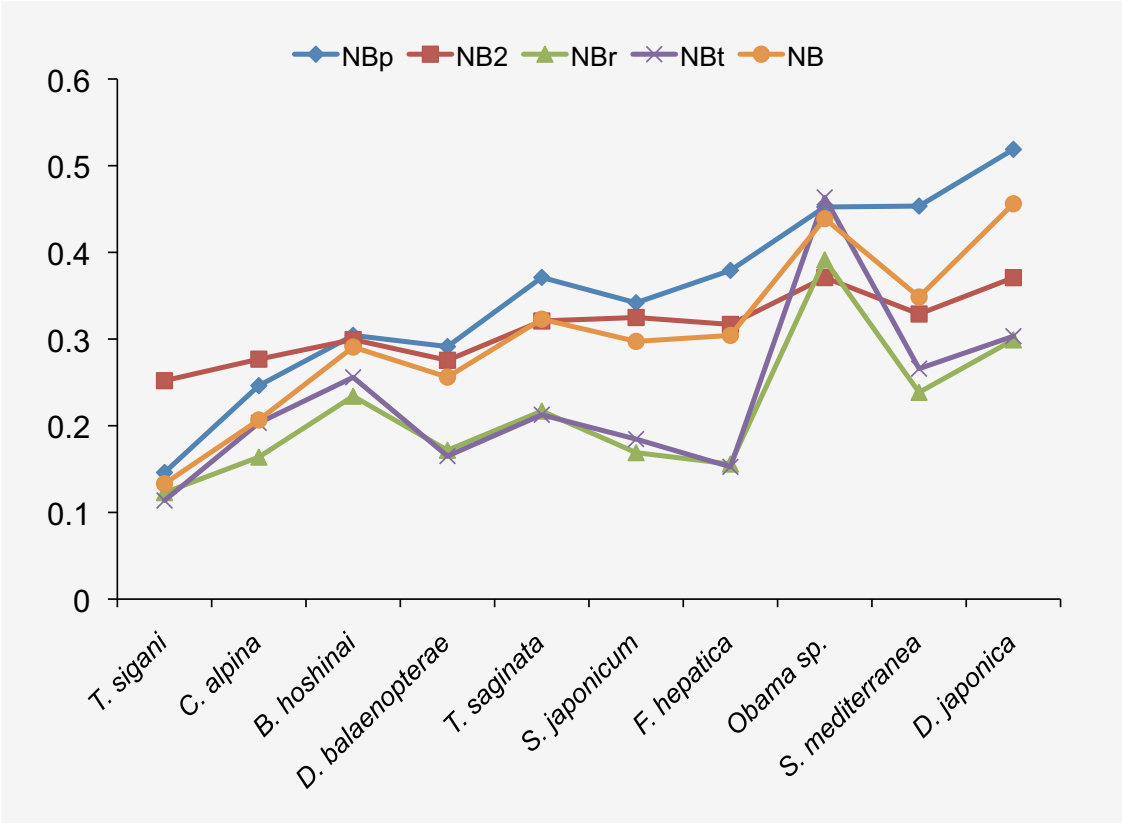
Supplementary Figure 5. Secondary structure of the 22 tRNA of *Obama* sp.



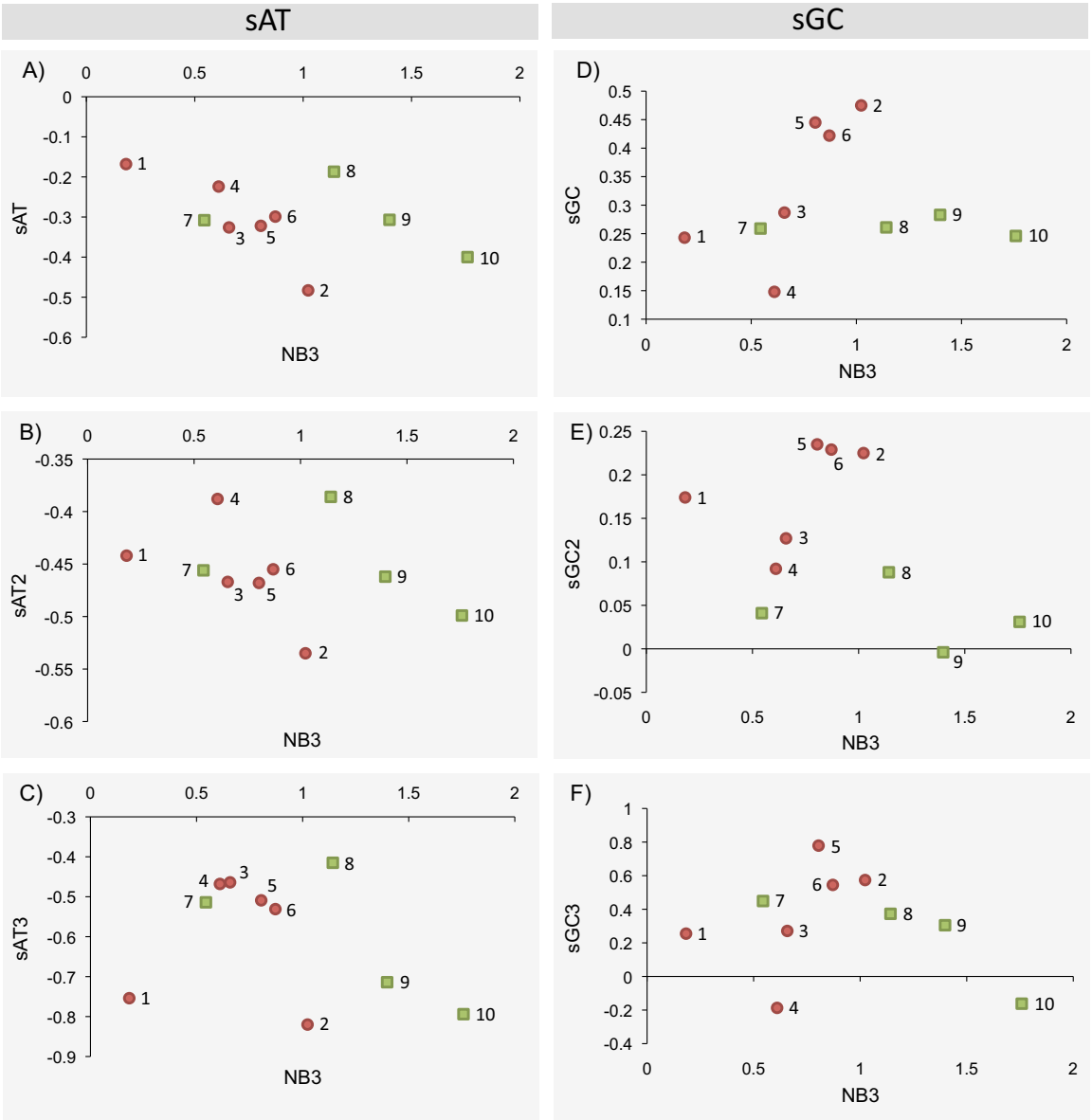
Supplementary Figure 6. Comparison between the Sakai and Sakaizumi (2012) *trnA*, *trnC* and *trnL1* secondary structure for *Dugesia japonica* based on their annotation and the secondary structure based on our new proposed annotation.



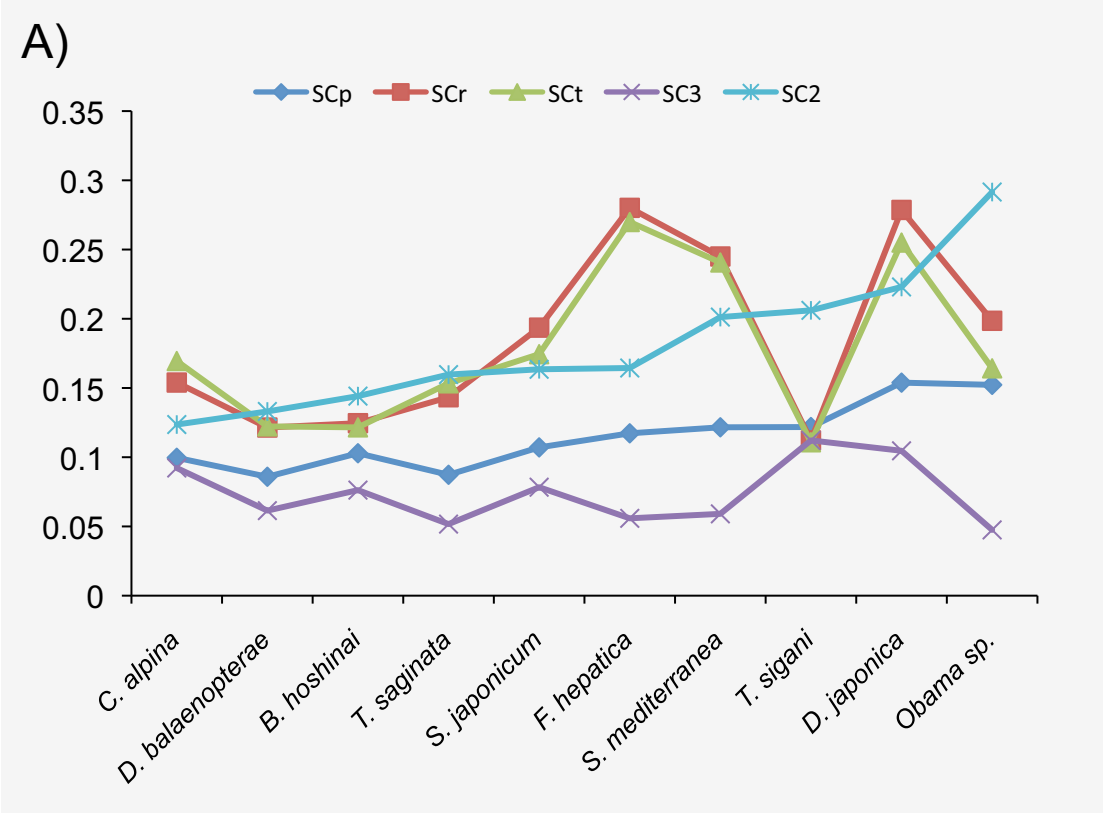
Supplementary Figure 7. Values of the different NB-based statistic across species excluding the NB3 (NB at the third position of four-fold degenerate codons).

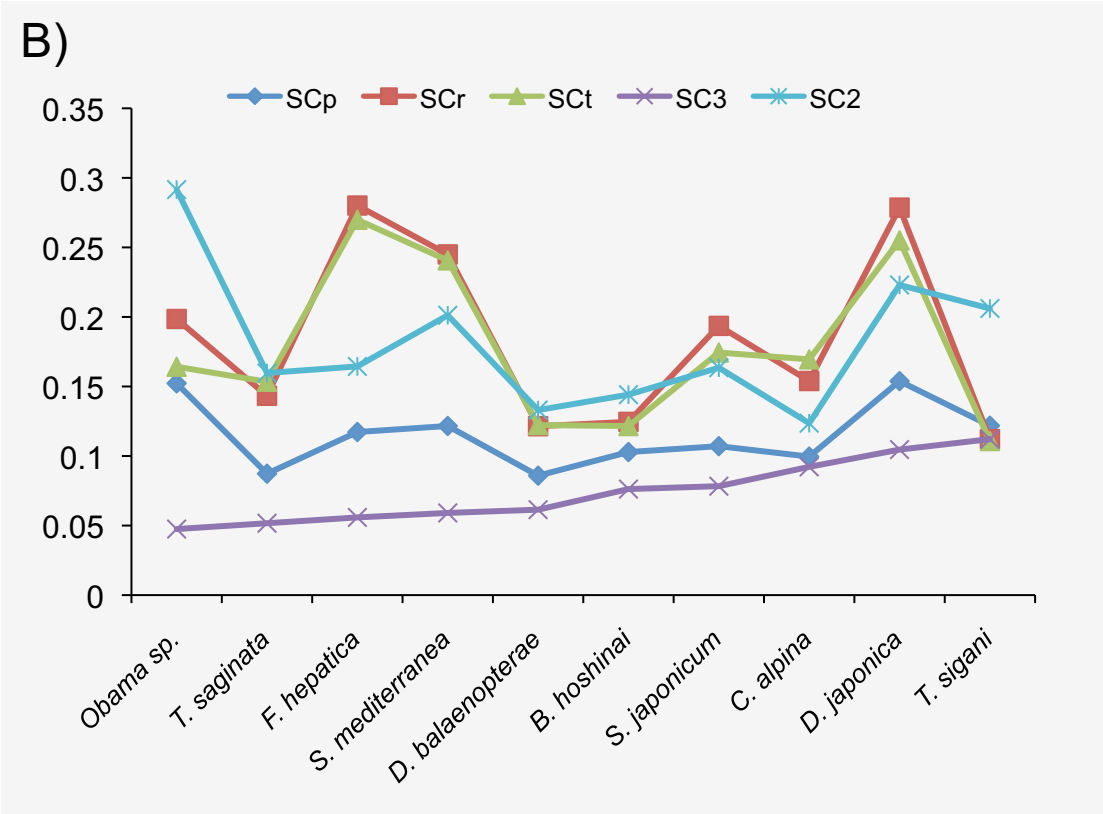


Supplementary Figure 8. Relationship between sAT, sGC values and NB3. sAT general skew; sAT2, sAT skew at the second positions; sAT3, sAT at the third positions. sGC, general skew; sGC2, sGC skew at the second positions; sGC3, sGC at the third positions. Green squares and red circles indicate free-living and parasitic platyhelminthes, respectively. The surveyed species are shown in numbers: 1, *T. sigani*; 2, *F. hepatica*; 3, *D. balaenopterae*; 4, *B. hoshinai*; 5, *T. saginata*; 6, *S. japonicum*; 7, *C. alpina*; 8, *Obama* sp.; 9, *S. mediterranea*; 10, *D. japonica*.



Supplementary Figure 9. SC values across species adjusted for the observed mutation bias. Ordered ascending based on the Chi scales values for A) second positions of the PCG and B) for the third position of four-fold degenerate codons equifrequency.





Supplementary Table 1. Locality and habitat information on species collected for this study.

Species	Triclad family	Locality	Habitat	Coordinates	Collectors (collection date)
Crenobia alpina (Dana, 1766)	Planariidae	Rasos de Peguera, Catalonia	Spring	398813.00 E, 4661021.72 N	M. Riutort, E. Solà, M. Álvarez-Presas, L. Leria (02/11/2010)
Polycelis felina (Dalyell, 1844)	Planariidae	Viladrau, Catalonia	Spring	449819.14 E, 4633075.96 N	M. Riutort, E. Mateos, M. Álvarez-Presas (26/10/2010)
Dugesia subtentaculata (Draparnaud, 1801)	Dugesiidae	Santa Fe del Montseny, Catalonia	River	455638.94 E, 4624799.19 N	M. Riutort, E. Solà, M. Álvarez-Presas (22/10/2010)
Microplana terrestris (Müller, 1774)	Geoplanidae	Kirby Wiske, UK	Garden	602480.45 E, 6013160.19 N	J. McDonald (19/10/2010)
Obama sp. (6) Carbayo *et al.*, 2013	Geoplanidae	Torruella de Fluvià, Catalonia	Garden	502271.00 E, 4670489.97 N	M. Riutort (12/10/2010)

Supplementary Table 2. Data of mitochondrial proteins used to conduct the tBLASTx analyses in order to detect whether the mitochondrial genes were present in the 454 sequencing reads.

Protein name	Locus	*Fasciola hepatica* Protein product	*Gyrodactylus derjavinoides* Protein product	*Taenia solium* Protein product
cytochrome c oxidase subunit III	*cox3*	NP_0666217.2	YP_001974722.1	NP_659225.1
cytochrome b	*cob*	NP_0662218.2	YP_001974723.1	NP_659226.1
NADH dehydrogenase subunit 4L	*nad4l*	NP_0662219.1	YP_001974724.1	NP_659227.1
NADH dehydrogenase subunit 4	*nad4*	NP_066220.2	YP_001974725.1	NP_659228.1
ATP synthase F0 subunit 6	*atp6*	NP_066221.2	YP_001974726.1	NP_659229.1
NADH dehydrogenase subunit 2	*nad2*	NP_066222.2	YP_001974727.1	NP_659230.1
NADH dehydrogenase subunit 1	*nad1*	NP_066223.2	YP_001974728.1	NP_659231.1
NADH dehydrogenase subunit 3	*nad3*	NP_066224.2	YP_001974729.1	NP_659232.1
cytochrome c oxidase subunit I	*cox1*	NP_066225.2	YP_001974730.1	NP_659233.1
cytochrome c oxidase subunit II	*cox2*	NP_066226.2	YP_001974731.1	NP_659234.1
NADH dehydrogenase subunit 6	*nad6*	NP_066227.2	YP_001974732.1	NP_659235.1
NADH dehydrogenase subunit 5	*nad5*	NP_066228.2	YP_001974733.1	NP_659224.1

Supplementary Table 3. Primers designed for the reamplification of *Crenobia alpina*.

Name		Sequence 5'-3'	Annealing T (ºC)	Genes
Tinc	F	GATTGCTACGGGTTTGG	49	*cob*[a]; *nad4l*[a]
Gana	R	CACATTCCTCTTATCCC	42.2	
Joan	F	GTGAAGGTTTTGGGG	44.1	*nad4l*; *nad4*[a]
Dora	R	CCCTTCCAACACTCC	44	
Ste	F	GGTTGGTGTTTTCGG	45.3	*nad4*[a]; *trnM (cau)*; *trnH (gug)*
Phen	R	CAACCAAAACCGCCAAG	42.8	
Dar	F	GGGTTGAAAGATGTGCGG	54.2	*cox1*[a]
Win	R	CCAAAACCGCCAATC	48.2	
Ice	F	GTATTTCTTTGGGGTTGG	46.8	*cox1*[a]
Age	R	CTCCCCAGCCATTCC	50.1	
Dino	F	GGGTTCTTTATTGTCTTTGCTTAGCG	47.2	*cox1*[a]; *trnE (uuc)*; *nad6*; *nad5*[a]
Saure	R	CAGCGAGCATTGTGAATAGTCC	45.7	
Rap	F	CCCAGTATCCTTTTTC	39.4	*nad5*[a]
Tor	R	ACAAGCATAAAGTATTCCC	43.2	
Chi	F	TCTTTTGTCCGCTTCTG	47.4	*nad5*[a]; *trnS2 (ugc)*; *trnD (guc)*; *trnR (ucg)*; *cox3*; *trnI (gau)*; *trnQ (uug)*[a]
Cago	R	CCGAAATACAAACCTTC	42.6	
Angi	F	CACTCTTCTTTGCGTTG	45.2	*cox3*[a]; *trnI (gau)*; *trnQ (uug)*; *trnK (cuu)*; *atp6*; *trnV (uac)*
Laia	R	CAACAACCCCCAAAAC	47	
Ptero	F	GGGTGTATGTGGACTTTTG	47.8	*atp6*[a]; *trnV (uac)*; *nad1*; *trnW (uca)*; *cox2*
Dactil	R	GAAACAATCTAACTGCTCC	43.7	
Wil	F	CTTTGCTTGGTCCATTG	47.6	*nad1*[a];
Son	R	CCACGACGCTTCTCCTC	52.3	
Eva	F	GAGTGTGGTTTTGATGG	44.2	*nad3*[a]; *trnA (ugc)*; *nad2*[a]
Ona	R	CCCAGAAAACACAAAGAAAC	48.7	
Cholo	F	GTGTTCTCTTATGTCTCC	38.3	*nad2*[a]; *trnF (gaa)*; RR[a]
Epus	R	CCCCTTATTTTCCAC	40.1	
Trilo2	F	GGGAAATAGAAGGAGGG	45.9	
Bite	R	CTAAGGGGAGGGTTGGG	51.9	
Brady	F	GTTGAAGAATGAGACTG	37.1	*trnC (gca)*; *rrnS*[a]
Pus	R	GAATAGTGACGGGCGGTG	54.2	
New	F	GAAAGATAGATAGAGGGG	39.5	*rrnS*[a]; *trnL1 (uag)*; *trnY (gua)*; *trnG (ucc)*; *rrnL*[a]
York	R	CCTTCATATTAAACCCGTTC	53.7	
Artro	F	GTATCCCCTGCTCGTTG	49.2	*rrnL*[a]; *trnL2 (uaa)*; *trnT (ugu)*; *trnN (guu)*; *cob*[a]
Pode	R	CAACCCTCTTCCCCAC	52.5	

[a] The gene is covered partially by the primers.

Supplementary Table 4. Primers designed for the reamplification of *Obama* sp.

Name		Sequence 5'-3'	Annealing T (ºC)	Genes
Grand	F	GAAAGKKAGGAGGTG	40.8	*cob*[a]; *nad4l*[a]; *nad4*[a]
Jete	R	CTTTAHATCAWACTGAC	37.1	
Kete	F	CATGGTTTTTGTTCTTC	50.6	*nad4*[a]; *trnF (gaa)*; *cox1*[a]
Peten	R	CCAAAACCACCAATC	51.9	
Ni	F	GGTTTTATTGTTTGAGC	49.9	*cox1*[a]; *trnE (uuc)*; *nad6*; *nad5*[a]
Jinsky	R	CCATCYCAACCAAAC	48.6	
Pau	F	CTGCTTTAGTTCATTC	44.6	*nad5*[a]; *trnS2 (uga)*; *trnD (guc)*; *trnR (ucg)*; *cox3*[a]
Lova	R	GWAAACCATGAAAACCAG	50.2	
Bat	F	GCAGYTTGATATTGRC	46.4	*cox3*[a]; *trnI (gau)*; *trnQ (uug)*; *trnK (cuu)*; *atp6*; *trnV (uac)*; *nad1*[a]
Man	R	CGAATCTGBATATABCTC	40.4	
Porde	F	GGTTCTTTDGARTTTGC	48.7	*nad1*[a]; *trnW (uca)*; *cox2*; *trnP (ugg)*; *nad3*[a]
Bra	R	GMARACGAGAMATATAC	23.7	
Enri	F	GARGAATTRCGTHGTGG	39.2	*nad3*[a]; *trnA (ugc)*; *nad2*[a]
Kito	R	GAAGATYCAARCC	29	
Ene	F	GGYTTGRTCTTC	27	*nad2*[a]; *trnM (cau)*; *trnH (gug)*; *trnC (gca)*; *rrnS*[a]
Sim	R	GYTGCTGGCACYC	35	
Valen	F	GTTAGTGTACGGTTG	42.2	*rrnS*[a]; *trnL1 (uag)*; *trnY (gua)*; *trnG (ucc)*; *trnS1 (ucu)*; *rrnL*[a]
Tin	R	CGGTCTAAACTCAAATC	49.8	
Demi	F	CGAAAAGACCCTACAG	50.7	*rrnL*[a]; *trnT (ugu)*; *trnL2 (uaa)*; *trnN (guu)*; RR; *cob*[a]
Plie	R	GTAATAACAGTAGCDCC	42.9	

[a] The gene is covered partially by the primers.

Supplementary Table 5. Different Statistics for the 454 sequencing.

	Polycelis felina	*Dugesia subtentaculata*	*Crenobia alpina*	*Obama* sp.
Total number (raw) of reads	20,249	5,060	31,282	11,671
Mb	2.82	0.61	6.19	3.77
N50 (in bp)	246	146	377	427
Average length (in bp)	139.26	119.82	197.96	323.22
Number of removed reads				
E. coli genome seq.	3	11	3	221
Uni Vecdb	346	124	545	102
Low quality	132	27	205	114
Length< 50 bp	4,716	1,131	5,399	720
Number of reads used for the assembling	15,052	3,767	25,130	10,614
Mb	2.58	0.54	5.88	3,59
N50 (in bp)	270	161	382	425

Supplementary Table 6. Summary of tBLASTn hits for raw reads against the mitochondrial proteins of the three parasitic flatworms: *Fhep*, *Fasciola hepatica*; *Gder*, *Gyrodactylus derjavinoides*; *Tsol*, *Taenia solium*.

	Polycelis felina			*Dugesia subtentaculata*			*Crenobia alpina*			*Obama* sp.		
	Fhep	*Gder*	*Tsol*	*Fhep*	*Gder*	*Tsol*	*Fhep*	*Gder*	*Tsol*	*Fhep*	*Gder*	*Tsol*
cox3	0	1	0	0	0	0	2	0	0	2	0	0
cob	1	1	1	6	7	7	55	67	62	58	64	64
nad4l	0	0	0	0	0	0	0	0	0	0	0	0
nad4	1	2	1	0	1	1	27	16	15	11	17	10
atp6	0	0	0	0	0	0	0	2	1	1	2	1
nad2	0	0	0	0	0	0	0	0	0	0	0	0
nad1	0	0	0	1	1	3	48	45	52	41	44	43
nad3	0	0	0	0	0	0	0	0	8	0	0	0
cox1	1	1	1	7	7	7	126	132	117	176	191	187
cox2	0	0	0	6	6	6	43	47	45	30	33	29
nad6	0	0	0	0	0	0	0	0	0	0	0	0
nad5	5	4	4	1	1	1	51	49	48	37	40	41

Supplementary Table 7. Annotation table for the mitochondrial genome of *C. alpina*.

Gene	Start	End	Start codon	Stop Codon	Size (bp)
cob	1	1113	ATG	TAG	1113
nad4l	1157	1390	TTG	TAG	234
nad4	1359	2714	ATG	TAA	1356
trnM (cau)	2722	2789			68
trnH (gug)	2925	2988			64
cox1	3359	5113	ATG	TAG	1755
trnE (uuc)	5117	5177			61
nad6	5179	5664	ATG	TAA	486
nad5	5665	7314	ATG	TAG	1650
trnS2 (uga)	7247	7479			53
trnD (guc)	7483	7544			62
trnR (ucg)	7545	7601			57
cox3	7633	8430	ATG	TAG	798
trnI (gau)	8429	8492			64
trnQ (uug)	8493	8547			55
trnK (cuu)	8553	8620			68
atp6	8624	9277	TTG	TAA	654
trnV (uac)	9280	9344			65
nad1	9511	10235	ATG	TAA	825
trnW (uca)	10234	10298			65
cox2	10302	11021	ATG	TAR	720
trnP (ugg)	11030	11093			64
nad3	11118	11450	TTG	TAA	333
trnA (ugc)	11449	11509			61
nad2	11515	12486	TTA	TAG	972
trnF (gaa)	12485	12550			66
trnC (gca)	14582	14646			65
rrnS	14676	15308			633
trnL1 (uag)	15310	15374			65
trnY (gua)	15372	15441			70
trnG (ucc)	15444	15507			64
trnS1 (ucu)	15509	15575			67
rrnL	15637	16494			858
trnL2 (uaa)	16656	16719			64
trnT (ugu)	16720	16777			64
trnN (guu)	16784	16852			69

Supplementary Table 8. Annotation table for the mitochondrial genome of *Obama* sp.

Gene name	Start	End	Start Codon	Stop Codon	Size (bp)
cob	1	1182	ATG	TAG	1182
nad4l	1166	1399	TTA	TAG	234
nad4	1368	2789	ATG	TAA	1422
TrnF (gaa)	2789	2853			65
cox1	2854	4590	GTG	TAA	1737
trnE (uuc)	4593	4656			64
nad6	4658	5134	TTG	TAA	477
nad5	5138	6772	TTG	TAA	1635
trnS2 (uga)	6773	6834			62
trnD (guc)	6835	6901			67
trnR (ucg)	6897	6961			65
cox3	6962	7753	ATG	TAA	792
trnI (gau)	7754	7819			66
trnQ (uug)	7821	7887			67
trnK (cuu)	7880	7944			65
atp6	7945	8619	TTG	TAA	675
trnV (uac)	8619	8681			63
nad1	8684	9577	ATG	TAA	894
trnW (uca)	9580	9647			68
cox2	9648	10427	TTG	TAA	780
trnP (ugg)	10429	10491			63
nad3	10510	10851	TTG	TAA	342
trnA (ugc)	10852	10918			67
nad2	10925	11923	ATG	TAA	999
trnM (cau)	12076	12138			63
trnH (gug)	12144	12212			69
trnC (gca)	12213	12275			63
rrnS	12311	12972			662
trnL1 (uag)	12973	13036			64
trnY (gua)	13037	13099			63
trnG (ucc)	13096	13160			65
trnS1 (ucu)	13161	13229			69
rrnL	13259	14178			920
trnT (ugu)	14181	14237			57
trnL2 (uaa)	14238	14301			64
trnN (guu)	14302	14366			65

Supplementary Table 9. Annotation table for the mitochondrial genome of *S. mediterranea*.

Gene	Start	End	Start Codon	Stop Codon	Size (bp)
cob	1	1101	TAT*	TAG	1101
nad4l	1068	1361	ATG	TAG	294
nad4	1312	2688	ATG	TAG	1377
cox1	3391	5023	TAT*	TAA	1633
trnE (uuc)	5024	5085			62
nad6	5096	5542	ATG	TAG	447
nad5	5539	7134	TTA*	TAA	1596
trnS2 (uga)	7138	7196			59
trnD (guc)	7197	7258			62
trnR (ucg)	7258	7319			62
cox3	7317	8144	TAT*	TAA	828
trnI (gau)	8138	8205			68
trnQ (uug)	8204	8268			65
trnK (cuu)	8268	8331			64
atp6	8334	8969	ATG	TAG	636
trnV (uac)	8971	9033			63
nad1	9030	9920	ATG	TAA	891
trnW (uca)	9924	9988			65
cox2	9989	10867	TTG*	TAA	879
trnP (ugg)	10971	11041			71
trnS1 (ucu)	11071	11123			53
nad3	11126	11395	TTG*	TAG	270
trnA (ugc)	11400	11468			69
nad2	11547	12416	ATG	TAA	870
trnM (cau)	24128	24190			63
trnH (gug)	24193	24259			67
trnF (gaa)	24263	24328			66
rrnS	24330	25036			707
trnL1 (uag)	25038	25100			63
trnY (gua)	25106	25171			66
trnG (ucc)	25177	25245			69
rrnL	25256	26160			905
trnL2 (uaa)	26161	26223			63
trnT (ugu)	26224	26277			54
trnC (gca)	26292	26351			60
trnN (guu)	26361	26424			64

* Start codon not found.

Section IV

General discussion

General discussion

In the present thesis we have focused our research efforts on the study of the distribution patterns and diversification processes of the freshwater planarian genus *Dugesia*. With this aim, we have used molecular phylogenetics and biogeographical methodologies.

Freshwater flatworms are so unresearched from the evolutionary point of view that every single new contribution to their knowledge is of great value. This lack of knowledge may be explained because of freshwater flatworms are not the most fancy group to carry out biological studies as far as they are very challenging and hard to explore for different reasons, such as the lack of useful fossils for the whole phylum or their reproductive and karyological plasticity.

Research on the correlation between geological-climatic events and biogeographical patterns on freshwater platyhelmints using molecular phylogenetics has been previously done only once (Lázaro *et al.*, 2011). This approach was carried out on the *Dugesia* sister genus *Schmidtea*, an inhabitant of the Western Mediterranean. Our results on the historical biogeography for the whole distribution of *Dugesia* in general and for the Greek region in particular, strongly support such historical correlation between paleogeographical events and the genus distribution and diversification patterns.

However, some of our results must be taken cautiously due to the inherent problems of biogeographical approaches. One example is the case of the sister relationship between the Cretan species *Dugesia cretica* and the rest of Greek species that had been initially interpreted by our work as the result of the well-known geological event of isolation of Crete from a former landmass called Ägäis in first place. Therefore, this was considered a suitable calibration point for the estimation of divergence times for *Dugesia* (Chapter 1). Later, we obtained samples from more distant places such as Turkey, Israel and Iran which proved to be closely related with different lineages of the Greek animals. On the other hand, during the development of the biogegraphic studies we also found 'outliers' that have been explained by a probable human-mediated dispersion. This kind of situations made us very cautious again about biogeographical interpretations. However, as seen between the first and the second

historical biogeography works, the increase of information permits more accurate and reliable answers to the curious diversification and distribution patterns of *Dugesia*.

Dugesia is a diverse genus of freshwater planarian with 81 formally described and valid species (four are new species described here) and many other that are proposed as candidate species but not formally described yet or others considered as *species inquirenda* (Sluys *et al.*, 2013; Annexes 1 − Tables 1 and 2). According to the homogeneity of the *Dugesia* morphology, its extremely wide distribution and its putative old age (Chapter 2), it is very likely that the speciation of this group has been mainly driven by vicariant events across its present distribution after ancient events of wide dispersal. On the other hand, the karyological changes also seem to have played a role in *Dugesia* diversification, in some cases leading to sympatric speciation (e.g. *D. hepta* and *D. benazzii* on Sardinia). Interestingly, when we look at its old origin and the fact that it has kept and homogeneous morphology among species until the present, it seems that the selective pressure on these creatures is weak (or the other way around). Therefore, the observed big number of *Dugesia* species would be essentially a consequence of its antiquity and wide distribution.

The diversity of *Dugesia* may actually be much larger than that already known. Nonetheless, the description of new species is a hard process, from the histological sections preparation to the species delineation and description by the researcher. Such difficulties probably hinder an increased rate of species erection. On the other hand, many *Dugesia* populations are fissiparous, this means that they do not have a copulatory apparatus on which base a formal description. This fact prevents the species identification of these specimens. Fortunately, it has been shown that the molecular data is a good tool to facilitate and direct the species description and partially overcome such problematics. In the present thesis we have carried out an integrative approach for *Dugesia* species delineation, using both morphological description and a molecular-based method for species delimitation, the General Mixed Yule-Coalescent method (GMYC) (Chapter 3). The results of this work have resulted in four new described species and 12 candidate species to be described. Unfortunately, many of these undescribed candidate species were asexual populations. In laboratory conditions it is possible to eventually resexualize fissiparous populations. The results from methods such as the GMYC would point which *Dugesia* specimens would be more likely to be new species and worthy to keep and care of in order to induce them to resexualize and to carry out morphological analyses. According to our results the molecular-based

delimitation methods are very suitable and reliable approaches for *Dugesia* species delimitation, pointing to new species from asexual specimens, helping in distinguishing putative cryptic species and providing more robustness in the delimitation of already valid species.

In the same integrative work we also described two species of a whole new genus, *Recurva*. Although we did not carry out any molecular-based species delimitation analysis, we obtained a phylogenetic tree, placing the new genus from Greece as the sister group of the genus *Schmidtea*. The description of six new species from the Greek area (2 *Recurva*, 4 *Dugesia*), plus 12 candidate species, plus the 9 already known, make of this region a Mediterranean hotspot of dugesiid biodiversity. However, this could be a bias because of the intensive research on this region. As mentioned before, the focus in the future on more different areas across the *Dugesia* distribution range will probably lead to an increased number of freshwater flatworms species.

An overview on the general results of the present thesis indicates the helpfulness and convenience of using molecular data in biogeographical and systematic research on freshwater flatworms. Actually, the use of this kind of data source may be inseparable of any evolutionary or diversity study on triclads. On the other hand, the use of more molecular markers such as whole mitochondrial genomes (Chapter 4) or nuclear sequences obtained from next-generation sequencing technologies is promising and necessary in order to get more accurate information from molecular data and more reliable results in freshwater flatworms evolutionary research.

Through the following sections I will summarize and extract issues coming from the study of *Dugesia* and other platyhelminthes during the development of this thesis, as well as concerns and future perspectives for the research on this gliding and wonderful group of animals.

1.1 The research on freshwater flatworms and the importance of phylogenetics, systematics and species description

Freshwater flatworms are still a relatively understudied group, leaving aside the intensive and competitive research on their amazing regeneration capabilities. The model organisms in this field *Schmidtea mediterranea* and *Dugesia japonica* both

gather a big number of research works in comparison with other species (365 and 570 results respectively in ISI web of Knowledge; search by 'Topic'; date 05/08/2014), along *Girardia tigrina*, also used in a wide diversity of experimental studies (927 results; also counting the results for *Dugesia tigrina*). As already mentioned in the first part of this general discussion, during the development of the present thesis, we accidentally found and described what may be the closest genus to *Schmidtea*, *Recurva* (Sluys *et al.*, 2013). Up to the start of this thesis it was thought to be *Dugesia* (Álvarez-Presas *et al*., 2008). This discovery could perhaps lead in the future to further interesting comparative regeneration studies along dugesiids in the light of their known phylogenetic relationships.

The model species *Dugesia japonica* according to our results could be a big conundrum (Chapter 2). This species described from China (including Hong Kong), Japan, Korea, Taiwan, and a part of Primorsky in the Russian Far East (Kawakatsu and Mitchell, 2004 and references therein) is now deciphered as different deep lineages. *D. japonica* is probably an old morphologically static group although further morphological analyses would be interesting. This fact must be taken into consideration because different research groups (e.g. in regeneration) are working on different Asian *Dugesia*, often considering all them to be the same species or lineage. Some of these research teams do not carry out any kind of species identification analysis after collecting the animals (e.g. Sakai and Sakaizumi, 2012; Yuan *et al.*, 2014) or they assign the lab strain just on the basis of their karyotype (i.e. haploid number n = 8). Accuracy would be desired for proper comparison of the results of different research teams.

On the other hand, other studies do not pay attention to the species they are working with, being very common to find papers dealing with *Girardia tigrina* but calling it *Dugesia tigrina* (e.g. Prados *et al*., 2013; Ramakrishnan *et al*., 2014). This synonymia was stablished more than 20 years ago (De Vries and Sluys, 1991).

It is convenient to describe as much accurately as possible the organisms the different research groups are working with in order to avoid confusion and make results and observations more comparable and useful, especially for people working in other fields. Such species or lineage identification could be done by regular collaboration with systematists or phylogenetists.

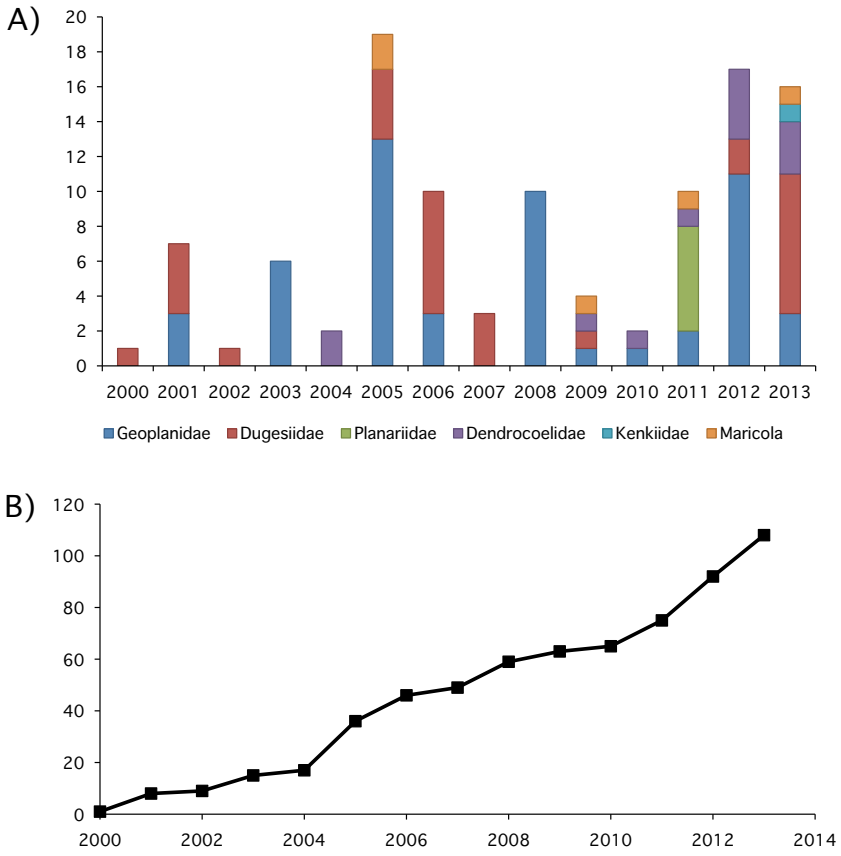


Fig. 1.1 A) Number of triclad species described per year since the 2000. Grouped according the Continenticola families plus Maricola, B) Total increase of new triclad species described since 2000. Up to 108 species.

Despite the great interest of generating the mentioned regenerative scientific information, along other kind of experimental researches (e.g. toxicological or even drug addiction), the evolutionary and systematic knowledge about freshwater triclads is still growing at a relative slow rate. Still, during the last decade a considerable number of freshwater triclads have been described (Fig. 1.1). However, there is a certain foreseeing of a much bigger diversity than that already known (Carbayo and Froehlich, 2008; Sluys *et al*., 2013). Therefore, there is a need to accelerate such evolutionary and systematic knowledge (Wilson, 2003). Hopefully, the advent and attractiveness of the use of new data (e.g. molecular data or ecological niche modelling) in an integrative taxonomic framework will help to achieve this major challenge. On the other hand, apart from the description of new species and regenerative studies, new biological knowledge about these animals (e.g. ecological, behavioural) may help to get a better

picture of their natural history. Such new knowledge would be desirable and helpful for anyone working on triclads.

1.2 The urgency to describe new species; the biodiversity crisis

As far as I have spent lots of hours dedicated to freshwater flatworms I have become concerned about their preservation status, more difficult to assess in comparison with other bigger and/or more attractive organisms because of their inconspicuous nature. Many different papers dealing somehow with the assessment of the biodiversity recurrently start their lines remembering the urgent need to describe and catalogue as much organisms as possible before they go extinct forever, swept away by the present biodiversity crisis (Wheeler *et al*., 2004). Indeed, this issue is serious, even considering freshwater flatworms.

During a sampling trip to Greece, we collected specimens from the type locality of *Dugesia elegans* De Vries, 1984 in Petaloudes Park, a retreat and touristic place on the island of Rhodes. As far as it is known, it is the only locality where this species can be found. Once in the lab, we analyzed 15 individuals resulting in 13 belonging to *D. sicula*, and two to *D. elegans*, the local species. *D. sicula* is a species presenting a big number of fissiparous populations widespreaded along the Mediterranean basin. Because its wide distribution range and its low genetic variability, it has been proposed that *D. sicula* wide distribution along the Mediterranean could be a consquence of human activity (Lázaro and Riutort, 2013). Considering the ratio 13/2 (introduced/local species) it seems probable that the *D. sicula* fissiparous populations are able to out-compete local species, ultimately reducing their population numbers. This situation could be likely the same in any place were *D. sicula* settles.

There are more evidences of threatened freshwater triclad species. *Schmidtea mediterranea*, the regeneration model organism, is recorded from few scattered populations very restricted in space (Lázaro *et al*., 2011). It is probable that former populations have been diminished by habitat destruction (e.g. urbanization), specially in the Mediterranean coast. In the same region other freshwater creatures such as the Spanish toothcarp (*Fartet* in catalan) and the Valencia toothcarp (*Samaruc* in catalan)

are now cataloged as endangered or critically endangered. On the other hand, it has also been proposed that *S. mediterranea* have been out-competed by other triclad species (Lázaro and Riutort, 2013).

According to the IUCN Red List database only one species of freshwater flatworm is extinct, *Romankenkius pedderensis* BALL, 1974B, which in fact is not extinct (Grant *et al*., 2006; Forteath *et al*., 2012). This database does not include any flatworm species as threatened, but the heads of the Red List admit it to be biased to terrestrial organisms, specially to vertebrates. The main reason of this lack of indexed species in these database is probably the extreme dificulty to assess the conservation status of flatworm species. However, it is known that many Mediterranean species live in restricted habitat extensions, which makes them more fragile and vulnerable to extinction events. Although we do not have any certainty about such putative threatened species, it is extremely likely that the extinction of many freshwater flatworms species has gone unnoticed. Therefore, there is an actual urgency to identify and catalogue as many planarians species as possible in order to take profit of the only chance we have to get the most accurate picture of their evolutionary history. Every single species or population disappearing is a forever lost valuable piece to untangle the evolutionary tree of these creatures.

1.3 The antiquity of the genus *Dugesia* and the Platyhelminthes

The observation of the wide distribution range of the genus *Dugesia* and the consideration of its limited vagility may suggest an old origin of the genus. These facts have lead some planariologists to think and propose the origin of this genus to have taken place on the ancient Gondwana superterrane (Ball, 1974a; 1975; Sluys *et al*., 1998), during the Mesozoic Era. In the present thesis we include a study dealing with the origin of *Dugesia* that did not found much support for the previous hypothesis but for an older origin, on the supercontinent Pangaea (Chapter 2). Probably, the divergence of *Dugesia* from its sister genera *Schmidtea* and *Recurva* occurred about 255 Mya. These antiquity is striking, specially considering how little these animals have changed morphologically along this time. Additionally, a study on the phylogeography of another dugesiid species *Schmidtea mediterranea* proposed an age for this species of

20−4 Mya and the origin of the genus back to 40 Mya. This would support an idea of general morphological stasis in planarians, as also proposed for the genus *Girardia* by Sluys (2007). Interestingly, these cases lead us to wonder how old the family Dugesiidae, the Tricladida order or the whole Platyhelminthes may be. Unfortunately, the fossil record has only provided us with a few scattered clues on the antiquity of the phylum. Therefore, a part from rare new fossils of platyhelminths found in the future, the only way to infer the origin of the different groups within the phylum will heavily rely on the historical biogeography; the correlation of historical events with the observed diversity and distribution of the present species. Hopefuly, new methods and data will increase in the following decades, allowing much accurancy in the estimations of the Platyhelminthes lineages antiquity.

1.4 The limitations of *Dugesia* morphology based studies

1.4.1 Limitations in species delimitation

Freshwater flatworms have been and are still described on the basis of their inner morphology. The special conditions in which they have to be preserved in order to be analyzed, the long and careful required sectioning, and the copulatory apparatus reconstruction from all the prepared slides make the formal description of *Dugesia* species a time consuming and painstaking process. Furthermore, it is not unusual for any of the two first steps (specimen fixation and sectioning) to fail because different unfortunate reasons, leading to the impossiblity to describe the prepared specimens. However, it is a necessary process due to the generalized idea of the need of different morphological diagnostic characters to describe new species and to identify already described ones.

Once morphological sections are sucessfully obtained new difficulties can arise. It is not uncommon to find very similar copulatory apparatus among the *Dugesia* species, making their identification even more difficult. Such limitations in available morphological characters may lead to the description of new species on the basis of just one morphological feature. For instance, *Dugesia astrocheta* MARCUS, 1953 is a species very similar to *D. sicula* (De Vries, 1988a; Sluys, 2007). They are differentiated just because the former lacks transverse muscles between the ventral nerve cords and the

ventral part of the gut. A second example involves the species *D. nansheae* De Vries, 1988 and *D. afromontana* that are split based on their body coloration (Stocchino *et al*., 2012). Eventually, this morphological similarities in the genus *Dugesia* could lead to oversplit species when intraspecific variation or preservation artifacts are considered or, on the opposite case, to be unable to detect very similar species (cryptic species).

One example of an overlooked species during a first check includes a species described in the present thesis, *Dugesia parasagitta* Sluys & Solà, 2013. When analyzing samples from the Greek island Corfu, a first check of the morphology pointed all collected populations to be the same species, *D. sagitta* (Schmidt, 1861). Later, we carried out the molecular species delimitation method GMYC and three different entities out from the *D. sagitta* populations were identified as putative species. Finally, a double-check on the morphology of the Corfu animals allowed us to erect a new species on the basis of characters that went unnoticed when analyzed for the first time.

On the other hand, the morphological oversplit of *Dugesia* species have also been proposed for *D. maghrebiana* Stocchino, Manconi, Corso, Sluys, Casu & Pala, 2009 and *D. sicula* (Lázaro, 2012) on the basis of their similarity in a molecular phylogenetic tree. However, morphological characters were considered to be enough for the independent and formal description of the former species. Although still under study and not presented in this thesis, we also found molecular evidences for the synonymia of *D. biblica* Benazzi & Banchetti, 1972 with *D. sicula*. In the following section we also propose another case of putative species oversplit between *D. arabica* Harrath & Sluys, 2013 and *D. aethiopica* Stocchino, Corso, Manconi & Pala, 2002 according to the molecular data.

Indeed, these situations may be explained by a rapid and recent accumulation of new and different morphological states. Both cases would imply that the few characters of the copulatory apparatus are able to change at relatively fast rates. Nonetheless, we consider these examples as very interesting and proper to be tested in further research as molecular work will ultimately help to overcome this under or oversplit problems due to a possible certain degree of morphological plasticity.

1.4.1.1 A preliminary example of morphological and molecular disagreement in species delimitation: *D. aethiopica* and *D. arabica*

In this section I present a preliminary approach concerning the discordance between molecules and morphology in two *Dugesia* species. During the development of the

present thesis we have received samples of *Dugesia* from different collaborators. One of these *Dugesia* packages contained samples from many localities in Western Yemen from which *Dugesia arabica* HARRATH & SLUYS, 2013 was described.

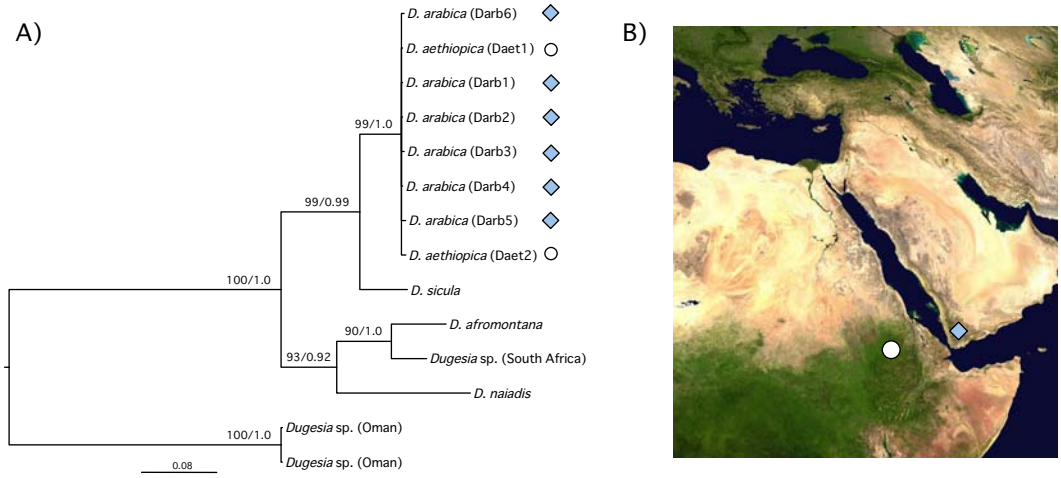


Fig. 1.2 A) Phylogenetic tree based on the mitochondrial gene Cox1 including specimens of *D. arabica* and *D. aethiopica* along its sister species. Topology of MrBayes 3.2. Values on the branches are showing the support of Bootstrap (>75)/Posterior Probability (>0.90). B) Localities of *D. aethiopica* and *D. arabica*. White circle: *D. aethiopica* type locality. Blue rhombus: Region of various localities of *D. arabica*. Photography: World Wind (NASA) (Public Domain).

The Cox1 sequence of these samples rapidly showed an striking molecular resemblance of those specimens to *D. aethiopica* specimens from the Lake Tana, Ethiopia (Fig. 5.2B). Interestingly, populations from Dhamar, in Yemen, were assigned to *D. aethiopica* in a previous paper (Sluys, 2007). In that paper, it was already pointed that these populations could be a different species because of the absence of a cavity in the parenchyme of the penis papilla. Finally, the Arabian populations were described as a different species, *D. arabica* (Harrath *et al*., 2013). The differences between *D. aethiopica* and *D. arabica* are the absence of such parenchymatic cavity in the penis papilla, the asymmetrical openings of the oviducts, and the absence of a subepithelial longitudinal muscle layer on the bursal canal.

In order to show this unclear molecular distinction between both species we obtained a phylogenetic tree including one specimen from each of six different sampling localities from Yemen included in the paper of Harrath and collaborators (2013) and two specimens of *D. aethiopica* (Annexes 1 − Table 3) (Fig. 1.2A).

As a preliminary conclusion the very little genetic distance between *D. aethiopica* and *D. arabica* specimens points them to be the same species. However, it is rather problematic to decide whether the described differences in morphology correspond to intraespecific or interespecific differences. As morphology and molecules do not evolve at the same rate, it is possible that any of both changes 'faster' than the other. If *D. aethiopica* and *D. arabica* are the same species, the morphological differences between them could be due to preservation artifacts or to a extreme plasticity of the described diagnostic characters. On the other hand, if such differences are a consequence of a speciation process, the molecular mix with their respectively sister species could be due to incomplete lineage sorting. Therefore, this is an interesting group to be analyzed with a multilocus approach using many different loci (both mitochondrial and nuclear) in a coalescent framework. We did not carry out proper molecular-based species delimitation analyses to test whether they could actually be considered different evolving lineages. Therefore, we here adopt a conservative approach in the light of our limited data and methods and suggest to keep them as separated species until further evidences are obtained.

1.4.1.2 Perspectives in *Dugesia* species delimitation: beyond morphology and molecules

As mentioned before, the gathering of evidences such as the dating analyses and the general morphological homogeneity of the genus have pointed *Dugesia* as a morphological static creature. It has kept the same general external and inner appearance for millions of years and the inner morphology (e.g. the copulatory apparatus) has hardly change across the different species (Chapter 2). The apparently non-adaptative nature of the different morphological characters suggests a stochastic accumulation of phenotypic differences, having little to do with selective pressure on these traits. Thus, the isolation of populations in different isolated freshwater bodies may be responsible of the fixation of random trait changes of their inner morphology.

The morphological description of freshwater flatworms has welcomed the incorporation of molecular data to help in the proposal of candidate new species, in revealing cryptic species and supporting those already known. However, the use of more data sources beyond the morphology and the molecular-based delimitation

methods that may help to delinate and identify species with more robustness would be of major interest.

As discussed in the following section comparing the morphological characters used by Sluys and collaborators (1998) in a phylogenetic reconstruction with a molecular-based phylogenetic tree, it seems that the karyology may be informative in delineating some groups or species. Most of the *Dugesia* species present a karyotype of n = 8, but there are some exceptions with a chromosmal haploid number of n = 7, that has been found to be a synapomorphic condition for *D. batuensis* and *D. ryukyuensis* and a differential trait of *D. hepta* from its sister species *D. benazzii*. Furthermore, another molecularly defined lineage contains many species that share an haploid number of n = 9. Therefore, more effort on karyological analyses would be of great help as they support the membership of a species to a certain clade or its differentiation from its sister species or group.

The ecological factors have never been taken into serious consideration for species delimitation in freshwater flatworms. *Dugesia* species seem to be very generalist as this would explain its wide distribution range and its apparent homogeneous nature. Environmental factors such as the temperature seems to overlap across species as it has been exposed in the Ecology section of the Introduction. Thus, it seems that *Dugesia* species can live under a wide range of water temperatures. Still, the absence of representatives on the polar and subarctic regions on Eurasia clearly suggests a null tolerance of the whole genus to very cold waters. However, it would be interesting to try to measure or register environmental conditions such as chemical composition of the water, the ground, type of rocks, surrounding flora, or the mean strength of water current, among others. These data could be analyzed and correlated statistically with the different *Dugesia* lineages under study. It would be also necessary to be sure that such correlation results are not due to circumstancial conditions.

A different but more work-demanding strategy in the attempt to plot limits between species could imply keeping different *Dugesia* populations suspected to be different species on the bases of geographical or genetic data and try to breed them in the laboratory. Thus, adopting a biological concept based approach. This kind of experiments have been already carried out for Proseriata free-living flatworms (Curini-Galleti *et al*., 2012). Nonetheless, this would only be possible when dealing with sexual reproducing populations, which many times is not a common situation. Although interesting and arguably accurate, this method may be not the best when there is a need

to delimit species at an increased rate. On the other hand, one critizism to this approach may argue that separated species in nature (e.g. distant drainage basins) may interbreed in lab conditions producing hybrids but this situation would be very unlikely in the wild. Thus, the species criteria adopted by the researcher would be pivotal in the rise of a new species as it should be decided which importance or sense have this approach and its results.

In summary, up-to-date molecular and morphological data seem to be the best available sources to delimit *Dugesia* species. Although it is generally accepted that the most important data source to describe species is the morphology, other data could be considered in combination with morphology or other datatypes or only by themselves to delimit freshwater flatworm species. However, as much as different methodologies and different data are used, species delimitation will always relay on a certain degree of subjectiviness. Nonetheless, every approach or methodology searching for species limits on the basis of different data (e.g. genetics or ecology) is helping to find the best approximate answers.

1.4.2 Limitations in phylogenetics

There are few characters available to be used in morphology-based phylogenies of *Dugesia* (15 in Sluys *et al*., 1998). Unfortunately, they were not enough to solve the phylogeny of the genus apart from large monophyletic species groups (Fig. 1.6). The lack of more morphological characters probably prevented the delimitation of smaller clades. On the other hand, some of these features seem to show some plasticity as they are found in far-related species. This problem may arise from the morphologically uniform nature of the genus.

We obtained a phylogenetic tree that only includes those morphology-based described species for which we have molecular data available (Fig. 5.3). The aim of this analysis is to map the states of those *Dugesia* characters used by Sluys and collaborators (1998) and two new features in comparison with the molecular tree. The phylogenetic tree was obtained by a Bayesian approach using the software BEAST v 1.7.3 with four genes (Cox1, 18S, 28S, ITS−1). Nonetheless, we must consider that the results may be biased because this analysis only encompasses one third of the total number of known *Dugesia* species.

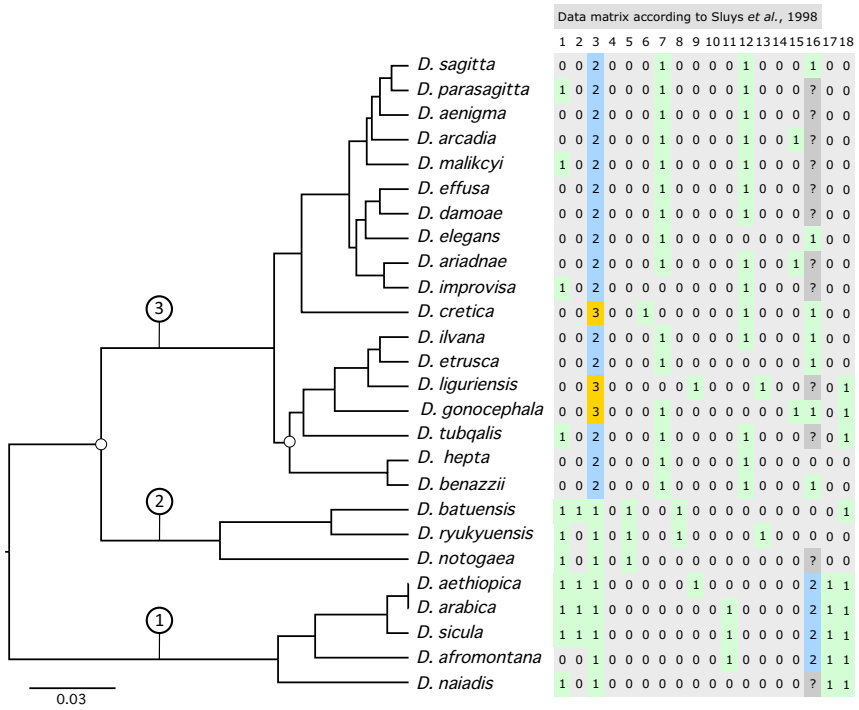


Fig. 1.3 Phylogenetic tree including 26 *Dugesia* species based on a concatenated dataset including Cox1, 18S, 28S and ITS−1 genes compared with a morphological data matrix based on that of Sluys *et al.*, 1998 with two extra characters. The numbers in the circles indicate the main phylogenetic clades. White circles on the nodes show posterior probablity supports under 0.95. Character description, 1−16 from Sluys *et al.*, 1998; 17−18 for the present discussion: **1** Ejaculatory duct centrally (0), ventrally (1) located in penial papilla; **2** Opening of ejaculatory duct terminal (0), subterminal (1); **3** Diaphragm absent (0), small (1), large (2), pointed (3); **4** Double diaphragm absent (0), present (1); **5** Duct between seminal vesicle and diaphragm absent (0), present (1); **6** Adenodactyls sensu stricto absent (0), present (1); **7** Penial folds absent (0), present (1); **8** Penial valve absent (0), asymmetrical (1), symmetrical (2); **9** Glandular parenchymatic zones in penial papilla absent (0), present (1); **10** Nipple on penial papilla absent (0), present (1); **11** Entrance of oviducts into bursal canal symmetrical (0), asymmetrical (1), common oviduct (2); **12** Openings vasa deferentia into seminal vesicle not close to diaphragm, i.e., in anterior section of seminal vesicle (0), close to diaphragm, i.e., in posterior section of vesicle (1); **13** Musculo−glandular area of swelling in atrial wall absent (0), area (1), swelling (2); **14** Glandular vestibulum absent (0), present (1); **15** Outer pharynx musculature normal (0), with extra third longitudinal layer (1); **16** Haploid number of chromosmes: n = 7 (0), n = 8 (1), n = 9 (2); **17** Pleated and/or folded bursal canal absent (0); present (1); **18** Seminal vesicle shape 'rounded' (0), elongated (1).

Looking at the results it seems that most of the characters are not giving much phylogenetic information. However, some of them are pointing big groups:

(i) The character state 'large diaphragm' (character 3) is shared by the species within the clade 3, with the exception of *D. cretica* that has developed a pointed diaphragm. The 'small diaphragm' would be paraphyletic but shared by the groups 1 and

2. The paraphyletic condition of this state suggests the plesiomorphic condition of this state.

(ii) The new character considered here 'Pleated and/or folded bursal canal' (character 17) probably is a synapomophy of group 1.

(iii) Finally, the only non−morphological character also gives some phylogenetic information, the karyotypes. The haploid number n = 9 (character 16) would correspond to a monophyletic clade of species (group 1). On the other hand, the haploid number n = 7 appears twice along the tree, shared by two sister species (*D. batuensis* and *D. ryukyuensis*) and far apart in *D. hepta*.

As already noted by Sluys and collaborators (1998) for the morphological data, the comparison between the molecular tree and the morphological dataset points a few characters to be able to distinguish only between big phylogenetic groups. However, such comparison is interesting as it shows which characters are more plastic and tend to evolve more independently and which are more reliable to give phylogenetic information and clues about the species affinities.

1.5 Limitations and perspectives in evolutionary research on planarians

There are plenty of questions about the natural history of freshwater flatworms that are still waiting to be properly answered. Although during the present thesis we worked to find answers to some of these questions, there is still a large amount of investigation to be done. Yet, the evolutionary research on freshwater flatworms have to take into account some limitations or the lack of certain knowledge about the biology of these animals.

The point is, how uncertainties on biological aspects of freshwater planarians and lack of enough data can impact the evolutionary research on these creatures. In this section I present four different issues that may concern any planariologist working on historical biogeography and systematics of freshwater flatworms using molecular data: the dispersal capabilities, the putative differences in subtitution rates of *Dugesia* depending on their reproduction type, the intrinsec problems of historical biogeography, and the limited set of molecular markers for planarians.

1.5.1 Dispersal capabilities

There is a basic knowledge about freshwater flatworms, however in some aspects is still quite poor. For instance, most of the ecological studies on some of these animals were done before the 90's, and very few updates have been carried out to the present. It is widely assumed that these are fragile invertebrates of low vagility (Ball, 1974a). However, this low dispersal capability may not be as limited as we previously thought. We consider that further studies on their dispersion capabilities would be of great help to contextualize better biogeographical and phylogeographical approaches of these animals. Such studies could include extensive population analyses on adjacent basins with a well-known hydrological history or extensive studies within large basins (e.g. Nile or Amazon). On the other hand, *in situ* or *in vitro* resistant related research may be carried out in order to analyze how they react in front of different degrees of dissecation or salinity. It would also be of great interest to find out how and how often they could manage to move in phreatic waters, and to explore if they are able to move through water-saturated sediments. New approaches on dispersive abilities would be of major help for biogeographical and phylogeographical studies with these animals.

Although we temptatively assume the low dispersal capability of *Dugesia*, we now consider as not discardable even the dispersal in small pools of freshwater in rafts, eventually refilled with rainwater. Situations like that would allow the overseas dispersion. However, we still consider this possibility as extremely unlikely.

On the other hand, in freshwater flatworms biogeography we also have to take into account the human-mediated dispersal that can ultimately complicate its inference. This way of dispersal has been found for different freshwater flatworm species (e.g. Ball, 1969b; Lázaro *et al*., 2011).

1.5.2 Substitution rates

Dugesia probably shifts from one mode of reproduction to another in the wild, even changing its karyotypes from triplois to diploids and viceversa. In *Dugesia* triploidy is associated with a mainly-fissiparous way of reproduction, while the diploidy is essentially related to sexual reproduction by cross-fertilization. Such reproductive shift involving changes in the ploidies have been recently observed in laboratory conditions

in *D. ryukyuensis* (Chinone *et al*., 2014) but it is still unknown if it occurs in wild populations. However, it seems likely that *Dugesia* would be able to do this in nature. To this lack of knowledge, we have to add another uncertainty. We do not know for how long asexual populations are reproducing asexually and what would trigger the sexual shift, if it is actually done (Kobayashi *et al*., 2012). It could be from days to many years, and very different among lineages or populations. For instance, in laboratory conditions, fissiparous strains of *Dugesia aethiopica*, *D. sicula* and *D. afromontana* have developed a copulatory apparatus after 7−8 months, 1 year and 2 years respectively. However, the development of the copulatory apparatus in most of the cases did not imply the adoption of a putative sexual reproductive mode (Stocchino and Manconi, 2013). In the wild, many populations have also been found to be ex-fissiparous (i.e. hyperplasic ovaries and an increased body size), such as some *D. sicula* populations (Pala *et al*., 1995; Lázaro and Riutort, 2013). However, we neither know if this new adquisition of a copulatory apparatus from an asexual individual would eventually imply the natural shift to a succesful sexual reproduction (Benazzi and Ball, 1972; Grasso and Benazzi, 1973; Benazzi and Deri, 1980). Interestingly, those *D. ryukyuensis* shifting from fissiparous triploids to truth sexual reproduction in laboratory conditions had been fed with the sexual dendrocoelid *Bdellocephala brunnea* IJIMA & KABURAKI, 1916 for three weeks until they developed the copulatory apparatus (Chinone *et al*., 2014).

Asexual strains have been long considered an evolutionary *cul-de-sac* because they are supposed to accumulate deleterious mutations (Kondrashov, 1988). However, the apparent reproductive and karyological plasticity of *Dugesia* would allow the genus to skip this fate. The long term asexual populations would survive thanks to the neoblasts, the pluripotent population of stem cells that are responsible of providing all specialized cells in *Dugesia* body. It is reasonable to assume that a neoblast accumulating deleterous mutations will die and will be replaced by a 'healthy' neoblast, avoiding the deadly effect of deleterious mutations. The replacement ability of a neoblast is well-known, it has been demonstrated a single of them to be capable of replacing all the cells of a whole planarian body (Wagner *et al*., 2011).

The phylogenetic trees in the present thesis contain both fissiparous and sexual-reproducing specimens. We do not know how mutations are fixed in asexual specimens and inherited by their offsprings, either in asexual (fissiparous or parthenogenetic) or resexualized populations. The possibility of different substitution rates depending on the

reproductive mode must be explored, as it could be widely different along branches of phylogenetic trees containing sexual and asexual populations of *Dugesia*. This research would be important because these disparate substitution rates could not fit even relaxed molecular clocks when tree dating analyses are carried out (Lázaro *et al*., 2011).

In summary, it is necessary to know how this putative fixation of nucleotidic changes would impact on *Dugesia* substitution rates. This could be explored by simulations or by empirical tests, and if significant trying to develop a specific substitution model for implementation in freshwater flatworms phylogenetic and divergence time estimation analyses. An interesting putative subject for this study may be the species *D. arabica* in Yemen. This species is known to present different natural asexual (triploid) and sexual (diploid) reproducing populations in the wild, as well as mixoploid (diploid and triploid) populations that reproduce both sexually and asexually (by fissiparity and parthenogenesis).

1.5.3 Biogeographical uncertainties

Freshwater flatworms are considered fragile organisms that cannot disperse overland or through air, and with no resistant or dispersal phases in their life-cycle (Reynoldson, 1966). Because of this, they have been considered good indicators of paleogeographical relationships. However, when dealing with historical biogeography inference of flatworms a major inconvenient appears, the so-mentioned lack of fossils. Such non-existent fossil record limits the reconstruction of the flatworm tree of life to be exclusively based on the extant species, with no additional information from the strata. On the other hand, it forces historical biogeography to make use of the paleogeographical events to calibrate phylogenetic trees with little alternative choice.

Both undersampled or phylogenies with no information on extinct lineages and the calibration of phylogenetic trees using paleogeographical events imply associated limitations in historical biogeography.

First, phylogenies can now be inferred with an important degree of confidence, failing when, for instance, biological events such as evolutive radiations have taken place. On the other hand, phylogenies are unavoidable and instrinsicaly excluding extinct or not sampled lineages. These unrecorded and missed representatives in the tree could have been distributed in different areas giving reliable and important information about the lineage history. Therefore, historical biogeography exclusively inferred from

contemporary taxa may lead to wrong interpretations (Lieberman, 2002; Quental and Marshall, 2010; Crisp *et al*., 2011).

Second, the use of paleogeographical events to calibrate phylogenetic trees may seem proper for poor dispersal species, such as freshwater fauna sensitive to marine environments, but it is still difficult to know with certainty if the split of two landmasses happened at the same time of the divergence of a species on them. On the other hand, calibration using geological events could be prone to circular arguments, specially if no external calibration to the point of interest or substitution rates of other creatures are used (Datta-Roy and Kranth, 2009; Kondandaramaiah, 2011). Unfortunately, the geological-based calibrations are the best approaches available for such complicated creatures.

Another point to take into consideration in historical biogeographical approaches is the uncertainty that is also present in the geological field regarding the timing and succession of some paleogeographical events. Some of these events processes and timings have been challenged and investigated recently, being changed in the last years. Here I present three different examples of such still-changing paleogeographical hypotesis.

The collision of India with Eurasia has been thought to be direct to the Tibetan area to the Paleocene/Early Eocene about 50−55 Mya (Patriat and Achache, 1984; Zhu *et al*., 2004; Leech *et al*., 2005). However, in 2007 a quite different new proposal of the track of the subcontinent migrating northwards has been proposed, it first collided with an intraoceanic island arc over 55 Mya and later it collided with the Tibetan area over 35 Mya (Abrajevitch *et al*., 2005; Ali and Aitchison, 2006; Aitchison *et al*., 2007). Even this new proposal has two alternatives, it collided directly with Eurasia or it migrated very close to Southeast Asia around 55 Mya, allowing fauna interchange, and being its final impact with the Tibetan area (Ali and Aitchison, 2008). This second event has already been supported by historical biogeographical inferences (Klaus *et al.*, 2010). However, the debate on how and when India collided with Asia is still ongoing (Van Hinsbergen *et al*., 2012a,b; Aitchison and Ali, 2012; Ali and Aitchison, 2014). A second example involves the proposal of the Kerguelen Plateau as a putative land bridge that connected India and Antartica until as late as c. 80 Mya. It has been postulated that it was used as a pathway by different fauna until then (Hay *et al*., 1999). However, this connection has been challenged recently (Ali and Aitchison, 2008; 2009). Most of this plateau in the mid-Late Cretaceous was submerged and the terminations of the terrain

were separated from India or Antarctica. A second connection from the Late Cretaceous has also been proposed between Southern Madagascar and Eastern Antartica's Riiser-Larsen Peninsula, the Gunnerus Ridge (Case, 2002; e.g. Yoder and Nowak, 2006; Prasad and Sahni, 2009). Although it is included in many discussions of different works (e.g. Tierney *et al*., 2008; Upchurch, 2008), this ancient pathway has also been challenged recently, considered an untenable bioconnection (Ali and Krause, 2011).

Therefore, datings or phylogenies with no historical explanation or hardly fitting ones will hopefully find a better future historical explanation thanks to the ever-improving geological knowledge. Other studies would be simply based on misassumptions or highly impacted by the undersampling and/or extinction bias. On the other hand, future paleontological descoveries will for sure shed more light on the historical biogeography of clades that fossilize, making their reconstruction and divergence time estimations more accurate.

1.5.4 Availability of molecular markers

One of the main limitations during the development of the present thesis has been the lack of effective primers to amplify successfully different regions of the planarian genomes. This lack of available molecular markers for the Lophotrocozoa has been a general inconvenient for a while. In the different works presented here, we used in total up to four markers, one mitochondrial (Cox1) and three nuclears (18S, 28S, ITS−1). We have tried to obtain more markers by sequencing the complete mitochondrial genome of different triclads. Unfortunately we did not succeed in the obtention of a *Dugesia* species mitogenome. However, we obtained and annotated two new complete mitochondrial genomes (*Crenobia alpina* and *Obama* sp.) that will be of great help in the development of primers to amplify regions of these molecules or to perform whole genome comparisons, specially useful in intraspecific studies. Although we did not apply the advent of next-generation high throughput sequencing to obtain nuclear markers for the present thesis, future works on freshwater flatworms are now using such advances to carry out more reliable evolutionary analyses of flatworms.

1.6 General perspectives

Although some of the sections in the present general discussion may seem to discourage researchers to work on systematics and historical biogeography of freshwater flatworms, the idea is essentially the opposite. Indeed, it is necessary to carry out these kind of approaches dealing with other freshwater flatworms to compare the results obtaining more reliable interpretations about their evolutionary history. In this way, we will be able to get a more accurate picture of these fascinating creatures.

Planarians do not fossilize and they present confusing and astonishing karyological and reproductive features. However, these problematics, among others, are what actually make them more challenging and therefore interesting. With the present work we have been able to improve or update previous hypotheses on *Dugesia* historical biogeography in the light of new data not available at that time. On the other hand, we have been able to apply new species delimitation methodologies that have facilitated the description of *Dugesia* species. In the future, more methodologies, datasources, and the unstoppable everincreasing general knowledge (e.g. paleogeographical) will allow to overcome (at least partially) some of the mentioned problematics in the *Dugesia* evolutionary analyses and to challenge or support the hypotheses presented in this thesis.

Section V

Conclusions

Conclusions

1. The genus *Dugesia* presents a high number of species in the Greek area, as the number of formally described species in this region has been increased from 9 to 13. Additionally, a whole new genus with two new species and a putative third one has been described. Two confirmed and 10 unconfirmed candidate *Dugesia* species have also been proposed.

2. The high number of species known from Greece also suggests a much bigger number of *Dugesia* diversity to be described across the genus distribution range. This idea is also supported by the high number of deep lineages on Madagascar and the Far East.

3. The species delimitation method General Mixed Yule-Coalescent (GMYC) is a convenient and effective approach to carry out delimitation of putative species in the genus *Dugesia*. The results from this methodology together with morphological analysis facilitates the detection of unnoticed character states and give robustness to the species delineation. On the other hand, it is specially useful to suggest putative new species from asexual populations.

4. The most plausible explanation according to our results places the origin of *Dugesia* on the supercontinent Pangaea in the Late Triassic, instead of on Gondwana as previously proposed. The presence of the genus on Eurasia and Africa would be explained by an ancient wide distribution on Pangaea.

5. *Dugesia* seems to be an extreme example of long-term morphological stasis presenting an homogeneous inner morphology and a very similar external appearance among species across hundreds of millions of years.

6. The diversification of *Dugesia* may have been strongly shaped by vicariant events due to the drifting of the tectonic plates. However, within landmasses, the isolation, severing, and contact of freshwater bodies may have played a major role in its speciation.

7. The accidental human-mediated transportation of *Dugesia* specimens between geographically distant places may be a situation more common than previously thought as we detected some putative cases in both biogeographical works.

8. The free-living platyhelminthes of the Geoplanoidea suborder present a more relaxed selective pressure on their mitochondrial genomes in comparison with the parasitic platyhelminthes. This surprising result does not match the assumption of a more relaxed evolutionary pressure on parasites according to their life cycle.

Section VI

References

References

Abrajevitch A, Ali JR, Aitchison JC, Davis AM, Liu J, Ziabrev S (2005). **Neotethys and the India-Eurasia collision: insights from a palaeomagnetic study of the Dazhuqu ophiolite, southern Tibet.** *Earth Planet Sci Lett* **233(1−2)**:87−102.

Agata K, Soejima Y, Kato K, Kobayashi C, Umesono Y, Watanbe K (1998). **Structure of the planarian central nervous system (CNS) revealed by neuronal cell markers.** *Zoolog Sci* **15(3)**:433−440.

Agata K, Watanabe K (1999). **Molecular and cellular aspects of planarian regeneration.** *Semin Cell Dev Biol* **10(4)**:377−383.

Akaike H (1974). **A new look at the statistical model identification**. *IEEE Trans Automat Contr* **19(6)**:716−723.

Ali JR, Aitchison JC (2006). **Positioning Paleogene Eurasia problem: solution for 60−50 Ma and broader tectonic implications.** *Earth Planet Sci Lett* **251(1−2)**:148−155.

Ali JR, Aitchison JC (2008). **Gondwana to Asia: plate tectonics, paleogeography and the biological connectivity of the Indian sub-continent from the Middle Jurassic through end Eocene (166−35 Ma).** *Earth Sci Rev* **88**:145−166.

Ali JR, Aitchison JC (2009). **Kerguelen Plateau and the Late Cretaceous southern-continent bioconnection hypothesis: tales from a topographical ocean.** *J Biogeogr* **36(9)**:1778−1784.

Ali JR, Aitchison JC (2014). **Greater India's norhtern margin prior to its collision with Asia.** *Basin Res* **26(1)**:73−84.

Ali JR, Krause DW (2011). **Late Cretaceous bioconnections between Indo-Madagascar and Antarctica: refutation of the Gunnerus Ridge causeway hypothesis.** *J Biogeogr* **388(10)**:1855−1872.

Aitchison JC, Ali JR, Davis AM (2007). **When and where did India and Asia collide?** *J Geophys Res* **112(B05)**:423.

Aitchison JC, Ali JR (2012). **India-Asia collision timing.** *Proc Natl Acad Sci U S A* 2 **109(40)**:E2645.

Álvarez-Presas M, Baguñà J, Riutort M (2008). **Molecular phylogeny of land and freshwater planarians (Tricladida, Platyhelminthes): from freshwater to land and back.** *Mol Phylogenet Evol* **47(2)**:555−568.

Anderson JM, Johann JC (1958). **Some aspects of reproductive biology in the fresh-water triclad turbellarian *Cura foremanii.*** *Biol Bull* **115(3)**:375−383.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunder NC (1987). **Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics.** *Annu Rev Ecol Syst* **18**:489−522.

Avise JC (2000). **Phylogeography: The history and formation of species**. Harvard University Press, Cambridge, Massachusetts.

Avise JC (2009). **Phylogeography: retrospect and prospect.** *J Biogeogr* **36(1)**:3−15.

Baguñà J, Saló E, Auladell C (1989). **Regeneration and pattern formation in planarians. III. Evidence that neoblasts are totipotent stem cells and the source of blastema cells.** *Development* **107**:77−86.

Baguñà J (1998). **Planarians.** In: *Cellular and molecular basis of regeneration: From invertebrates to humans* (ed. Ferretti, P and Géraudie, J), pp. 135−165. John Wiley & Sons Ltd, New York.

Ball IR (1969a). **An annotated checklist of the freshwater tricladida of the Nearctic and Neotropical Regions.** *Can J Zool* **47(1)**:59−64.

Ball IR (1969b). ***Dugesia lugubris* (Tricladida: Paludicola), a European immigrant into North American fresh waters.** *J Fish Res Board Can* **26(2)**:221−228.

Ball IR (1970). **Freshwater Triclads from the Oriental Region.** *Zool J Linn Soc* **49(4)**:271−294.

Ball IR (1974a). **A contribution to the phylogeny and biogeography of the freshwater triclads (Platyhelminthes, Turbellaria).** In: *Biology of the Turbellaria* (ed. Riser, NW and Morse, PM), pp. 339−401. McGraw-Hill, New York.

Ball IR (1974b). **A new genus of freshwater triclad from Tasmania with reviews of the related genera *Cura* and *Neppia* (Turbellaria, Tricladida).** *Life Sci Contrib R Ont Mus* **99**:1−48.

Ball IR (1975). **Nature and formulation of biogeographical hypotheses.** *Sys Zool* **24(4)**:407−430.

Ball IR (1983). **Planarians, plurality and biogeographical explanations.** In: Evolution, time and space: the emergence of the biosphere (ed. Simms RW, Price JH and Whalley PES), pp. 409−430. Academic Press London, London.

Ball IR, Fernando CH (1970). **Freshwater Triclads (Platyhelminthes. Turbellaria) from Anticosti Island.** *Nat Can* **97**:331−336.

Ball IR, Reynsoldson TB (1981). **British planarians.** Cambridge University Press, Cambridge.

Barrett CF, Freudenstein JV (2011). **An integrative approach to delimiting species in a rare but widespread mycoheterotrophic orchid.** *Mol Ecol* **20(13)**:2771−2786.

Benazzi M (1952). **Sulla possibilità di reproduzione sessuata senza accoppiamento in Tricladi di acqua dolce.** *Atti Soc Tose Sci Nat Ser B* **59**:107−110.

Benazzi M, Baguñà J, Ballester R, Puccinelli I, Papa RD (1975). **Further contribution to the taxonomy of the *Dugesia lugubris-polychroa group* with description of *Dugesia mediterranea* n. sp. (Tricladida, Paludicola).** *Boll Zool* **42(1)**:81−89.

Benazzi M, Ball IR (1972). **The reproductive apparatus of sexual specimens from fissiparous populations of Fonticola morgani (Tricladida, Paludicola).** *Can J Zool* **50**:703−704.

Benazzi M, Banchetti R (1972). **Descrizione di *Dugesia biblica*, nuova microspecie del "gruppo *Dugesia gonocephala*" trobata nel fiume Giordano (Israele).** *Atti Soc Ital Sci Nat* **79**:83−91.

Benazzi M, Deri P (1980). **Histo-cytological study of ex-fissiparous planarian testicles (Tricladida Paludicola).** *Monit Zool Ital* **14**:151-163.

Benazzi M, Grasso M (1977). **Comparative research on the sexualisation of fissiparous planarians treated with substances contained in sexual planarians.** *Monit Zool Ital* **11(1−2)**:9−19.

Benazzi M, Gremigni V (1982). **Developmental biology of triclad turbellarians (Planaria).** In: *Developmental biology of freshwater invertebrates* (ed. Harrison, FW and Cowden, RR), pp. 152−211. Alan R. Liss, New York.

Benton MJ, Ayala FJ (2003). **Dating the tree of life.** *Science* **300(5626)**:1698−1700.

Best JB, Morita M, Noel J (1968). **Fine structure and function of planarian goblet cells.** *J Ultrastruct Res* **24**:385−397.

Beukeboom LW, Vrijenhoek RC (1998). **Evolutionary genetics and ecology of sperm-dependent parthenogenesis.** *J Evol Biol* **11(6)**:755−82.

Beukeboom LW, Weinzierl RP, Reed KM, Michiels NK (1996). **Distribution and origin of chromosomal races in the freshwater planarian *Dugesia polychroa* (Turbellaria: Tricladida).** *Hereditas* **124(1)**:7−15.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007). **Cryptic species as a window on diversity and conservation.** *Trends Ecol Evol* **22(3)**:148−155.

Bishop P (1995). **Drainage rearrangement by river capture, beheading and diversion.** *Progr Phys Geogr* **19(4)**:449−473.

Blair JE (2009). **Animals (Metazoa).** In: *The timetree of life* (ed. Hedges, SB and Kumar, S), pp. 223−230. Oxford University Press, Oxford.

Bond JE, Stockman AK (2008). **An integrative method for delimiting cohesion species: finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring.** *Syst Biol* **57(4)**:628−646.

Britten RJ (1986). **Rates of DNA sequence evolution differ between taxonomic groups.** *Science* **231(4744)**:1393−1398.

Bromham L, Phillips MJ, Penny D (1999). **Growing up with dinosaurs: Molecular dates and the mammalian radiation.** *Trends Ecol Evol* **14(3)**:113−118.

Bromham L (2003). **Molecular clocks and explosive radiations.** *J Mol Evol* **57(Suppl. l)**:S13−20.

Bromham L, Penny D (2003). **The modern molecular clock.** *Nat Rev Genet* **4(3)**:216−224.

Brøndsted HV (1969). **Planarian regeneration.** Pergamon Press, Oxford.

Brundin L (1966). **Transantarctic relationships and their significance as evidenced by chironomid midges with a monograph of the subfamilies Podonominae and Aphroteniinae and the austral Heptagynae.** *K Sven Vetenskapsakad Hand* **11**:1−472.

Calow P, Beveridge M, Sibly R (1979). **Heads and tails: adaptational aspects of asexual reproduction in freshwater triclads.** *Am Zool* **19(3)**:715−727.

Calow P, Read DA (1986). **Ontogenetic patterns and phylogenetic trends in freshwater flatworms (Tricladida): Constrains or selection?** *Hydrobiologia* **132(1)**:263−272.

Calow P, Woollhead AS (1977). **Locomotory strategies in freshwater triclads and their effect on the energetics of degrowth.** *Oecologia* **27(4)**:353−362.

Carbayo F, Froehlich ME (2008). **Estado do conhecimento dos macroturbelários (Platyhelminthes) do Brasil.** *Biota Neotrop* **8(4)**:77−197.

Cardona A, Hartenstein V, Romero R (2005). **The embryonic development of the triclad** *Schmidtea polychroa. Dev Genes Evol* **215(3)**:109−131.

Carpenter K, Morita M, Best J (1974). **Ultrastructure of the photoreceptor of the planarian *Dugesia dorotocephala*. I. Normal eye.** *Cell Tiss Res* **148(2)**:143−158.

Carranza S, Littlewood DTJ, Clough KA, Ruiz-Trillo I, Baguñà J, Riutort M (1998). **A robust molecular phylogeny of the Tricladida (Platyhelminthes: Seriata) with a discussion on morphological synapomorphies.** *Proc Biol Sci* **265(1396)**:631−640.

Carstens BC, Knowles LL (2007). **Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanopus* grasshoppers.** *Syst Biol* **56(3)**:400−411.

Carstens BC, Dewey TA (2010). **Species delimitation using a combined coalescent and information-theoretic approach: an example from North American *Myotis* bats.** *Syst Biol* **59(4)**:400−414.

Carstens BC, Pelletier TA, Reid NM, Satler JD (2013). **How to fail at species delimitation.** *Mol Ecol* **22(17)**:4369−4383.

Case JA (2002). **A new biogeographic model for dispersal of Late Cretaceous vertebrates into Madagascar and India.** *J Vert Paleontol* **22(3 Suppl.)**:42A.

Castle WE (1941). **The morphology and life-history of *Hymanella retenuova*, a new species of triclad from New England.** *Am Midl Nat* **26**:85−97.

Chen F, Li W (2001). **Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees.** *Am J Hum Genet* **68(2)**:444−456.

Charni M, Harrath AH, Sluys R, Tekaya S, Zghal F (2004). **The freshwater planarian *Dugesia sicula* Lepori, 1948 (Platyhelminthes, Tricladida) in Tunisia: ecology, karyology, and morphology.** *Hydrobiologia* **517(1−3)**:161−170.

Chinone A, Nodono H, Matsumoto M (2014). **Triploid planarian reproduces truly bisexually with euploid gametes produced through a different meiotic system between sex.** *Chromosoma* **In press.**

Cox CB, Moore PD (2005). **Biogeography: an ecological and evolutionary approach**, 7th edn. Blackwell, Oxford, UK.

Crisp MD, Trewick SA, Cook LG (2011). **Hypothesis testing in biogeography.** *Trends Ecol Evol* **26**:66−72.

Curini-Galletti M, Sanna D, Lai T, Cossu P, Scarpa F, Dedola GL, Delogu V, Casu M (2012). **Cryptic species and evolutionary patterns in interstitial microturbellaria: Monocelislineata (Proseriata) a case study.** *12th International Symposium on Flatworm Biology*, Naturhistoriska riksmuseet, Stockholm, Sweden.

Dahm AG (1958). **Taxonomy and ecology of five species groups in the family Planariidae.** Nya Litografen, Malmö, Sweden.

Dalyell IG (1814). **Observations on some interesting phaenomena on animal physiology exhibited by several species of planaria.** Edinburgh.

Dana JPM (1766). **De hirudinis nova specie, noxa, remediisque adhibendis (*H. alpina*).** *Mélanges Phylos Math Soc Roy Turin* **3**:199−212.

Datta-Roy A, Karanth KP (2009). **The Out-of-India hypothesis: what do molecules suggest?** *J Biosci* **34(5)**:687−697.

Davies RW, Reynoldson TB (1971). **The incidence and intensity of predation on lake-dwelling triclads in the field.** *J Anim Ecol* **40(1)**:191−214.

De Beauchamp P (1913). **Turbellaria Tricladida.** In: *Turbellariés, Trématodes et Gordiacés. Voyage de Ch. Alluaud et R. Jeannel en Afriques Orientale (1911−1912). Résultats Scientifiques* 3−16.

De Queiroz K (1998). **The general lineage concept of species, species criteria, and the process of speciation.** In: *Endless Forms: Species and Speciation* (ed. Howard, SJ and Berlocker, SH), pp. 57−78. Oxford University Press, New York.

De Queiroz K (2005). **Ernst Mayr and the modern concept of species.** *Proc Natl Acad Sci U S A* **102(Suppl 1)**:6600−6607.

De Queiroz K (2007). **Species concepts and species delimitation.** *Syst Biol* **56(6)**:879−886.

De Vries EJ (1984). **On the species of the *Dugesia gonocephala* group (Platyhelminthes, Turbellaria, Tricladida) from Greece.** *Bijdr Dierk* **54(1)**:101−126.

De Vries EJ (1987). **A contribution to the taxonomy, phylogeny and biogeography of the genus *Dugesia* sensu stricto (Platyhelminthes, Tricladida, Paludicola).** PhD thesis 1987, University of Amsterdam, Amsterdam.

De Vries EJ (1988a). **A synopsis of the nominal species of the subgenus *Dugesia* (Platyhelminthes: Tricladida: Paludicola) from Africa and Madagascar.** *Zool J Linn Soc* **92(4)**:345−382.

De Vries EJ (1988b). **Further contribution to the taxonomy and biogeography of the subgenus *Dugesia* (Platyhelminthes: Tricladida: Paludicola) in the Mediterranean region and the Middle East.** *Isr J Zool* **35(3)**:109−136.

De Vries EJ, Sluys R (1991). **Phylogenetic relationships of the genus *Dugesia* (Platyhelminthes, Tricladida, Paludicola).** *J Zool* **223(1)**:103−116.

Dentzien-Dias PC, Poinar G, de Figueiredo AEQ, Pacheco ACL, Horn BLD, Schultz CL (2013). **Tapeworm eggs in a 270 million-year-old shark coprolite.** *PLoS One* **8(1)**:e55007.

Dietz RS (1961). **Continental and ocean basin evolution by spreading of the sea floor.** *Nature* **190**, 854−57.

Doyle JJ (1992). **Gene trees and species trees: Molecular systematics as one-character taxonomy.** *Syst Bot* **17(1)**:144−163.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006). **Relaxed phylogenetics and dating with confidence.** *PLoS Biol* **4(5)**:e88.

Ducey PK, Cerqua J, West L-J, Warner M (2006). **Rare egg capsule production in the invasive terrestrial planarian *Bipalium kewense*.** *Southwest Nat* **51(2)**:252−254.

Duminil J, Kenfack D, Viscosi V, Grumiau L, Hardy OJ (2012). **Testing species delimitation in sympatric species complexes: the case of an African tropical tree, *Carapa* spp. (Meliaceae).** *Mol Phylogenet Evol* **62(1)**:275−285.

Durand JD, Templeton AR, Guinand B, Imsiridou A, Bouvet Y (1999). **Nested clade and phylogeographic analyses of the chub, *Leuciscus cephalus* (Teleostei, Cyprinidae), in Greece: implications for the Balkan Peninsula biogeography.** *Mol Phylogenet Evol* **13(3)**:566−80.

Edgecombe GD, Giribet G, Dunn CW, Hejnol A, Kristensen RM, Neves RC, Rouse GW, Worsaae K, Sørensen (2011). **Higher-level metazoan relationships: recent progress and remaining questions.** *Org Divers Evol* **11(2)**:151−172.

Edwards DL, Knowles LL (2014). **Species detection and individual assignment in species delimitation: can integrative data increase efficacy?** *Proc Biol Sci* **281(1777)**:20132765.

Egger B, Gschwentner R, Rieger R (2007). **Free-living flatworms under the knife: past and present.** *Dev Genes Evol* **217(2)**:89−104.

Felsenstein J (1978). **Cases in which parsimony or compatibility methods will be positively misleading.** *Syst Biol* **27(4)**:401−410.

Felsenstein J (1981). **Evolutionary trees from DNA sequences: A maximum likelihood approach.** *J Mol Evol* **17(6)**:368−376.

Fitch WM (1970). **Distinguishing homologous from analogous proteins.** *Syst Zool* **19(2)**:99−113.

Fitch WM (1971). **Toward defining the course of evolution: minimum change for a specified tree topology.** *Syst Zool* **20(4)**:406−416.

Forteath GNR, Osborn AW (2012). **Survival of endemic invertebrates of Lake Pedder and Lake Edgar subsequent to inundation.** *Rec Queen Vict Mus* **116**.

Gadagkar SR, Rosenberg MS, Kumar S (2005). **Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree.** *J Exp Zool B Mol Dev Evol* **304B(1)**:64−74.

Giribet G (2008). **Assembling the lophotrochozoan (=spiralian) tree of life.** *Philos Trans R Soc Lond B Biol Sci* **363(1496)**:1513−1522.

Girard CF (1850). **A brief account of the fresh water species of Planariae inhabiting the United States.** *Proc Boston Soc Nat Hist* **3**:264−265.

Glenner H, Hansen AJ, Sørensen MV, Ronquist F, Huelsenbeck JP, Willerslev E (2004). **Bayesian inference of the metazoan phylogeny; a combined molecular and morphological approach.** *Curr Biol* **14(18)**:1644−1649.

Goldman N (1993). **Statistical tests of models of DNA substitution.** *J Mol Evol* **36(2)**:182−198.

Goodman M, Czelusniak J, Moore W, Romero-Herrera AE, Matsuda G (1979). **Fitting the gene lineages into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences.** *Syst Zool* **28(2)**:132−163.

Gourbault N (1969). **Expansion de *Dugesia tigrina* (Girard), planaire américaine introduite en Europe.** *Ann Limnol* **5(1)**:3−7.

Grant L, Sluys R, Blair D (2006). **Biodiversity of Australian freshwater planarians (Platyhelminthes: Tricladida: Paludicola): New species and localities, and a review of paludicolan distribution in Australia.** *System Biodivers* **4(4)**:435−471.

Grasso M, Benazzi M (1973). **Genetic and physiologic control of fissioning and sexuality in planarians.** *J Embryol Exp Morphol* **30(2)**:317−328.

Grasso M, Montanaro L, Quaglia A (1975). **Studies on the role of neurosecretion in the induction of sexuality in a planarian agamic strain.** *J Ultrastruct Res* **52(3)**:404−408.

Gubitz T, Thorpe RS, Malhotra A (2000). **Phylogeography and natural selection in the Tenerife gecko *Tarentola delalandii:* Testing historical and adaptive hypotheses.** *Mol Ecol* **9(9)**:1213−1221.

Haria AH (1995). **Hydrological and environmental impact of earthworm depletion by the New Zealand flatworm (*Artioposthia triangulata*).** *J Hydrol (Amst)* **171(1−2)**:1−4.

Harrath AH, Sluys R, Aldahmash W, Al-Razaki A, Alwasel S (2013). **Reproductive strategies, karyology, parasites, and taxonomic status of *Dugesia* populations from Yemen (Platyhelminthes: Tricladida: Dugesiidae).** *Zool Sci* **30**:502−508.

Harrington RC, Near TH (2012). **Phylogenetic and coalescent strategies of species delimitation in snubnose darters (Percidae: *Etheostoma*).** *Syst Biol* **61(1)**:63−69.

Hausdorf B (2011). **Progress toward a general species concept.** *Evolution* **65(4)**:923−931.

Hausdorf B, Henning C (2007). **Biotic element analysis and vicariance biogeography.** In: *Biogeography in a changing world* (ed. Ebach, MC and Tangney, RS), pp. 95−115. CRC Press, Boca Raton, Florida.

Hauser J (1987). **Sexualization of *Dugesia anderlani* by feeding.** *Acta Biol Leopold* **9(1)**:111−128.

Hay WW, DeConto RM, Wold CN, Willson KM, Voigt S, Schulz M, Wold-Rossby A, Dullo WC, Ronov AB, Balukhovsky AN, Soeding E (1999). **An alternative global Cretaceous paleogeography. Evolution of Cretaceous ocean/climate systems.** In: *Evolution of the Cretaceous ocean-climate system* (ed. Barrera, E and Johnson, C), pp. 1−48. Geological Society of America Special Paper, 332.

Heads M (2005). **Dating nodes on molecular phylogenies: a critique of molecular biogeography.** *Cladistics* **21**:62−78.

Hess HH (1962). **History of ocean basins.** In: *Petrological studies: a volume in honor of A.F. Buddington* (ed. Engel, AEJ., James, HL and Leonard, BF), pp. 599−620. Geological Society of America, New York.

Hesse R (1897). **Untersuchungen über die Organe der Lichtempfindung bei niederen Tieren II. Die Augen der Plathelminthes, insbesondere der tricladen. Turbellarien.** *Z Wiss Zool* **65**:528−582.

Hey J (2006). **On the failure of modern species concepts.** *Trends Ecol Evol* **21(8)**: 447−450.

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback (2001). **Bayesian inference of phylogeny and its impact on evolutionary biology.** *Science* **294(5500)**:2310−2314.

Huelsenbeck JP, Andolfatto P, Huelsenbeck ET (2011). **Structurama: Bayesian inference of population structure.** *Evol Bioinform* **7**:55−59.

Hyman LH (1951). **The Invertebrates: Platyhelminthes and Rhynchocoela The Acoelomate Bilateria.** McGraw-Hill, London and New York.

Hugall A, Moritz C, Moussalii A, Stanisic J (2002). **Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875).** Proc Natl Acad Sci U S A **99(9)**:6112–6117.

Ichikawa A, Kawakatsu M (1964). **A new freshwater planarian, *Dugesia japonica*, commonly but erroneously known as Dugesia gonocephala (Duges)**. *Annot Zool Jpn* **37(3)**:185−194.

Ijima I, Kaburaki T (1916). **Preliminary descriptions of some Japanese Triclads.** *Annot Zool Jpn* **9(2)**:153−171.

Ishii S (1980a). **The ultrastructure of the protonephridial flame cell of the freshwater planarian *Bdellocephala brunnea*.** *Cell Tissue Res* **206(3)**:441−449.

Ishii S (1980b). **The ultrastructure of the protonephridial tubules of the freshwater planarian *Bdellocephala brunnea*.** *Cell Tissue Res* **206(3)**:451−458.

Issac NJ, Mallet J, Mace GM (2004). **Taxonomic inflation: its influence on macroecology and conservation.** *Trend Ecol Evol* **19(9)**:464−469.

Jörger KM, Schrödl M (2013). **How to describe a cryptic species? Practical challenges of molecular taxonomy.** *Front Zool* **10**:59.

Justine J, Winsor L, Gey D, Gros P, Thévenot J (2014). **The invasive New Guinea flatworm *Platydemus manokwari* in France, the first record for Europe: time for action is now.** *PeerJ* **2**:e29.

Kawakatsu M, Mitchell RW (1989). **Taxonomic and Karyological studies of *Dugesia batuensis* Ball, 1970 (Turbellaria, Tricladida, Paludicola), from the Batu Caves, Malaysia.** *J Speleol Soc Jpn* **14**:1−14.

Kawakatsu M, Mitchell RW (2004). **Platyhelminthes: Dugesiidae and Dimarcusidae.** In: *Freshwater invertebrates of the malaysian Region* (ed. Yule, CM and Sen, YH), pp. 55−65. Academy of Sciences Malaysia, Kuala Lumpur, Malaysia.

Kawakatsu M, Oki I, Tamura S, Sugino H (1976). **Studies on the morphology, karyology and taxonomy of the japanese freshwater planarians *Dugesia japonica* Ichikawa et Kawakatsu, with a description of a new subspecies *Dugesia japonica ryukyuensis* subspec. nov.** *Bull Fuji Women's College* **14:**81−126.

Kawakatsu M, Tamura S, Takai M, Yamamoto K, Ueno R, Oki I (1993). **The first record of occurrence of a naturalized population of *Dugesia tigrina* (Girard, 1850) at Nagasaki, Kyûshû, Japan (Turbellaria: Tricladida: Paludicola).** *Bull Biogeogr Soc Jpn* **48**:28−34.

Kawakatsu M, Tamura S, Yamayoshi T, Oki I (1980). **The freshwater planarians from Thailand and South India.** *Annot Zool Jpn* **53(4)**:254−268.

Keller J (1894). **Die ungeschlechtliche Fortpflanzungder Süsswasser-Turbellarien.** *Jen Zeit Naturw* **28**:370–407.

Kenk, R (1974). **Index of the genera and species of the freshwater triclads (Turbellaria) of the world.** *Smithson Contr Zool*, **183**.

Kimura M (1968). **Evolutionary rate at the molecular level.** *Nature* **217**:624−626.

Kimura M (1983). **The Neutral Theory of Molecular Evolution.** Cambridge University Press, Cambridge, UK.

Kishino H, Thorne JL, Bruno WJ (2001). **Performance of a divergence time estimation method under a probabilistic model of rate evolution.** *Mol Biol Evol* **18(3)**:352−361

Klaus S, Schubart CD, Streit B, Pfenninger M (2010). **When Indian crabs were not yet Asian − biogeographic evidence for Eocene proximity of India and Southeast Asia.** *BMC Evolutionary Biology* **10**:287.

Klima J (1961). **Elektronenmikroskopische Studien über die Feinstruktur der Tricladen *(Turbellaria).*** *Protoplasma* **54(1)**:101−162.

Knowles LL, Carstens BC (2007). **Delimiting species without monophyletic gene trees.** *Syst Biol* **56(6)**:887−895.

Kobayashi K, Koyanagi R, Matsumoto M, Padilla Cabrero J, Hoshi M (1999). **Switching from asexual to sexual reproduction in the planarian *Dugesia ryukyuensis*: bioassay system and basic description of sexualizing process.** *Zoolog Sci* **26(2)**:392−398.

Kobayashi K, Maezawa T, Nakagawa H, Hoshi M (2012). **Existence of two sexual races in the planarian species switching between asexual and sexual reproduction.** *Zool Sci* **29(4)**: 265−272.

Kolaczkowski B, Thornton JW (2004). **Performance of maximum parsimony and maximum likelihood phylogenetics when evolution is heterogeneous.** *Nature* **431**:980−984.

Kondandaramaiah U (2011). **Tectonic calibrations in molecular dating.** *Curr Zool* **57(1)**:116−124.

Kondrashov AS (1998). **Deleterious mutations and the evolution of sexual reproduction.** *Nature* **336**:435−440.

Koy KA, Plotnich RE (2008). **Effect of the odorant of the predatory flatworm *Dugesia gonocephala* on the foraging behavior of *Caenorhabditis elegans*.** *J Nematol* **40(4)**:286−289.

Kubatko L, Degnan J (2007). **Inconsistency of phylogenetic estimates from concatenated data under coalescence.** *Syst Biol* **56(1)**:17−24.

Lázaro EM (2012). **Filogeografia de planàries d'aigua dolça (Tricladida, Platyhelminthes) a la Mediterrània Occidental.** PhD thesis, Universitat de Barcelona, Barcelona.

Lázaro EM, Harrath H, Stocchino G, Pala M, Baguñà J, Riutort M (2011). ***Schmidtea mediterranea* phylogeography: an old species surviving on a few Mediterranean islands?** *BMC Evol Biol* **11**:274.

Lázaro EM, Riutort M (2013). ***Dugesia sicula* (Platyhelminthes, Tricladida): the colonizing success of an asexual planarian.** *BMC Evol Biol* **13**:268.

Lázaro EM, Sluys R, Pala M, Stocchino GA, Baguñà J, Riutort M (2009). **Molecular barcoding and phylogeography of sexual and asexual freshwater planarians of the genus *Dugesia* in the Western Mediterranean (Platyhelminthes, Tricladida, Dugesiidae).** *Mol Phylogenet Evol* **52(3)**:835−845.

Leaché AD, Fujita MK, Minin VN, Bouckaert RR (2013). **Species delimitation using genome-wide SNP data.** *Syst Biol* **In press.**

Leech ML, Singh S, Jain AK, Klemperer SL (2005). **The onset of India-Asia continental collision: early, steep subduction required by the timing of UHP metamorphism in the western Himalaya.** *Earth Planet Sci Lett* **234(1−2)**:83−97.

Leliart F, Verbruggen H, Wysor B, De Clerk O (2009). **DNA taxonomy in morphologically plastic taxa: Algorithmic species delimitation in the *Boodlea* complex (Chlorophyta: Cladophorales).** *Mol Phylogent Evol* **53(1)**:122−133.

Leloup E (1944). **Recherches sur les triclades dulçicoles epigés de la Forêt de Soignes.** *Mém Mus R His Nat Bel* **102**:1−112.

Lepori NG (1948). **Descrizione di *Dugesia sicula*, nuova sp. di Triclade d'acqua dolce dei dintorni di Catania.** *Arch Zool Ital* **33**:461−472.

Lepori NG (1951). **Sulle caratteristiche morfologiche e sulla posizione sistematica della planaria di Sardegna e Corsica già ascritta *Dugesia* (Euplanaria) *gonocephala* (Dugès).** *Atti Soc Tose Sci Nat Ser B* **58**:28−47.

Lieberman BS (2002). **Phylogenetic biogeography with and without the fossil record: gauging the effects of extinction and paleontological incompleteness.** *Palaeogeogr Palaeoclimatol Palaeoecol* **178**:39−52.

Lischetti AB (1919). **Un verme del genero Planaria, enemigo natural de las larvas del mosquito.** *Physis (B Aires)* **4**:591−595.

Liu S-Y, Seleck C, Friedrich B, Lutz R, Vila-Farré M, Dahl A, Brandl H, Lakshmanaperumal N, Henry I, Rink JC (2013). **Reactivating head regrowth in a regeneration-deficient planarian species.** *Nature* **500**:81−84.

Lomolino MV, Riddle BR, Brown JH (2006). **Biogeography**, 3rd edn. Sinauer Associates, Sunderland, Massachusetts.

MacDonald GM (2003). **Biogeography: Introduction to space, time and life.** John Wiley, New York.

MacRae E (1967). **The fine structure of sensory receptor processes in the auricular epithelium of the planarian, *Dugesia tigrina*.** *Z Zellforsch Mikrosk Anat* **82(4)**:479−494.

Maddison WP (1997). **Gene trees in species trees.** *Syst Biol* **46(3)**:523−536.

Magallon SA (2004). **Dating lineages: molecular and paleontological approaches to the temporal framework of clades.** *Int J Plant Sci* **165(4 Suppl.)**:S7−S21.

Mallet J (2007). **Hybrid speciation.** *Nature* **446**:279−283.

Marcus R (1946). **Sobre Turbellaria brasileiros.** *Bol Fac Fil Ci Letras, Univ São Paulo, Zool* **11:**5−253.

Marcus E (1953). **Turbellaria Tricladida.** *Inst Parcs Nat Congo Belge Expl Parc Nat Upemba Mission G F de Witte* **21**:1−62.

Mayer F, Dietz C, Kiefer A (2007). **Molecular species identification boosts bat diversity.** *Front Zool* **4**:4.

Milllington A, Blumer M, Schickhoff U (2011). **The SAGE handbook of biogeography.** SAGE Publications Ltd.

Moore WS (1995). **Inferring Phylogenies from mtDNA Variation: mitochondrial-gene trees versus nuclear-gene trees.** *Evolution* **49(4)**:718−726.

Morando M, Avila LJ, Sites JW (2003). **Sampling strategies for delimiting species: genes, individuals, and populations in the *Liolaemus elongatus-kriegi* complex (Squamata: Liolaemidae) in Andean-Patagonian South America.** *Syst Biol* **52(2)**:159−185.

Moritz C (1994). **Defining "Evolutionarily Significant Units" for conservation.** *Trends Ecol Evol* **9(10)**:373−375.

Mossel E, Vigoda E (2005). **Phylogenetic MCMC algorithms are misleading on mixtures of trees.** *Science* **309(5744)**:2207−2209.

Müller OF (1774). **Vermium terrestrium et fluviatilium, seu animalium infusoriorum, helminthicorum et testaceorum, non marinorum, succincta historia.** Vol. Primi pars altera. Havniæ & Lipsiæ.

Neppi V (1904). **Über einige exotische Turbellarien.** *Zool Jahrb Abt Syst Geogr Biol Tiere* **21**:303−326.

Newmark PA, Sánchez-Alvarado A (2002). **Not your father's planarian: a classic model enters the era of functional genomics.** *Nat Rev Gen* **3**:210−219.

Ogren R, Kawakatsu M, Froehlich EM (1993). **What is the silicified Rhynchodemid species from calico mountain nodules?** In: Additions and corrections of the previous land planarian indices of the world (Turbellaria, Tricladida, Terricola). *Bull Fuji Women's College* **31**:61−86.

O'Meara BC (2010). **New heuristic methods for joint species delimitation and species tree inference.** *Syst Biol* **59(1)**:59−73.

Padial JM, Miralles A, De la Riva I, Vences M (2010). **The integrative future of taxonomy.** *Front Zool* **7**:1−14.

Pagel M (1999). **Inferring the historical patterns of biological evolution.** *Nature* **201**:877−884.

Pala M, Casu S, Vacca RA (1980). **Primi dati sulla distribuzione geografica di planarie del gruppo Dugesia lugubris-polychroa (Turbellaria, Tricladida) in Sardegna, con particolare riguardo alla presenza di una mutazione tetraploide nel fiume Cixerri (Siliqua, Cagliari).** *Boll Soc Sarda Sci Nat* **19**:183−188.

Pala M, Vacca RA, Casu S, Stocchino G (1995). **The freshwater planarian *Dugesia sicula* from Sardinia (Platyhelminthes, Tricladida).** *Hydrobiologia* **310**:151−156.

Pallas PS (1774). **Spicilegia zoologica quibus novae imprimis et obscurae animaliu. speciosiconibus atque conamentariis illustratur.** Fasc X, Berolini.

Pamilo P, Nei M (1988). **Relationships between gene trees and species trees.** *Mol Biol Evol* **5(5)**:568−583.

Patriat P, Achache J (1984). **India-Asia collision chronology has implications for crustal shortening and driving mechanism of plates.** *Nature* **311**:615−621.

Pérez-Losada M, Hoeg JT, Crandall KA (2004). **Unraveling the evolutionary radiation of the thoracican barnacles using molecular and morphological evidence: A comparison of several divergence time estimation approaches.** *Syst Biol* **53(2)**:244−264.

Peterson AT (2001). **Predicting species geographic distributions based on ecological niche modeling.** *Condor* **103(3)**:599−605.

Peterson KJ, Lyons JB, Nowak KS, Takacs CM, Wargo MJ, McPeek MA (2004). **Estimating metazoan divergence times with a molecular clock.** *Proc Natl Acad Sci U S A* **101(17)**:6536−6541.

Peterson KJ, Cotton JA, Gehling JG, Pisani D (2008). **The Ediacaran emergence of bilaterians: congruence between the genetic and the geological fossil records.** *Philos Trans R Soc Lond B Biol Sci* **363(1496)**:1435−1443.

Pierce WD (1960). **Silicified turbellaria from Calico Mountains nodules.** *Bull South Calif Acad Sci* **59**:138−143.

Poinar G (2003). **A rhabdocoel turbellarian (Platyhelminthes, Typhloplanoida) in Baltic amber with a review of fossil and sub-fossil platyhelminths.** *Invertebr Biol* **122(4)**:308−312.

Poinar G (2004). **Correction of a proposed generic name for a fossil turbellarian.** *Invertebr Biol* **123(2)**:181.

Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogles AP (2006). **Sequence-based species delimitation for the DNA taxonomy of undescribed insects.** *Syst Biol* **55(4)**:595−605.

Prados J, Álvarez B, Howarth J, Stewart K, Gibson CL, Hutchinson CV, Young AM, Davidson C (2013). **Cue competition effects in the planarian.** *Anim Cogn* **16(2)**:177−86.

Prasad GVR, Sahni A (2009). **Late Cretaceous continental vertebrate fossil record from India: palaeobiogeographical insights.** *B Soc Géol Fr* **180(4)**:369− 381.

Price JP, Clague DA (2002). **How old is the Hawaiian biota? Geology and phylogeny suggest recent exchange.** *Proc R Soc Lond B* **269**:2429−2435.

Quental TB, Marshall CR (2010). **Diversity dynamics: molecular phylogenies need the fossil record.** *Trends Ecol Evol* **25**:434−441.

Ramakrishan L, Amatya C, DaSaer CJ, Dalhoff Z, Eggerichs MR (2014). **Galantamine reverses scopolamine-induced behavioral alterations in *Dugesia tigrina*.** *Invert Neurosci* **In press.**

Rambaut A, Bromham L (1998). **Estimating divergence dates from molecular sequences.** *Mol Biol Evol* **15(3)**:442−448.

Reddien PW, Sánchez−Alvarado A (2004). **Fundamentals of planarian regeneration.** *Annu Rev Cell Dev Biol* **20**:725−757.

Ree RH, Moore BR, Webb CO, Donoghue MJ (2005). **A likelihood framework for inferring the evolution of geographic range on phylogenetic trees.** *Evolution* **59(11)**:2299−2311.

Ree RH, Smith SA (2008). **Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis.** *Syst Biol* **57(1)**:4−14.

Reynoldson TB (1961). **Environment and reproduction in freshwater triclads.** *Nature* **189**:329−330.

Reynoldson TB (1966). **The distribution and abundance of lake-dwelling triclads − towards a hypothesis.** *Adv Ecol Res* **3**:1−71.

Ribas M (1990). **Cariologia, sistemàtica i biogeografia de les planàries d'aigües dolces als Països Catalans.** PhD thesis, Universitat de Barcelona, Barcelona.

Riddle BR, Dawnson MN, Hadly EA, Hafner DJ, Hickerson MJ, Mantooth SJ, Yoder A (2008). **The role of molecular genetics in sculpting the future of integrative biogeography.** *Progr Phys Geogr* **32(2)**:173−202

Rieseberg LH, Wendel JF (1993). **Introgression and its consequences in plants.** In: *Hybrid zones and the evolutionary process* (ed. Harrison, RG), pp. 70−109. Oxford University Press, Oxford.

Riutort M, Lázaro EM, Álvarez-Presas M, Solà E, Paps, J (2012). **Evolutionary history of the Tricladida and the Platyhelminthes: an up-to-date phylogenetic and systematic account.** *Int J Dev Biol* **56(1−3)**:5−17.

Rokas A, Carroll SB (2005). **More genes or more taxa? The relative contribution of gene number and taxon number to phylogenetic accuracy.** *Mol Biol Evol* **22(5)**:1337−1344.

Rokas A, Williams B, King N, Carroll SB (2003). **Genome-scale approaches to resolving incongruence in molecular phylogenies.** *Nature* **425**:798−804.

Ronquist F, Klopfstein S, Vilhelmsen L, Schulmeister S, Murray DL, Rasnitsyn AP (2012). **A total-evidence approach to dating with fossils, applied to the early radiation of the hymenoptera.** *Syst Biol* **61(6)**:973−999.

Romero R (1987). **Anàlisi cel·lular i quantitatiu del creixement i de la reproducció a diferents espècies de planàries.** PhD thesis, Universitat de Barcelona, Barcelona.

Rosenberg MS, Kumar S (2003). **Heterogeneity of nucleotide frequencies among evolutionary lineages and phylogenetic inference.** *Mol Biol Evol* **20(4)**:610−621.

Rosell JA, Olson ME, Weeks A, De-Nova JA, Lemos RM, Camacho JP, Feria TP, Gómez-Bermejo R, Montero JC, Eguiarte LE (2010). **Diversification in species complexes: tests of species origin and delimitation in the *Bursera simaruba* clade of tropical trees (Burseraceae).** *Mol Phylogenet Evol* **57(2)**:798−811.

Sakai M, Sakaizumi M (2012). **The complete mitochondrial genome of *Dugesia japonica* (Platyhelminthes; Order Tricladida).** *Zoolog Sci* **29(10)**:672−680.

Sakurai T (1981). **Sexual induction by feeding in an asexual strain of the freshwater planarian, *Dugesia japonica japonica*.** *Annot Zool Jpn* **54(2)**:103−112

Sánchez-Alvarado A (2000). **Regeneration in the metazoans: why does it happen?** *Bioessays* **22(6)**:578−590.

Sanmartín I, Ronquist F (2004). **Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns.** *Syst Biol* **53(2)**:216−243.

Satler JD, Carstens BC, Hedin M (2013). **Multilocus species delimitation in a complex of morphologically conserved trapdoor spiders (Mygalomorphae, Antrodiaetidae, *Aliatypus*).** *Syst Biol* **62(6)**:805−823.

Schmidt O (1861). **Über Planaria torva Autorum.** *Z Wiss Zool* **11**:89−94.

Shockaert ER, Hooge M, Sluys R, Schilling S, Tyler S, Artois T (2008). **Global diversity of free living flatworms (Platyhelminthes, "Turbellaria") in freshwater.** *Hydrobiologia* **595**:41−48.

Schlick-Steiner BC, Steiner FM, Seifert B, Stauffer C, Christian E, Crozier RH (2010). **Integrative taxonomy: a multisource approach to exploring biodiversity.** *Annu Rev Entomol* **55**:421−438.

Schwarz GE (1978). **Estimating the dimension of a model.** *Ann Stat* **6(2)**:461−464.

Simpson GG (1951). **The species concept.** *Evolution* **5**:285−298.

Skaer RJ (1961). **Some aspects of the cytology of *Polycelis nigra*.** *Q J Microsc Sci* **102**:295−317.

Slowinski J, Page RD (1999). **How should species phylogenies be inferred from sequence data?** *Syst Biol* **48(4)**:814−825.

Sluys R (1989). **Phylogenetic relationships of the triclads (Platyhelminthes, Seriata, Tricladida)**. *Bijdr Dierk* **59(1)**:3−25.

Sluys R (2001). **Towards a phylogenetic classification and characterization of dugesiid genera (Platyhelminthes, Tricladida, Dugesiidae): A morphological perspective.** In: *Interrelationships of the Platyhelminthes* (ed. Littlewood, DTJ and Bray, RA), pp. 57−73. Taylor & Francis, London.

Sluys R (2007). **Annotations on freshwater planarians (Platyhelminthes Tricladida Dugesiidae) from the Afrotropical Region.** *Trop Zool* **20**:229−257.

Sluys R, Kawakatsu M, Winsor L (1998). **The genus *Dugesia* in Australia, with its phylogenetic analysis and historical biogeography (Platyhelminthes, Tricladida, Dugesiidae).** *Zool Scr* **27(4)**:273−289.

Sluys R, Kawakatsu M, Yamamoto K (2010). **Exotic freshwater planarians currently known from Japan.** *Belg J Zool* **140(Suppl.)**:103−109.

Sluys R, Kawakatsu M, Riutort M, Baguñà J (2009). **A new higher classification of planarian flatworms (Platyhelminthes, Tricladida).** *J Nat Hist* **43(29−30)**:1763−1777.

Sluys R, Solà E, Gritzalis K, Vila-Farré M, Mateos E, Riutort M (2013). **Integrative delineation of species of Mediterranean freshwater planarians (Platyhelminthes: Tricladida: Dugesiidae).** *Zool J Linn Soc* **169(3)**:523−547.

Smith AB, Pisani D, Mackencie-Dodds JA, Stockley B, Webster BL, Littlewood DT (2006). **Testing the molecular clock: molecular and paleontological estimates of divergence times in the Echinoidea (Echinodermata).** *Mol Biol Evol* **23(10)**:1832−1851.

Spellerberg IF, Sawyer JWD (1999). **An introduction to applied biogeography.** Cambrige University Press, Cambridge.

Spiegelman M, Dudley PL (1973). **Morphological stages of regeneration in the planarian *Dugesia tigrina:* A light and electron microscopic study.** *J Morphol* **139(2)**:155−183.

Stage HH, Yates WW (1939). **Ground beetle predatory on the eggs of *Aedes* mosquitoes.** *Proc Entomol Soc Wash* **41**:204−206.

Stanovic S (1960). **The Balkan Lake Ohrid and its living World. Monographie Biologicae Vol, IX.** Dr. W. Junk, Den Haag, Netherlands.

Stocchino GA, Corso G, Manconi R, Pala M (2002). **African planarians: *Dugesia aethiopica* sp. n. (Platyhelminthes, Tricladida) from Lake Tana (NW Ethiopia).** *Ital J Zool* **69(1)**:45−51.

Stocchino GA, Manconi R, Corso G, Sluys R, Casu S, Pala M (2009). **African planarians: Morphology and karyology of *Dugesia maghrebiana* sp. n. (Platyhleminthes, Tricladida) from Tunisia.** *Ital J Zool* **76(1)**:83-91.

Stocchino GA, Sluys R, Deri P, Renata M (2013). **Integrative taxonomy of a new species of planarian from the Lake Ohrid basin, including an analysis of biogeographical patterns in freshwater triclads from the Ohrid region (Platyhelminthes, Tricladida, Dugesiidae)**. *ZooKeys* **313**:25−43.

Stocchino GA, Sluys R, Manconi R (2012). **A new species of *Dugesia* (Platyhelminthes, Tricladida, Dugesiidae) from Afromontane forest in South Africa, with an overview of freshwater planarians from the African continent.** *Zootaxa* **3551**:43−58.

Takahata N (1989). **Gene genealogy in three related populations: Consistency probability between gene and population trees.** *Genetics* **122(4)**:957−966.

Teshirogi W (1986). **On the origin of neoblasts in freshwater planarians (Turbellaria).** *Hydrobiologia* **132(1)**:207−216.

Tierney SM, Smith JA, Chenoweth L, Schwarz MP (2008). **Phylogenetics of allodapine bees: a review of social evolution, parasitism and biogeography.** *Apidologie* **39(1)**:3−15.

Török LJ, Röhlic P (1959). **Contributions to the fine structure of the epidermis of *Dugesia lugubris*. O. Schm.** *Acta Biol Acad Sci Hung* **10**:23−48.

Ullyott P, Beauchamp RSA (1931). **Mechanisms for the prevention of self-fertilization in some species of fresh-water triclads.** *Q J Microsc Sci* **74**:477−490.

Ullyott P (1936). **A note on the zoogeographical history of north western Europe.** *Proc Prehist Soc* **2**:169−177.

Upchurch P (2008). **Gondwanan break-up: legacies of a lost world?** *Trends Ecol Evol* **23(4)**:229−236.

Upeniece I (2001). **The unique fossil assemblage from the Lode Quarry (Upper Devonian, Latvia).** *Mitt Mus Natkd Bcrl Geowiss Reihe* **4(1)**:101−119.

Van Hinsbergen DJJ, Lippert PC, Dupont-Nivet G, McQuarrite N, Doubrovine PV, Spakman W, Torsvik TH (2012a). **Greater India Basin hypothesis and a two-stage Cenozoic collision between India and Asia.** *Proc Natl Acad Sci U S A* **109(20)**:7659−7664.

van Hinsbergen DJJ, Lippert PC, Dupont-Nivet G, McQuarrite N, Doubrovine PV, Spakman W, Torsvik TH (2012b). **Reply to Aitchison and Ali: Reconciling Himalayan ophiolite and Asian magmatic arc records with a two-stage India-Asia collision model.** *Proc Natl Acad Sci U S* **109(40)**:E2646.

Vieites DR, Wollenberg KC, Andreone F, Köhler J, Glaw F, Vences M (2009). **Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory.** *Proc Natl Acad Sci U S A* **106(20)**:8267−8272.

Vila-Farré M (2011). **Planarias ibéricas: taxonomía y biogeografía.** PhD thesis, Universitat de Barcelona, Barcelona.

Vila-Farré M, Sluys R, Almagro Í, Handberg-Thorsager M, Romero, R (2011). **Freshwater planarians (Platyhelminthes, Tricladida) from the Iberian Peninsula and Greece: diversity and notes on ecology.** *Zootaxa* **2779**:1−38.

Wagner DE, Wang IE, Reddien PW (2011). **Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration.** *Science* **332(6031)**:811-816.

Waters JM, Craw D, Youngson JH, Wallis GP (2001). **Genes meet geology: Fish phylogeographic pattern reflects ancient, rather than modern, drainage connections.** *Evolution* **55(9)**:1844−1851.

Weiss A (1909). **Turbellaria, Rhabdocoelida and Tricladida paludicola.** In: *Die Fauna Südwest Australiens: Ergebnisse der Hamburger südwest−australischen Forschungsreise 1905* (ed. Michaelsen, W and Hartmeyer, R), volume 2, pp. 403−410. Fischer, Jena.

Welton L, Siler C, Oaks J, Diesmos A, Brown R (2013). **Multilocus phylogeny and Bayes estimates of species boundaries reveal hidden evolutionary relationships and cryptic diversity in Southeast Asian monitor lizards.** *Mol Ecol* **22(13)**:3495−3510.

Wheeler QD, Raven PH, Wilson EO (2004). **Taxonomy: impediment or expedient?** *Science* **303(5656)**:285.

Wiens JJ, Donoghue MJ (2004). **Historical biogeography, ecology and species richness.** *Trends Ecol Evol* **19(2)**:639−644.

Wiens JJ, Penkrot TA (2002). **Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*).** *Syst Biol* **52(1)**:69−91.

Wiley EO (1978). **The evolutionary species concept reconsidered.** *Syst Zool* **27(1)**:17−26.

Wilhelmi J (1908). **Sinnesorgane der Auriculargegend bei Süβwassertricladen.** *Zool Anz* **33**:388−393.

Wilson EO (2003). **The Encyclopedia of Life.** *Trends Ecol Evol* **18(2)**:77−80.

Wojcicki M, Brooks DR (2006). **PACT: an efficient and powerful algorithm for generating area cladograms.** *J Biogeogr* **32(5)**:755−74.

Woodworth WM (1897). **Contribution to the morphology of the Turbellaria, II: On some Turbellaria from Illinois.** *Bull Mus Comp Zool* **31**:1−16.

Wright JF (1974). **Some factors affecting the distribution of *Crenobia alpina* (Dana), *Polycelis felina* (Dalyell) and *Phagocata vitta* (Dugès) (Platyhelminthes) in Caernarvonshire, North Wales.** *Freshw Biol* **4(1)**:31−59.

Wu CI (1991). **Inferences of species phylogeny in relation to segregation of ancestral polymorphisms.** *Genetics* **127(2)**:429−435.

Yang Z (1994). **Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods.** *J Mol Evol* **39(3)**:306−314.

Yang Z, Rannala B (2006). **Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds.** *Mol Biol Evol* **23(1)**:212−226

Yang Z, Rannala B (2010). **Bayesian species delimitation using multilocus sequence data.** *Proc Natl Acad Sci U S A* **107(20)**:9264−9269.

Yoder AD, Yang Z (2000). **Estimation of primate speciation dates using local molecular clocks.** *Mol Biol Evol* **17(7)**:1081−1090.

Yoder AD, Nowak MD (2006). **Has vicariance or dispersal been the predominant biogeographic force in Madagascar? Only time will tell.** *Annu Rev Ecol Evol Syst* **37**:405−431.

Yuan ZQ, Zhang JY, Zhao BS (2014). **Molecular cloning and expression pattern of DjStag gene in the planarian *Dugesia japonica* during embryonic development.** *Genet Mol Res* **13(1)**:188−197

Zhang H-C, Sun J, Chen G-Q, Liu D-Z (2010). **Molecular systematics of eight populations of *Dugesia japonica* in Henan Province based on COI sequences.** *Dong wu fen lei xue bao* **35(4)**:770−775.

Zhou WW, Wen Y, Fu J, Xu YB, Jin JQ, Ding L, Min MS, Che J, Zhang YP (2012). **Speciation in the *Rana chensinensis* species complex and its relationship to the uplift of the Qinghai-Tibetan Plateau.** *Mol Ecol* 2012, **21(4)**:960−973.

Zhu BD, Kidd WSF, Rowley DB, Currie BS, Shaffique N (2005). **Age of initiation of the India-Asia collision in the east-central Himalaya.** *J Geol* **113**:265−285.

Zuckerkandl E, Pauling L (1962). **Molecular disease, evolution, and genetic heterogeneity.** In: *Horizons in Biochemistery* (ed. Kasha, M and Pullman, B), pp. 189−225. Academic Press, New York.

Zuckerkandl E, Pauling L (1965). **Evolutionary divergence and convergence in proteins.** In: *Evolving Genes and Proteins* (ed. Bryson, V and Vogel HJ), pp. 97−166. Academic Press, New York.

Section VII

Annexes

1

Annex 1 − Tables

Annexes 1 – Table 1. List of valid *Dugesia* species. It includes their known distribution range and the synonyms proposed up-to-date. Updated from Kenk (1974). All the references are available at http://continenticola.lifedesks.org.

Currently valid *Dugesia* species	Distribution	Synonyms
D. aenigma De Vries, 1984	Greece: Cephalonia	
D. aethiopica Stocchino, Corso, Manconi & Pala, 2002	Ethiopia: Lake Tana	
D. andamanensis (Kaburaki, 1925)	Andaman Islands	*Planaria andamanensis* n. sp. Kaburaki, 1925, 1925:29 *Euplanaria (Euplanaria) andamanensis*: Kenk, 1930:292 *Dugesia andamanensis*: de Beauchamp, 1951b:96
D. annadalei (Kaburaki, 1918)	Burma/Myanmar	*Planaria annandalei* n. sp. Kaburaki, 1918a:191 *Euplanaria (Euplanaria) annandalei*: Kenk, 1930:292 *Dugesia annandalei*: de Beauchamp, 1951:96 *Dugesia (Dugesia) annandalei*: Ball, 1974:376
D. afromontana Stocchino & Sluys, 2012	South Africa	
D. arabica Harrath & Sluys, 2013	Yemen	
D. arcadia De Vries, 1988	Greece: Peloponnese	
D. ariadnae De Vries, 1984	Greece: Naxos	
D. ariesiana Sluys & Grant, 2007	Australia	
D. astrocheta Marcus, 1953	Democratic Republic of Congo	*Dugesia (Dugesia) astrocheta:* Ball, 1974:376
D. austroasiatica Kawakatsu, 1985	Southeast Asia; Japan	
D. bactriana De Beauchamp, 1959	Afghanistan; Pakistan	*Dugesia bactrians*: Ali, 1871:73 [spelling error] *Dugesia (?Dugesia) bactriana*: Ball, 1974:376.
D. bakurianica (Porfirjeva, 1958)	Caucasus	*Dugesia gonocephala bakurianica* Porfirjeva, 1958 *Dugesia gonocephala bacurianica* Porfirjeva, 1958:88 [spelling error]
D. batuensis Ball, 1970	Malaysia: Batu Caves	*Dugesia (Dugesia) batuensis*: Ball, 1974:376
D. benazzii Lepori, 1951	Corsica; Sardinia	*Dugesia (Dugesia) benazzii*: Ball, 1974:376
D. bengalensis Kawakatsu, 1983	India: West Bengal	
D. biblica Benazzi & Banchetti, 1972 *	Israel; Turkey	
D. borneana Kawakatsu, 1972	Borneo	
D. brigantii De Vries & Benazzi, 1983	Italy: Liguria	

D. burmaensis (Kaburaki, 1918)	Myanmar/Burma	*Planaria burmaensis* Kaburaki, 1918 *Planaria (Euplanaria) burmaensis*: Kenk, 1930:292 *Dugesia (?Dugesia) burmaensis*: Ball, 1974:379.
D. capensis Sluys, 2007	South Africa	
D. colapha Dahm, 1971	Ghana	*Dugesia (Dugesia) colapha*: Ball, 1974 376
D. congolensis De Beauchamp, 1951	Democratic Republic of Congo	*Dugesia (Dugesia) congolensis*: Ball, 1974:376
D. cretica (Meixner, 1928)	Greece: Crete	*Euplanaria cretica* Meixner, 1928:585 *Euplanaria (Euplanaria) cretica*: Kenk, 1930:292 *Euplanaria gonocephala cretica*: de Beauchamp, 1939a:64 *Euriplanaria cretica*: Filipović, 1957:4 [spelling error] *Dugesia (Dugesia) cretica*: Ball, 1974:376.
D. damoae De Vries, 1984	Greece: Samos	
D. debeauchampi De Vries, 1988	Madagascar: Nosy Bé	
D. deharvengi Kawakatsu & Mitchell, 1989	Thailand	
D. didiaphragma De Vries, 1988	Democratic Republic of Congo	
D. ectophysa Marcus, 1953	Democratic Republic of Congo	*Dugesia (Dugesia) ectophysa*: Ball, 1974:376
D. effusa Sluys, 2013	Greece: Chios	
D. elegans De Vries, 1984	Greece: Rhodes	
D. etrusca Benazzi, 1944	Italy	*Dugesia (Euplanaria) etrusca* n. sp. Benazzi, 1944:19 *Dugesia (Dugesia) etrusca*: Ball, 1974:376.
D. golanica Bromley & Benazzi, 1991	Israel: Banias	
D. gonocephala (Dugès, 1830)	Europe	*Planaria gonocephala* n. sp. Dugès, 1830:83 *Dugesia gonocephala*: Girard, 1850:256 *Planaria torva* var. *gonocephala*: Diesing, 1850:206 *Goniocarena gonocepahala*: Schmarda, 1859:14 *Euplanaria (Euplanaria) gonocephala*: Hesse, 1897:549 *Euplanaria (Euplanaria) gonocephala*: Kenk, 1930:291 *Planaria genocephala* Kasymov, 1965:182 [spelling error] *Dugesia (Dugesia) gonocephala*: Ball, 1974:376 ?*Planaria trigonocephala* nom. nud. Böhmig, 1883:Ixxxxix
D. hepta Pala, Casu & Vacca, 1981	Sardinia	

D. hymanae (Šivickis, 1928)	Philippines	*Planaria hymani* n. sp. Šivickis, 1928:360 *Dugesia hymanae* nom. emend. Marcus, 1953:12
D. ilvana Lepori, 1948	Italy: Elba island	*Dugesia gonocephala* var. *ilvana nom.* nud. Benazzi, 1945:270 *Dugesia (Dugesia) ilvana*: Ball, 1974:376
D. improvisa Sluys & Solà, 2013	Greece: Naxos	
D. indica Kawakatsu, 1969	India: Madhya Pradesh	*Dugesia (Dugesia) indica*: Ball, 1974:376
D. indonesiana Kawakatsu, 1973	Indonesia	
D. iranica (Livanov, 1951)	Iran; Turkey	*Euplanaria iranica* n. sp. Livanov, 1951:110 *Dugesia (Dugesia) iranica* Ball, 1974:376.
D. izuensis Kato, 1943	Japan	*Dugesia (?Dugesia) izuensis*: Ball, 1974:376
D. japonica Ichikawa & Kawakatsu, 1964	China; Japan; Taiwan	*Dugesia (Dugesia) japonica* Ball, 1974:376 *Dorotocephala gonocephala* Brøndsted, 1969:84 [lapsus calami] Previously designated as *Planaria*, *Euplanaria*, or *Dugesia gonocephala*
D. krishnaswamyi Kawakatsu, 1975	South India: Tamir Nadu	
D. lamottei De Beauchamp, 1952	Guinea	*Dugesia (Dugesia) lamottei*: Ball, 1974:376
D. lanzai Banchetti & del Papa, 1971	Kenya; Uganda	
D. leclerci Kawakatsu & Mitchell, 1995	Indonesia: Sulawesi	
D. leporii Pala, Stocchino, Corsa & Casu, 2000	Sardinia	
D. libanica Bromley & Benazzi, 1991	Lebanon	
D. liguriensis De Vries, 1988	Italy: Liguria	
D. lindbergi De Beauchamp, 1959	Afghanistan	*Dugesia (?Dugesia) lindbergi*: Ball, 1974:376
D. machadoi De Beauchamp, 1952	Southwest Africa: Angola	*Dugesia (Dugesia) machadoi*: Ball, 1974:376
D. maghrebiana Stocchino, Manconi, Corso, Sluys, Casu & Pala, 2009	Tunisia	
D. malickyi De Vries, 1984	Greece	
D. mertoni (Steinmann, 1914)	Indonesia: Kai Islands	*Planaria mertoni* n. sp. Steinmann, 1914:111 *Euplanaria (Euplanaria) mertoni*: Kenk, 1930:292 *Dugesia (?Dugesia) mertoni*: Ball, 1974:376
D. milloti de Beauchamp, 1952	Madagascar; Comoro Islands	*Dugesia (Dugesia) milloti*: Ball, 1974:376 *Dugesia millotti*: Sluys, 1998:276
D. minotauros De Vries, 1984	Greece: Crete	
D. mirabilis De Vries, 1988	Kenya	
D. monomyoda Marcus, 1953	South Africa: Natal	*Dugesia (Dugesia) monomyoda*: Ball, 1974:376

D. myopa De Vries, 1988	Madagascar: Majunga	
D. nannophallus Ball, 1970	Sri Lanka	*Dugesia (Dugesia) nannophallus:* Ball, 1974:376
D. naiadis Sluys, 2013	Greece: Chios	
D. nanshae De Vries, 1988	Iraq	
D. neumanni (Neppi, 1904)	Ethiopia	*Planaria neumanni* n. sp. Neppi 1904:309 *Euplanaria (Euplanaria) neumanni*: Kenk, 1930:292 *Dugeusia neumanni*: Hughes, 1966:409 [spelling error] *Dugesia (Dugesia) neumanni*: Ball, 1974:376
D. notogaea Sluys & Kawakatsu, 1998	Australia	
D. novaguineana Kawakatsu, 1976	New Guinea: Taleo territory	
D. parasagitta Sluys & Solà, 2013	Greece: Corfu	
D. praecaucasica (Porfirjeva, 1958)	Caucasus	*Dugesia gonocephala praecaucasica* n. subsp. Porfirjeva, 1958:71 *Dugesia (Dugesia) praecaucasica*: Ball, 1974:376
D. ryukyuensis Kawakatsu, 1976	China; Ryukyu Islands; Taiwan	
D. sagitta (Schmidt, 1861)	Greece: Corfu	
D. siamana Kawakatsu, 1980	Thailand	
D. sicula Lepori, 1948	Morocco; Algeria; Tunisia; Greece; Sicily; Elba	*Dugesia (Dugesia) sicula*: Ball, 1974:376
D. subtentaculata (Draparnaud, 1801)	Morocco; Algeria; Iberian Peninsula; France	*Planaria subtentaculata* n. sp. Draparnaud, 1801:100 *Euplanaria subtentaculata*: Komárek, 1926:9 *Euplanaria gonocephala*; de Beauchamp, 1932:221 (in part) *Euplanaria gonocephala subtentaculata*: Stanković, 1934:182 *Dugesia gonocephala subtentaculata*: Angelier, 1959:14 *Dugesia iberica* Gourbault & Benazzi, 1979:330
D. sudanica Dahm, 1971	Sudan	
D. superioris Stocchino & Sluys, 2013	Albania	
D. tamilensis Kawakatsu, 1980	India: Madurai	
D. taurocaucasica (Livanov, 1951)	Crimea; Caucasus	*Euplanaria tauro-caucasica* n. sp. Livanov, 1951:112 *Dugesia gonocephala taurocaucasica*: Porfirjeva, 1958:84 *Dugesia (Dugesia) taurocaucasica*: Ball, 1974:376.
D. transcaucasica (Livanov, 1951)	Caucasus region	*Euplanaria transcaucasica* n. sp. Livanov, 1951:113 *Dugesia gonocephala transcaucasica*: Porfirjeva, 1958:80

	Dugesia (Dugesia) transcaucasica: Ball, 1974:376
D. tubqalis Harrath & Sluys, 2012	Morocco
D. uenorum Kawakatsu & Mitchell, 1995	Indonesia: Sulawesi
D. vestibularis De Vries, 1988	Iran

* It could be a synonym of *D. sicula* according to preliminary results not shown in the present thesis.

Annexes 1 – Table 2. List of doubtful *Dugesia* species. It includes their known distribution range and the synonyms proposed up-to-date. Updated from Kenk (1974). All the references are available at http://continenticola.lifedesks.org.

Species	Distribution	Synonyms
D. aborensis (Whitehouse, 1913)	Himalaya	*Planaria aborensis* n. sp. Whitehouse, 1913;317
D. absoloni (Komárek, 1919)	Balkan Peninsula	*Geopaludicolia absoloni* n. sp. Komárek, 1919:32 *Planaria (Geopaludicolia) absoloni*: Meixner, 1928:574 *Euplanaria (Geopaludicolia) absoloni*: Kenk, 1930:292 *Dugesia absoloni*: Codreanu, 1961:165
D. barroisi (Whitehouse, 1914)	Southwest Asia	*Planaria barroisi* Whitehouse, 1914 *Dugesia(?) barroisi*: Ball, 1974:376
D. bilineata (Kaburaki, 1918)	Burma/Myanmar	*Planaria bilineata* Kaburaki, 1918:193 *Dugesia(?) bilineata*: Kawakatsu, 1969:214
D. brachycephala (Bömig, 1897)	East Africa	*Planaria brachycephala* Böhmig, 1897 *Dugesia? brachycephala*: Ball, 1974:378
D. cinerea (Stimpson, 1857)	Japanese islands	*Planaria cinerea* Stimpson, 1857
D. rava (Weiss, 1909)	Australia	*Planaria rava* Weiss, 1909 *Dugesia(?) rava*: Ball, 1974:378.
D. salina (Whitehouse, 1914)	Southwest Asia	*Planaria salina* Whitehouse, 1914: 461 *Dugesia(?) salina*: Ball, 1974:378
D. seclusa De Beauchamp, 1940	Crozet Islands (Indian Ocean)	*Euplanaria seclusa* de Beauchamp, 1940:313 *Dugesia seclusa*: Lepori, 1948a:188 *Curtisai seclusa*: Westblad, 1952:49 *Euplanaria reclusa*: Jeannel, 1961:137 [spelling error] *Dugesia (Girardia) seclusa*: Ball, 1974:377
D. persica Behzad, Masoud & Mahmood, 1998	Iran	
D. tanganyikae (Leidlaw, 1906)	East Africa: Lake Tanganyika	*Planaria tanganyikae* Laidlaw, 1907:777 *Dugesia(?) tanganyikae*: Ball, 1974:378
D. tiberiensis (Whitehouse, 1914)	Southwest Asia	*Planaria tiberiensis* Whitehouse, 1914:459 *Planaria tiberiadis*: Por, 1968:59 [spelling error] *Dugesia(?) tiberiensis*: Ball, 1974:378
D. venusta (Bömig, 1887)	East Africa	*Planaria venusta* Böhmig, 1897:12 *Dugesia(?) venusta*: Ball, 1974:378

Annexes 1 − Table 3. Details on *D. arabica* and *D. aethiopica* used in the Discussion section 1.4.1.1.

Code	Species	Locality	Reproduction in field	Collector	Coordenates
Darb1	*D. arabica*	Ibb-Alsaradeh	Asexual	Halim Harrath	13°55'17.00"N 44° 7'25.63"E
Darb2	*D. arabica*	Ibb-Aljarana	Asexual	Halim Harrath	13°51'19.12"N 44° 2'18.23"E
Darb3	*D. arabica*	Ibb-Hadeed	Sexual (cocoons)	Halim Harrath	13°54'14.79"N 44° 5'20.03"E
Darb4	*D. arabica*	Ibb-Aljannet	Asexual	Halim Harrath	14° 1'58.79"N 44°11'31.32"E
Darb5	*D. arabica*	Otma	Sexual and asexual	Halim Harrath	14°27'37.02"N 43°54'6.51"E
Darb6	*D. arabica*	Ibb-Albustane	Sexual (cocoons)	Halim Harrath	13°56'3.75"N 44° 9'19.31"E
Daet1	*D. aethiopica*	Lake Tana, Ethiopia	Asexual	Renata Manconi	11°35'60.00"N 37°22'60.00"E
Daet2	*D. aethiopica*	Lake Tana, Ethiopia	Asexual	Renata Manconi	11°35'60.00"N 37°22'60.00"E

2

Annex 2 − Other publications

Int. J. Dev. Biol. 56: 5-17
doi: 10.1387/ijdb.113441mr

THE INTERNATIONAL JOURNAL OF DEVELOPMENTAL BIOLOGY
www.intjdevbiol.com

Evolutionary history of the Tricladida and the Platyhelminthes: an up-to-date phylogenetic and systematic account

MARTA RIUTORT*[,1], MARTA ÁLVAREZ-PRESAS[1], EVA LÁZARO[1], EDUARD SOLÀ[1] and JORDI PAPS[2]

[1]*Institut de Recerca de la Biodiversitat (IRBio) i Departament de Genètica, Universistat de Barcelona, Spain and*
[2]*Department of Zoology, University of Oxford, UK*

ABSTRACT Within the free-living platyhelminths, the triclads, or planarians, are the best-known group, largely as a result of long-standing and intensive research on regeneration, pattern formation and *Hox* gene expression. However, the group's evolutionary history has been long debated, with controversies ranging from their phyletic structure and position within the Metazoa to the relationships among species within the Tricladida. Over the the last decade, with the advent of molecular phylogenies, some of these issues have begun to be resolved. Here, we present an up-to-date summary of the main phylogenetic changes and novelties with some comments on their evolutionary implications. The phylum has been split into two groups, and the position of the main group (the Rhabdithophora and the Catenulida), close to the Annelida and the Mollusca within the Lophotrochozoa, is now clear. Their internal relationships, although not totally resolved, have been clarified. Tricladida systematics has also experienced a revolution since the implementation of molecular data. The terrestrial planarians have been demonstrated to have emerged from one of the freshwater families, giving a different view of their evolution and greatly altering their classification. The use of molecular data is also facilitating the identification of Tricladida species by DNA barcoding, allowing better knowledge of their distribution and genetic diversity. Finally, molecular phylogenetic and phylogeographical analyses, taking advantage of recent data, are beginning to give a clear picture of the recent history of the *Dugesia* and *Schmidtea* species in the Mediterranean.

KEY WORDS: *Metazoa, molecular phylogeny, Tricladida, Platyhelminthes, systematic*

Introduction

The Tricladida belong to the phylum Platyhelminthes, a phylum best known by their parasitic representatives and characterised by a general morphological simplicity. This simplicity has, from the very beginning, made the taxonomist's work difficult. Proof of this is the fact that the two first planarian genera, *Planaria* Müller, 1776 and *Polycelis* Ehrenberg, 1831, included many species not belonging to Tricladida (cf. Kenk 1974). The genus *Planaria*, which is now a valid genus of the Continenticola, was originally established by Müller (1776) to encompass all free-living lower worms. Thus, many species originally described as species of *Planaria* were later placed into other orders and suborders of Turbellaria or into the phylum Nemertina or Rhynchocoela. Similarly, the generic name *Polycelis*, introduced by Ehrenberg (1831), was originally applied by Diesing (1850) to all many-eyed turbellarians, which included terrestrial and polyclad turbellarians in addition to some freshwater planarians. The appearance of modern methodologies to define diagnostic, phylogenetic and systematic characters have gradually helped to partially solve this problem. Improved microscopy tools and staining procedures, the advent of electron microscopy, and more recently, the possibility of using monoclonal antibodies and confocal microscopy (unfortunately still in its infancy in Platyhelminthes studies) facilitate the discovery of defining features. At the same time, the use of molecular data to infer phylogenies has been crucial for understanding the origin and evolution of many Platyhelminthes features. Finally, molecular data are a key tool for understanding species' origins and demography, allowing the use of planarians as model organisms in the study of the origin and maintenance of biodiversity.

In this review we will use a top-down approach, beginning by revisiting the position of the phylum within the metazoans and re-

***Address correspondence to:** Marta Riutort. Dpt. de Genètica, Facultat de Biologia, Universistat de Barcelona, Avda. Diagonal, 643, 08028 Barcelona, España. Tel: +34-93-403-5432. e-mail: mriutort@ub.edu - web: www.ub.edu/genetica/evo-devo/riutort.htm

Final, author-corrected PDF published online: 16 March 2012

ISSN: Online 1696-3547, Print 0214-6282

viewing the historical importance of flatworms in understanding the origins of bilaterians and their complexity. Finally, we will discuss recent advances in the phylogeography of particular groups and how these new data are of interest for both planarian scientists and biodiversity researchers.

The Platyhelminthes

Morphology: from Gegenbaur to Ehlers

The phylum Platyhelminthes includes more than 20,000 species and is the fourth-largest animal phylum after arthropods, molluscs and chordates (Ruppert *et al.,* 2003). In addition, platyhelminths have played a key role in hypotheses regarding bauplan evolution, particularly the origin of bilateral symmetry, since the advent of evolutionary theory. They were originally named by German zoologist Karl Gegenbaur (1859), teacher and coworker of Ernst Haeckel, the same year that *The Origin of Species* was published. Their name is composed of the two Greek words *platy*, meaning "flat", and *helminth* meaning worm; thus, it is a direct translation of their vernacular name, "flatworms".

In Gegenbaur's time, the Animal Kingdom had already been divided into two major groups, the diploblastic animals and the triploblastic metazoans. The Diploblastica include sponges, cnidarians, ctenophores and placozoans; diploblasts have 2 embryonic layers (ectoderm and mesoderm) and have previously been referred to as the Radiata (due to their radial body symmetry) or the Coelenterata (although later on this term was restricted only to cnidarians and ctenophores). The Triploblastica, or Bilateria, include all other animals, which have a third developmental layer (mesoderm) and exhibit bilateral symmetry. The Platyhelminthes are bilaterians that are often described as having an austere architecture due to the absence of traits found in most bilaterians. Of particular note is the lack of coelom; however, other widespread characters are also missing in flatworms, such as the anus, mitosis in somatic cells and circulatory and respiratory systems. In contrast, many flatworms exhibit spiral embryonic cleavage, a type of development associated with several "complex" invertebrate phyla, such as annelids or molluscs. Because of this mix of simple and complex features, they have often been considered candidate representatives of the transition from diploblasts to triploblasts.

In the English translation of Gegenbaur's work (Gegenbaur *et al.,* 1878), the platyhelminths were included within the metazoan division "Vermes" together with annelids, or "nemathelmintes" (nematodes and nematomorphs), among many other worms. Gegenbaur split Platyheminthes into four groups: the Turbellaria (Rhabdocoela and Dendrocoela, the latter including genera such as *Planaria* and *Leptoplana*), the Trematoda, the Cestoda and the Nemertina. Interestingly, the latter group corresponds to the contemporary phylum nemertines, which were considered platyhelminths at that time and were subsequently often linked to flatworms; this is because both of them were thought to be acoelomates, but more than 100 years later, the rhynchocoel of nemertines was proven to be a derived coelom (see a recent comprenhensive review on nemertines in Turbeville 2002) . With regard to the coelom, Gegenbaur (1878) wrote the following: "*In a large number of Vermes this perienteric space (Coelom) is either altogether absent, or only rudimentarily present. This is the case in most of the Platyhelminthes and Nemathelminthes..*" He later states, "*In the land Planarians two cavities traversed by a reticulum of connective tissue extend along the body; they are largely broken up anteriorly. They are to be regarded as indications of a coelom of this kind*". Despite including platyhelminthes in the Coelomata category, Gegenbaur's tree of the animals still placed flatworms (vermes) as the intermediate group representing the transition from diploblasts to triploblasts (Fig. 1A).

Haeckel's view of platyhelminths (or "Platodes", as he sometimes called them) is similar to Gegenbaur's in some respects. For instance, both use the same phylum subdivisions (Turbellaria, Trematoda and Cestoda) and include flatworms within "Vermes", along with the nemertines and onychophorans, or leeches (cf. Haeckel 1866). Regarding the coelom, Haeckel (vol II, p. 148, (1876)) differentiates himself from Gegenbaur by considering flatworms to be acoelomates: "*For all the lower Worms which are comprised in the class of Flat-worms (Platyhelminthes), (the*

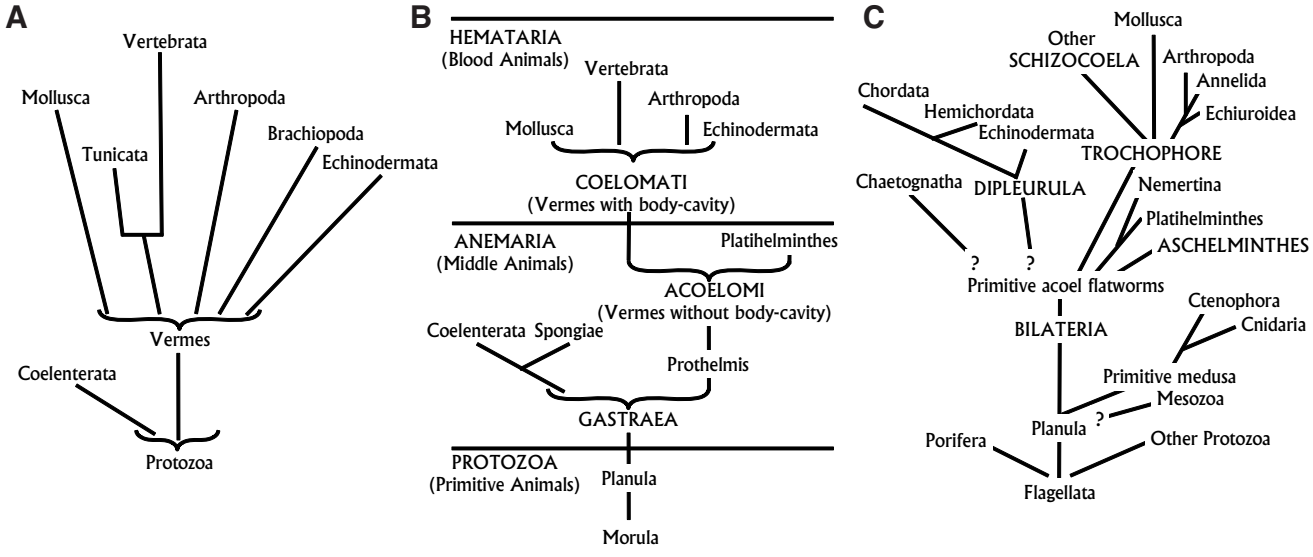


Fig. 1. Different views on the position of Platyhelminthes in the animal kingdom, based on morphology. (A) *Gegenbaur (p.70, 1878).* **(B)** *Haeckel (Haeckel, 1874).* **(C)** *Hyman (1940).*

Gliding-worms, Sucker-worms, Tape-worms), differ very strikingly from other Worms, in the fact that they possess neither blood nor body-cavity (no coelome); they are, therefore, called Acoelomi (..) But all other Worms (like the four higher tribes of animals) possess a genuine body-cavity and a vascular system connected with it, which is filled with blood; hence we class them together as Coelomati." Based on his famous Gastraea theory, Haeckel places flatworms in a critical position in the animal tree, deriving an ancestral flatworm-like animal from the ancestral gastraea (see Prothelmis, also called Archelminthes, in Fig. 1 B): "*The main division of Bloodless Worms (Acoelomi) contains, according to our phylogenetic views, besides the still living Flat-worms, the unknown and extinct primary forms of the whole tribe of Worms, which we shall call the Primaeval Worms (Archelminthes) (..) that may be directly derived from the Gastrea.*".

Haeckel was not the only pre-cladistic zoologist to use life stages in his phylogenetic hypotheses. Many sources attribute the division of bilaterians into three groups based on presence, absence or type of body cavity to American zoologist Henrietta Libbie Hyman: acoelomates (no cavity), pseudocoelomates (cavity derived from the early blastocoel) and coelomates (cavity appears later in development and is limited by an epithelium). Moreover, Hyman is often credited with proposing acoelomates as the first evolutionary offshoot from bilaterians, followed by pseudocoelomates as a sister group to coelomates, which supports a trend of increasing complexity in evolution. In fact, Hyman used this tripartite division of bilaterians to structure her *magnum opus*, *The Invertebrates* (Hyman 1940), but this organisation was more pedagogical than grounded in her ideas on animal evolution (Garey 2002). Instead, her views on metazoan phylogenetics were mostly driven by larval stages, following the planuloid–acoeloid hypothesis of Von Graff (Graff 1882). She derived both the Radiata (cnidarians and ctenophorans) and the Bilateria from a planula-like organism, with bilaterians originating from an acoel flatworm that gave rise to the Protostomia and the Deuterostomia (Fig. 1C). Hyman, who was very fond of the Turbellaria, disliked the term "Vermes" (Hyman 1940, p32): "[the group Vermes] *can only be defined in general and mostly negative terms (i.e., as worm-like animals without skeleton or jointed appendages) and which unites animals of remote and indeterminable relationship while separating groups admittedly closely allied (..) is futile and confusing*". She recognised the acoelomate status of platyhelminthes and separated them from nemertines due to the presence of an anus in the latter.

A prediction of the above hypotheses is that the flatworm bodyplan should be simple and ancestral to bilaterians. This point of view is in conflict with later proposals, primarily based on cladistic principles, which supported the idea of platyhelminths as derived protostomates whose bauplan has been secondarily simplified. Cladistic studies proliferated in the second half of the past century, but far from resolving the question on animal evolution, the new unified criteria for inferring phylogenetic trees produced a myriad of trees and did not resolve outstanding questions (a review of them can be read in Valentine 2004). Those years saw a parade of possible sister groups to Platyhelminthes: the Gnathostomulida, the Nemertea, the Gnathifera, or even the annelids and molluscs, based on their shared spiral cleavage. This instability was probably caused by the fact that most of the traits used in those studies are now recognised as symplesiomorphies or homoplasies (Baguñà, Riutort 2004).

At the same time, the monophyly of the group was under debate. Due to the lack of synapomorphies, the platyhelminths were divided in three groups: the Acoelomorpha (Acoela and Nemertodermatida), the Catenulida and the Rhabditophora (Smith *et al.,* 1986). Later, molecular data confirmed the lack of monophyly, suggesting acoelomorphs as the first offshoot of the bilaterian stem (Carranza *et al.,* 1997; Ruiz-Trillo *et al.,* 1999), whereas the Rhabditophora and the Catenulida constituted a monophyletic group. During those years, the Platyhelminthes also acquired a new member: the newly described Xenoturbella. Its status as flatworm, however, was controversial (see a review in Nielsen 2010). Molecular data first rejected its classification as a flatworm, but recent analyses have linked Xenoturbella to the acoelomorphs. While the wandering of acoelomorphs and xenoturbellids across the evolutionary tree of the animal kingdom is an interesting story, and most likely one far from ending, it is outside of the scope of this paper; therefore, from here onwards, we will use the term Platyhelminthes to refer to the Catenulida and the Rhabditophora (Baguñà, Riutort 2004). As exemplified by acoelomorphs and xenoturbellids, the uncertainties about the position of platyhelminths raised by morphology were replaced by new ones as the field of Systematics entered the molecular age.

Molecules: from one to many genes

The history of metazoan molecular phylogeny can be divided into three main stages: the beginning, using 18S ribosomal RNA gene (18S) sequence, followed by a short multigenic period and the current phylogenomic era. The pioneering work of Field and collaborators (1988) joined metazoan phylogenetics and molecular biology through 18S sequencing. The newly born field would encounter hurdles during its first years, primarily due to a low sampling of phyla and problems such as Long Branch Attraction (LBA), among others (Abouheif *et al.,* 1998). However, as the years passed, the sampling coverage was increased to a great extent, and innovative methods and evolutionary models were developed to deal with systematic and stochastic errors. In the second half of the 1990s, two papers transformed our view of metazoan evolution, providing what was called the New Animal Phylogeny (see a review in Halanych 2004). In this new evolutionary tree of the animals, the Bilateria were divided into three superclades: the Lophotrochozoa (comprising platyhelminths, lophophorates, annelids and molluscs, among many other phyla), the Ecdysozoa (embracing arthropods, nematodes and other traditional pseucoelomate worms) and the Deuterostomia.

The Platyhelminths would enter the molecular age represented by a single 18S sequence from the tricladid *Girardia tigrina* (then *Dugesia tigrina*) (Field *et al.,* 1988). This first flatworm sequence was placed at the first bilaterian split, a position matching the hypotheses posed by Hyman and her precursors. The first broad phylogeny for the platyhelminths using 18S sequences was described in the study of Carranza and collaborators (Carranza *et al.,* 1997), which indicated the separation of acoelomorphs—and, surprisingly, the catenulids as well—from the rest of the phylum; the main novelty in this study was that for the first time, the bulk of Platyhelminthes were placed within the Protostomates. Later works would recover the catenulids as a sister group to rhabditophorans within the lophotrochozoans and confirm the divorce of acoelomorphs and the platyhelminths (Jondelius *et al.,* 2002; Ruiz-Trillo *et al.,* 1999).

The second age of molecular phylogenetics would add more markers to the standard 18S analysis, first the 28S sequence (Mallatt *et al.,* 2010; Mallatt, Winchell 2002; Medina *et al.,* 2001; Paps *et al.,* 2009b; Telford *et al.,* 2003) and later protein coding genes, analysed alone or concatenated. However, the combination of the two ribosomal genes did not overcome the inference artefacts of 18S alone (Mallatt *et al.,* 2010; Paps *et al.,* 2009b), and the signal from protein coding genes when analysed individually lacked statistical support, with two notable exceptions: the alpha subunit of the sodium-potassium ATPase (Anderson *et al.,* 2004) and the myosin heavy chain type II (Ruiz-Trillo *et al.,* 2002). Finally, before systematics entered the high-throughput era, a handful of works used concatenated alignments (from 7 to 23 genes) for a significant number of phyla (Bourlat *et al.,* 2008; Paps *et al.,* 2009a; Sperling *et al.,* 2009). All these works placed platyhelminths within the lophotrochozoans once again, although they were variably positioned and had varying degrees of support.

The access to sequencing facilities and the decreasing cost of high-throughput sequencing has made a great quantity of partial genomic data available. This has resulted in new challenges and approaches to deal with the vast amount of information produced by these methods. Nevertheless, it has been shown that the perils of phylogenetic inference (i.e. LBA) are not completely removed by the use of more characters. Phylogenomic studies of metazoans have culminated in two major sequencing efforts, represented by the studies of Dunn and collaborators (Dunn *et al.,* 2008) and Hejnol and collaborators (2009). While both studies analysed a large number of markers (150 and 1,500 genes, respectively), some of the inter-phyla relationships are weakly supported. Re-analysis of this data with filtered markers and alternative evolutionary models obtained stronger statistical support (Philippe *et al.,* 2011). Again, all analyses place the platyhelminths among the lophotrochozoans, but their position within this group and, hence their evolutionary history, remains elusive.

Platyhelminthes in the 21st Century: simple or simplified?

The position of the Platyhelminthes within lophotrochozoans is vital for understanding the origins of their body plan in the context of animal evolution. Despite the lack of resolution of most molecular trees, the position of the platyhelminths can be summarised in two possible scenarios (Fig. 2). The first is the Platyzoa hypothesis (Giribet *et al.,* 2000), which suggests that the platyhelminths form a group (the Platyzoa) together with the gastrotrichs and the Gnathifera (rotifers, gnathostomulans and cycliophorans, among others). The platyzoans would be defined as ciliated, non-segmented acoelomates or pseudocoelomates that lack a vascular system and have a straight gut (when present), with or without an anus (Cavalier-Smith 1998). This platyzoan clade would be a sister group to the Spiralia, the coelomated lophotrochozoans with spiral cleavage and trochophora larva (i.e., annelids, molluscs, nemertines). The second scenario would place the Platyhelminthes alone as a sister group to the Spiralia, deeply nested within the lophotrochozoans and splitting off after an extensive ladder of many platyzoan and lophophorate phyla (Paps *et al.,* 2009b; 2009a).

The Platyzoa hypothesis divides the lophotrochozoans into two branches, one with all of the "simple" phyla together and the other

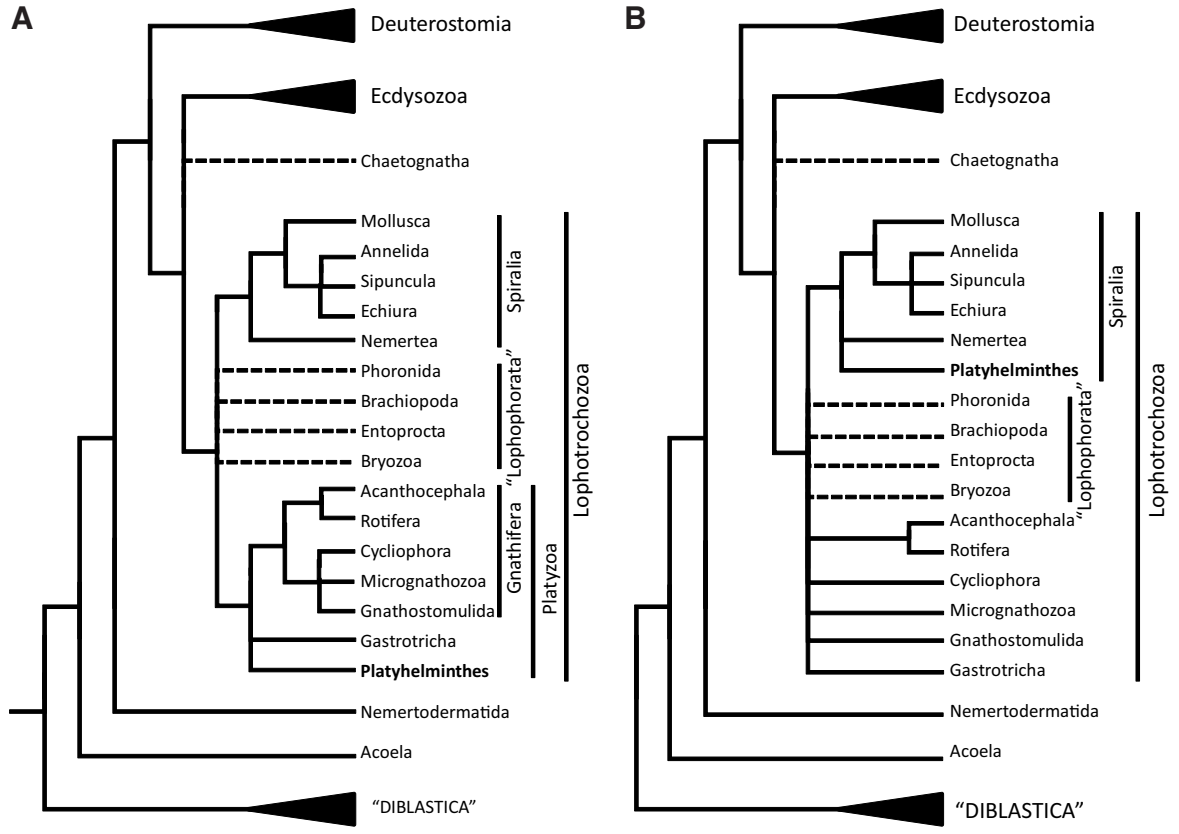


Fig. 2. The positions of Platyhelminthes based on molecular data. (A) *Platyzoa hypothesis.* **(B)** *The Platyhelminths sister group to Spiralia nested within the Lophotrochozoa.*

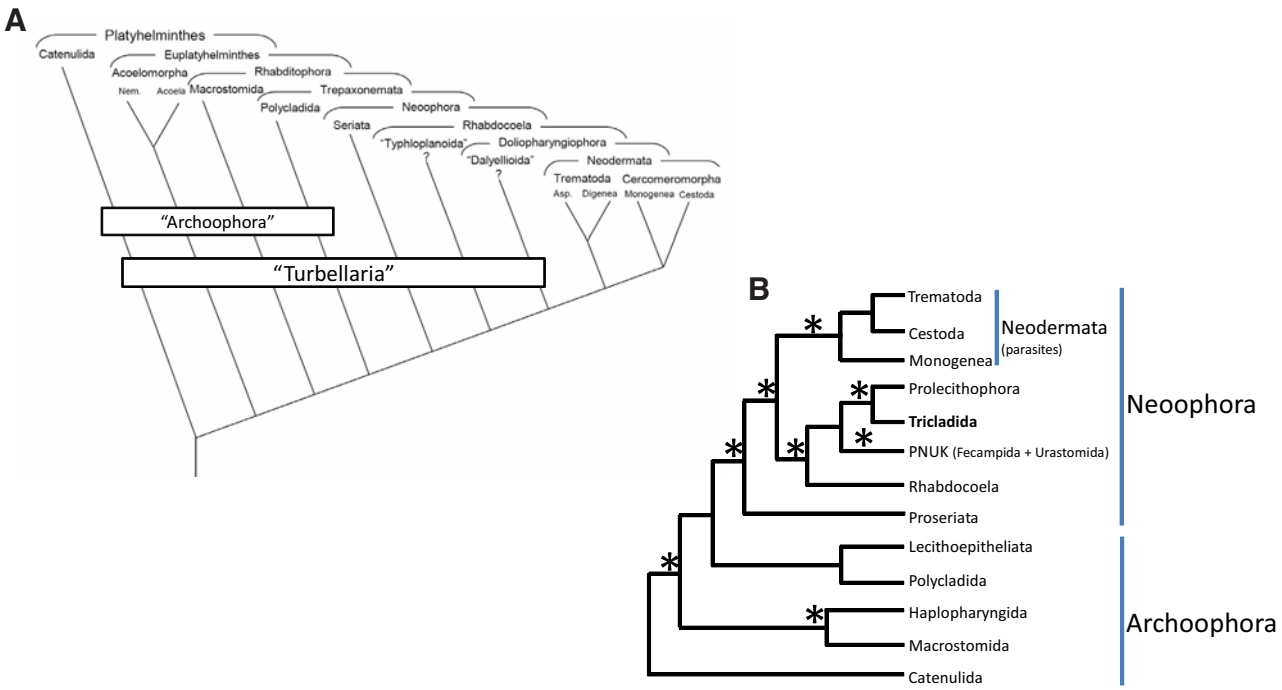


Fig. 3. Internal relationships of the Platyhelminthes. (A) *Ehlers (1985) scheme.* **(B)** *Tree summarising the relationships obtained in different studies based on ribosomal genes. * Strongly supported nodes.*

leading to more "complex" animals; this scenario leaves open the question of the last common ancestor of the Lophotrochozoa, as it could either be a complex animal that underwent simplification in the platyzoan lineage or the opposite. Therefore, without knowing the ancestral level of complexity, we cannot conclude whether flatworms are simple or simplified. The second hypothesis, with flatworms branching deep inside relatively more complex lophotrochozoans (i.e. lophophorates) and next to the Spiralia, points to a simplification process of the phylum; however, it is still possible that flatworms were originally simple and there was a parallel increase of complexity in the groups surrounding them.

Internal relationships within the Platyhelminthes

Historically, the Platyhelminthes were divided in three classes, the free-living class "Turbellaria" and two parasitic classes, the Cestoda and the Trematoda (since further divided into three Classes collectively named Neodermata: the Trematoda, the Cestoda, and the Monogenea) (Gegenbaur 1859; Haeckel 1866; Hyman 1951). The class Turbellaria was subsequently divided into a series of orders (~11, depending on the authors). Comprehensive morphological analyses, however, showed that the parasitic groups evolved from free-living platyhelminths and that the "Turbellaria" thus constituted a paraphyletic group (Ehlers 1985), a situation indicated by the quotes around the name (Fig. 3). The earliest rigorous morphological study of the group is by Karling (1974), while Ehlers (1985) (Fig. 3A) conducted the first cladistic analysis, which has defined the structure of the Platyhelminthes' internal relationships for a long time. However, later morphological studies have refined some groupings (Smith *et al.,* 1986, Haszprunar 1996a,b; and Littlewood *et al.,* 1999b, this last including molecular data). These studies agreed in most regions of the tree, recognising three monophyletic groups within the flatworms: the Acoelomorpha (Acoela and Nemertodermatida), the Catenulida, and the Rhabditophora (the largest group, comprising approximately eight free-living orders and the three classes of parasites) (Table 1). Within the Rhabditophora, the turbellarian orders were divided into the archoophorans (with homocellular female gonads, entolecithal eggs, and cannonical spiral cleavage) and the neoophorans (with heterocellular female gonads, i.e., with separate germaria and yolk glands, and ectolecithal embryos), the archoophorans being paraphyletic (Fig. 3). However, doubts about the relationships among the three major

TABLE 1

MAIN GROUPS AND ORDERS CLASSICALY CONSIDERED WITHIN THE PLATYHELMINTHES

Class "Turbellaria"
Acoelomorpha
Order Acoela
Order Nemertodermatida
Order Catenulida
Rhabditophora
Order Macrostomida
Order Haplopharyngida
Order Lecytoepiteliata
Order Polycladida
Order Proseriata
Order Rhabdocoela
Order Tricladida
Order Prolecitophora
Class Trematoda
Class Cestoda
Class Monogenea

See Figure 3 for their phylogenetic relationships; some of them do not have a taxonomic rank. As of today, Acoelomorpha (Acoela and Nemertodermatida) do not belong to the Platyhelminthes.

groups (the Acoelomorpha, the Catenulida and the Rhabditophora) and even the monophyly of the whole phylum persisted (Smith *et al.,* 1986). As explained above, the use of molecular data demonstrated the phylum to be polyphyletic, and thus, their description to be in need of reconsideration. The phylum was therefore redefined as a monophyletic group consisting of the Catenulida and the Rhabditophora (Baguñà, Riutort 2004) the latter including all free-living orders (except Catenulida) and the three parasitic groups (Cestoda, Trematoda and Monogenea). This organisation poses a taxonomic (nomenclatural) problem for the group, as most books still give a class rank to the three parasitic groups, as well as to the disused "Turbellaria", while their free-living sister clades have the rank of order. This results in a strange taxonomical arrangement and points to the need for a revision of the taxonomy of the whole group to reflect our current understanding of their relationships.

Molecular data have also been used to assess the relationships among orders and classes within the phylum, and these studies have been extensively reviewed in Littlewood and Olson (2001) and Baguñà and Riutort (2004). As in other cases, the first works suffered from poor sampling and inadequate knowledge of the limitations of molecular phylogenetic inference, which led to incorrect conclusions in many cases. The development of more sophisticated analytical methods and more thorough sampling has produced better-resolved phylogenies. Unfortunately, most of these works were led by researchers with interests in particular groups that, together with the difficulty in sampling most orders, has led to biased representations of orders within the phylogenies. Moreover, interest in solving the complete flatworm phylogeny disappeared as soon as these workers were able to establish the closest relative to their groups of interest, leaving the tree still full of uncertainties (Fig. 3B). All of these projects were based on ribosomal genes, and no attempt has been made to use multiple markers. Perhaps the reduced cost and increased sensitivity of new high-throughput methodologies, which enable whole-genome sequencing from individual small organisms, will encourage researchers to fully resolve the Platyhelminthes tree. Despite this pessimistic scenario, a general picture of the main relationships has emerged based on multiple molecular studies (Fig. 3B) and can be summarised as follows:

1. The Catenulida are the most basal group within the Platyhelminthes (sensu (Baguñà, Riutort 2004).

2. Within the Rhabditophora, the orders Macrostomida, Haplopharyngida, Polycladida and Lecithoepitheliata are basal groups with uncertain relationships among them but a clear sistergroup relationship to the rest of the Rhabditophora; this gives support to the derived status of the neoophoran gonad and, hence, to the monophyly of the group.

3. Within the Neoophora, the order Seriata, which included the infraorders Tricladida and Proseriata, is eliminated. The Proseriata are now basal within the Neoophora (also pointed out by Rohde (1990) on morphological grounds), whereas the Tricladida have moved to a more derived position within the tree.

4. In all molecular studies, the sister group of the Tricladida is the Prolecitophora, another group that Ehlers situated basally within the Neoophora. This tight molecular relationship was never suggested at the morphological level and seems not to have any morphological synapomorphy to support it.

5. There is no doubt about the monophyly of the Neodermata (Trematoda, Cestoda, Monogenea), which implies that obligate parasitism, present in all its members, evolved only once. However, a few other species within the free-living lineages are also parasitic or commensal (some groups within the Fecampiida and Urastomidae). Many of these groups share features similar to those present in the Neodermata, such as sperm morphology and a considerable reduction of internal organs. This led several authors to propose that the sister group of the Neodermata would consist of one or some combination of these groups, assuming a relatively recent origin for the Neodermata. However, molecular analyses contradicted these hypotheses.

6. The parasitic and commensal species belonging to the Fecampida and the Urastomidae (*Piscinquilinus*, *Notentera*, *Urastoma* and *Kronborgia* (PNUK) (Littlewood *et al.,* 1999a, b) are grouped in a cluster with the Rhabdocoela, together forming the sister group of the Tricladida + the Prolecithophora clade. This situation leaves the Neodermata with no close "turbellarian" sistergroup and contradicts its Rhabdocoela membership, implying a more ancient origin than that proposed with morphological data and that some of the characters shared by the "turbellarian" parasites and the Neodermata evolved convergently, probably as an adaptation to their parasitic life history.

7. The clade including the Tricladida + the Prolecitophora, the Rhabdocoela and PNUK is well supported by molecular data, but the internal relationships among them are not well resolved. Noren and Jondelius (Noren, Jondelius 2002) found weak support for a clade constituted by the Tricladida + the Prolecitophora and PNUK

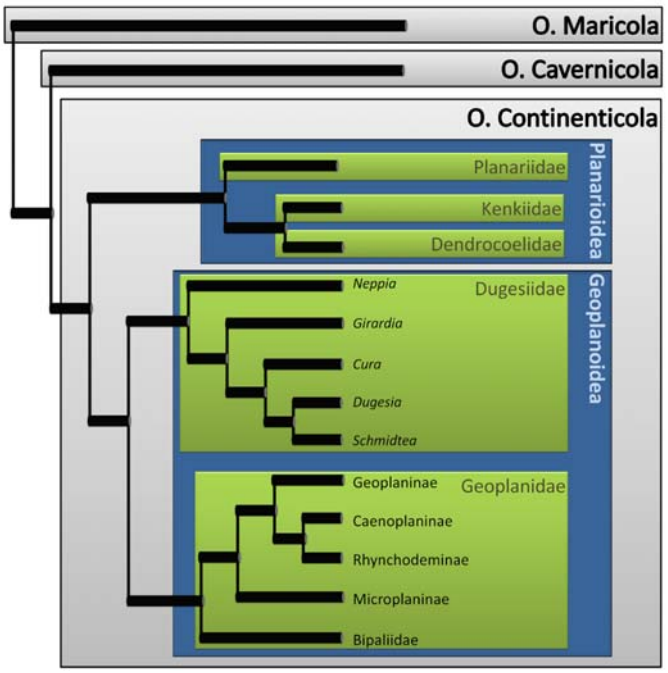


Fig. 4. Phylogenetic tree summarising the current understanding of the internal relationships of the Tricladida. *The tree is primarily based on molecular data, although the Kenkidae and Cavernicola relationship is based exclusively on morphology. The Dugesidae genera* Spathula, Romankenkius *and* Reynoldsonia *have not been included for the sake of clarity (they would be sister group to Microplaninae within Geoplanidae, see text); the genera* Bopsula, Eviella *and* Weissius *have not been included because they have never been subjected to molecular analyses, and there is no clear position for them on morphological grounds.*

(PNUK was named Fecampida in that work); they proposed a name for this clade (the Adiaphanida) based on the only identified feature shared by the three groups: most of the species in this clade have more or less opaque bodies. However, the lack of support for the group discourages the use of this name.

Phylogenetic knowledge below the order level is even more fragmentary, resulting in complex taxonomies. As discussed above, the morphological simplicity of the group makes systematic assessment difficult. Hopefully, future application of molecular techniques at this level will help to clarify the relationships and taxonomy.

The Tricladida (Lang, 1884)

Triclads occupy a derived position within the Platyhelminthes tree, with a clear sister group relationship to the Prolecitophora and close affinity to the Rhadocoela and the group of parasitic "Turbellaria" PNUK. However, these relationships do not reveal any ancestral characteristics because the group does not seem to share any morphological synapomorphies. The Triclads are characterised by a well-defined digestive cavity, consisting of a single anterior gut that splits to produce two posterior branches (which gives the name to the group). They also share other synapomorphies, such as the crossing over of pharynx muscles, embryological features, the cerebral position of female gonads, the serial arrangement of many nephridiopores and a marginal adhesive zone. Within them we can find the largest free-living rhabditophorans: some terrestrial planarians that can reach 1 meter in length and one described abyssal freshwater planarian from lake Baikal that reaches 30-40 cm in length and 10 cm in width.

Hallez (1894) divided the Tricladida into three ecological groups: the Paludicola (freshwater planarians), the Terricola (land planarians), and the Maricola (marine planarians). This division received a taxonomic rank (that varied between sub- and infraorder) and has been used since by all taxonomists, though a doubt was cast on the phylogenetic validity of these ecological groupings. Sluys (1990) proposed a fourth clade, the Cavernicola, grouping five species (belonging to four genera); four of them had been formerly assigned to the Maricola but with apparent closer affinities to the Paludicola. The systematic and phylogenetic relationships of these groups have been discussed on the basis of morphological and ultrastructural characters by Ball (1981), Sopott-Ehlers (1985), and Sluys (1989a). Within the triclads, Ball followed the division of the Tricladida proposed by Steinböck (1925) and considered the Terricola to be the sister group of a clade consisting of the Maricola and the Paludicola (Haploneura). While the Terricola were well defined by their complex diploneural nervous system, the Haploneura did not show clear synapomorphies. Moreover, no synapomorphies were found for the Maricola, but two presumed synapomorphies defined the Paludicola: their reduced precerebral diverticula and the position of the copulatory bursa anterior to the male copulatory apparatus (probursal condition). Sluys (1989a) presented a new phylogenetic scheme based on a reassessment of morphological characters. New traits were found to support the monophyly of the Terricola, the Maricola and the Paludicola, as well as to suggest a closer relationship between the Terricola and the Paludicola clades, changing the evolutionary scheme proposed by Ball.

A radically different view of the internal relationships of the Tricladida emerged from phylogenetic studies based on sequences of 18S ribosomal genes, showing the Terricola to be a sister group to the freshwater family Dugesiidae; this scenario was confirmed by the finding of a molecular synapomorphy, the presence of a shared 18S gene duplication (Carranza *et al.,* 1998). Therefore, the Paludicola emerged as paraphyletic because their previous sister group Terricola was now nested within them. The taxa Terricola and Paludicola became invalid and were replaced by a new taxon, the Continenticola (Carranza *et al.,* 1998). Later molecular studies (Baguñà *et al.,* 2001; Álvarez-Presas *et al.,* 2008) lent further support to the clustering together of the Terricola and the Dugesiidae and of the Terricola and the Dugesiidae to their sister-group, the Planariidae + the Dendrocoelidae (Planarioidea) (Fig. 4).

Regarding taxonomic ranks, the Tricladida were originally a suborder within the order Seriata (Ehlers 1985), and the groups within it had the rank of infraorder (Maricola, Paludicola, Cavernicola and Terricola). In the new classification of the group, taking into account the new understanding of their phylogenetic relationships (Sluys *et al.,* 2009), the Tricladida have order rank and include three suborders, the Maricola, the Cavernicola and the Continenticola (Fig. 4, Table 2).

The relationships within the Tricladida suborders have received uneven attention. Based on morphological characters, they have been considered in some detail within the Maricola (Sluys 1989b), the Cavernicola (Sluys 1990) and the former Paludicola (Ball 1974; De Vries, Sluys 1991; Sluys 1989a) but not the former Terricola. On the other hand, extensive molecular analyses have been performed on the Continenticola (including both former Paludicola and Terricola), but only a very preliminary study has been done for the Maricola, while the Cavernicola have not been studied.

TABLE 2

NEW CLASSIFICATION OF THE TRICLADIDA (LANG, 1884)

Suborder **Maricola** Hallez, 1892
Superfamily Cercyroidea Böhmig, 1906
Family Centrovarioplanidae Westblad, 1952
Family Cercyridae Böhmig, 1906
Family Meixneridídae Westblad, 1952
Superfamily Bdellouroidea Diesing, 1862
Family Uteriporidae Böhmig, 1906
Subfamily Uteriporinae Böhmig, 1906
Subfamily Ectoplaninae Bresslau, 1933
Family Bdellouridae Diesing, 1862
Subfamily Bdellourinae Diesing, 1862
Subfamily Palombiellinae Sluys, 1989
Superfamily Procerodoidea Diesing, 1862
Family Procerodidae Diesing, 1862
Suborder **Cavernicola** Sluys, 1990
Family Dimarcusidae Mitchell and Kawakatsu, 1972
Suborder **Continenticola** Carranza and al, 1998
Superfamily Planarioidea Stimpson, 1857
Family Planariidae Stimpson, 1857
Family Dendrocoelidae Hallez, 1892
Family Kenkiidae Hyman, 1937
Superfamily Geoplanoidea Stimpson, 1857
Family Dugesiidae Ball, 1974
Family Geoplanidae Stimpson, 1857
Subfamily Bipaliinae Von Graff, 1896
Subfamily Microplaninae Pantin, 1953
Subfamily Rhynchodeminae Von Graff, 1896
Subfamily Geoplaninae Stimpson, 1857

Sluys *et al.,* 2009

At present, the Maricola are divided into three superfamilies and are subdivided into six families (Table 2), although the number and groupings of the families have varied historically (see Sluys 1989b for a detailed account). In his monograph, Sluys (1989b) proposed the first and only cladistic phylogeny for the Maricola. He found synapomorphies giving support to the monophyly of the superfamilies, except for the family Meixneridae (now included within the Cercyroidea), which has uncertain relationships. The Cercyroidea are proposed to be the most basal maricolans, sister to a clade constituted by the Procerodoidea and the Bdellouroidea. The only molecular study (Charbagi-Barbirou *et al.,* 2011) finds, with strong support, the family Procerodidae (superfamily Procerodoidea) to be basal to the rest of the families, while the Cercyroidea occupy a derived position in the tree, contradicting the morphological data. With regard to the remaining relationships, the molecular tree recovers a paraphyletic Bdellouroidea (because it includes the Cercyroidea), within which the monophyly of its two component families (Uteriporidae and Bdellouridae) is not recovered. In fact, there are no morphological synapomorphies giving support to the monophyly of the Uterioporidae, but in the case of the family Bdellouridae, the molecular result is strongly contradicted by morphological data. The systematics of this group remains open and in need of extensive morphological and molecular studies.

Sluys (1990) defined the clade Cavernicola formed by five species (grouped into a single family, the Dimarcusidae) based on three morphological features related to the reproductive apparatus. Concerning their phyletic position within the Tricladida, he found many features inconsistent with its belonging to the Maricola (as initially proposed for four of the five species). Sluys (1990) suggested a closer relationship to the Paludicola than to the Maricola, due to the fact that the Cavernicola share one of the three apomorphies of freshwater triclads. However, at that time, the Terricola and the Paludicola were still considered to be independent sister groups; under the current Continenticola hypothesis, Sluys' proposal of a closer affinity of the Cavernicola to freshwater planarians could point to the inclusion of the Cavernicola within continenticolans. Unfortunately there is no easy way to obtain representatives of this group; hence, the final answer to their systematic status will have to wait.

Continenticola

This suborder combines the former Paludicola and Terricola; hence, to follow its history, we need to revise both groups. Within the Paludicola, Hallez (1894), in his revision of the group, recognised nine genera divided in two families, the Planaridae Stimpson, 1857 (now Planariidae) and the Dendrocoelidae Hallez, 1894, the second differing from the former in the possession of anterior adhesive organs. The first attempt to present a "natural" checklist of the Paludicola was made by Kenk (1930), who arranged those species that were sufficiently well investigated into two families, the Planariidae and the Dendrocoelidae, defined by the arrangement of the inner muscle layers of the pharynx. In the Planariidae, the circular and longitudinal muscles of the inner muscle zone of the pharynx form two separate layers, whereas in the Dencrocoelidae, the circular and longitudinal muscle fibres are intermingled. These internal characters confirmed Hallez's original division based exclusively on external features. The distribution of the genera was the same in both schemes, although many more genera had been described by the time Kenk proposed his revision. In 1974 Kenk produced an index of genera and species of freshwater planarians of the world, which was mainly a nomenclatural account in which he clarified the synonymies of many species and the multiple cases of organisms belonging to other orders of Platyhelminthes—or even to other phyla—that had been included in the Tricladida genera. The same year, Ball (1974) established the family Dugesiidae by extracting several genera from within the Planariidae, defining the new family by its unique eye structure, a multicellular pigment cup with numerous light receptive cells. He also proposed that the new family was the sister group of a clade composed of the other two (Planariidae and Dendrocoelidae) families sharing a common oviduct entering the atrium. Hyman (1937) had previously established a new family that included 3 genera of cave planarians, the Kenkiidae, and later findings included more genera within this family. Nonetheless, several authors did not accept their validity, and in the 1960s, their elimination was proposed (de Beauchamp 1961; Mitchell 1968). Hence, later analyses, such as those of Ball described above, considered them to be a subfamily (Kenkiinae) within the Planariidae. Finally, Kenk ((1975)) proposed that the subfamily Kenkiinae should be upgraded to the family level (Kenkiidae), and a recent detailed morphological study (Sluys, Kawakatsu 2006) showed that this family is more closely related to the Dendrocoelidae than to the Planariidae.

The Terricola were taxonomically divided into three families (Geoplanidae Stimpson, 1857, Bipaliidae Graff, 1896, and Rhynchodemidae Graff, 1896) for which no cladistic study has been undertaken. The Continenticola scenario raised by molecular data has resulted in a major taxonomic reorganisation for the Terricola. As a consequence of their sistergroup relationship to Dugesiidae, terrestrial planarians have seen their rank downgraded to the family level; hence, all its previous families became subfamilies (Table 2), and some subfamilies became tribes. The name selected for the family including all terrestrial planarians is Geoplanidae, as this was the older family designation for the terrestrial planarians (Stimpson, 1857), originally housing all of them. The families Dugesiidae and Geoplanidae have been taxonomically grouped into the Superfamily Geoplanoidea at an equivalent rank to the superfamily Planarioidea (the Planaridae + the Dendrocoelidae).

Again, there is a large asymmetry in the systematic and phylogenetic study of this group. Of the two superfamilies included within the Continenticola, the Planarioidea and the Geoplanoidea, the first has received little attention from a systematic point of view. Within it, only the Dendrocoelidae have been the object of a phylogenetic study based on morphological data (Sluys, Kawakatsu 2006).

Planarioidea *(Stimpson, 1857)*

The Planarioidea, including the families Planariidae, Dendrocoelidae and Kenkiidae, exhibit a Holarctic distribution. The Planariidae were the first paludicolan family established and included all genera known at that time. Sluys and Kawakatsu (2006) considered the Kenkiidae and the Dendrocoelidae to share the adhesive organ (previously interpreted as a convergent character, (de Beauchamp 1961; Mitchell 1968)), constituting a likely synapomorphy for their grouping. These workers also found a series of morphological characters that allowed them to define monophyletic groups within the family Dendrocoelidae. Despite inferring a phylogeny for these monophyletic groups, they do not provide any taxonomic rank for them. It is noteworthy that the Dendrocoelidae have undergone

an impressive radiation in Lake Baikal, from where more than 13 endemic genera have been described.

Geoplanoidea *(Stimpson, 1857)*

This superfamily was originally proposed by Stimpson (1857) to include the two families in which he divided the terrestrial planarians, the Geoplanidae and the Polycladidae (this latter group was later abandoned). Today, this superfamily houses all the freshwater planarians from the family Dugesiidae as well as all the terrestrial planarians (family Geoplanidae) (Sluys *et al.,* 2009). The Geoplanoidea are supported by molecular trees based on the two ribosomal genes and mitochondrial cytochrome oxidase I (COI) (Álvarez-Presas *et al.,* 2008), in addition to the presence of a ribosomal gene cluster duplication. In their initial proposal, Carranza and collaborators (1998) suggested the complex eye found in the Dugesidae and in terrestrial planarians as a possible morphological synapomorphy for the group. In fact, Ball (1981) had already considered the similar eye structure in the dugesiids and the terrestrial planarians as a weakness in his phylogenetic proposals. However, recent studies (Sluys, Kawakatsu 2006) have shown eyes with a similar structure to be present also in the dendrocoelids, thus casting doubt on the validity of this character as a defining feature for the Geoplanoidea. Nonetheless, Falleni and collaborators (2009) have found an ultrastructural character related to the morphology of the female gonad that can be considered as a synapomorphy for the group.

The most recent molecular study of the superfamily (Álvarez-Presas *et al.,* 2008) showed that a single transition occurred from freshwater to the terrestrial habitat (from a common ancestor with the Dugesiidae). The origin of this group is probably more than 100 million years old (Carranza *et al.,* 1999), and it was likely followed by rapid diversification. Fast and ancient radiations are respectively associated with short spans of time to accumulate good phylogenetic information in the molecules and long periods to overwrite it with later changes. Both of these phenomena make it difficult to obtain good support for this part of the history of the group. Nonetheless, the presence of three morphological synapomorphies for the terrestrial planarians (cf. Sluys *et al.,* 2009) further supports their unique origin. However, Álvarez-Presas and collaborators (2008) unexpectedly found that three species of freshwater planarians, belonging to the genera *Romankenkius* and *Spathula* (Dugesiidae), are situated within the clade of terrestrial planarians, implying a return to a freshwater environment from land and, from a systematic point of view, polyphyly of the Dugesiidae and paraphyly of the Geoplanidae. A more detailed morphological analysis and new molecular data are needed to test this hypothesis. Until then, this new situation has not been reflected in the new Tricladida classification, and the current family Dugesiidae includes all of the freshwater planarians from the superfamily Geoplanoidea. The family Geoplanidae is thus composed of only terrestrial species.

Family Geoplanidae

The Geoplanidae is divided into four subfamilies (Bipaliinae, Microplaninae, Rhynchodeminae, Geoplaninae) including over 800 described species, although this number is increasing due to extensive sampling and multiple studies being performed both in South America and in Europe. They have a cosmopolitan distribution (Winsor *et al.,* 1998), but most of the species are found in the southern hemisphere, while the Microplaninae are the subfamily with the most northerly distribution (including Africa and Europe).

There are no studies on the relationships among the subfamilies from a morphological point of view, although some hypotheses on the ancestry of certain groups have been posed. Based on their worldwide distribution, Winsor and collaborators (1998) proposed that the rhynchodemids are the earliest divergent terricolans, while Marcus and Froehlich (cf. Sluys 1989b), using characteristics of the copulatory organ, suggested that the Microplaninae are the earliest divergent terricolan clade. The only molecular study (Álvarez-Presas *et al.,* 2008) gave strong support to a basal position for the family Bipaliidae —never proposed on morphological grounds-and also revealed major problems with the classical taxonomy of the group. The Rhynchodeminae and the Microplaninae (constituting the Family Rhynchodemidae) did not group together; instead, the Rhynchodeminae showed a close relationship to the Caenoplaninae (with the Geoplaninae constituting family Geoplanidae), a situation that has been amended in the new taxonomy (Sluys *et al.,* 2009). Fig. 4 shows the summary tree of those analyses.

Interest in the terrestrial planarians has increased recently as a result of the introduction of non-native predatory species in regions where they have achieved pest status. For example, the New Zealand flatworm, *Arthurdendyus triangulatus*, has invaded the British Isles and continental Europe (Jones, Boag 1996), and other species have been introduced in North America (Dindal 1970). Land planarians have also been used effectively in Hawaii, the Maldives, Indonesia's Irian Jaya, and Guam to biologically control the introduced giant African snail species *Achatina fulica*. Some concern exists, however, that the introduction of these predatory land planarians has resulted in the extinction of some native land snails (Sugiura, Yamamura 2006). In addition, due to their fragility with respect to environmental changes and their predator status, terrestrial planarians have been proposed as excellent invertebrate bioindicators for biodiversity and conservation studies (Sluys 1999; Carbayo *et al.,* 2002) and have been demonstrated to be good models in comparative phylogeography studies over small scales in Australia (Sunnucks *et al.,* 2006) and the Brazilian Atlantic Forest (Álvarez-Presas *et al.,* 2011).

Family Dugesiidae

The family Dugesiidae has received more attention from a systematic point of view. This is probably because its members are among the most easily and frequently found planarians in Europe and North America (genera *Dugesia*, *Schmidtea* and *Girardia*); hence, taxonomists from these continents have dedicated many works to their species. However, the Dugesiidae include many more genera. Ball established the family in 1974 and included 11 genera: *Bopsula*, *Cura*, *Dugesia*, *Eviella*, *Girardia*, *Neppia*, *Reynoldsonia*, *Rhodax*, *Romankenkius*, *Schmidtea*, *Spathula* (Ball 1974, 1977). Later, *Rhodax* was moved to the new taxon Cavernicola (Sluys 1990), and more recently, a new genus, *Weissius,* has been added to the family (Sluys *et al.,* 2007). Ongoing morphological and molecular studies, both in the Mediterranean and in Australia, show that there is yet a broader diversity within this family and that more genera will likely be described in the future.

The family has a worldwide distribution and 3 of the 11 genera are present in the northern hemisphere: *Girardia*, *Dugesia* and *Schmidtea*. The two last genera mentioned originated on Laurasia:

Dugesia is distributed across Europe, Asia and Africa, whereas *Schmidtea* has a nearly exclusive European distribution (some populations have been described from North Africa). *Girardia* originally had an American distribution, although one species (*Girardia tigrina*) was introduced to Europe at the beginning of the 20th century. Similarly, *Schmidtea polychroa* was introduced in America.

The phylogenetic relationships within the family were first examined by Ball (1974), based on various morphological characters, thus producing a very preliminary scheme. It was 30 years before the first phylogeny was produced based on a species-level analysis of a large number of morphological features (Sluys 2001). That study was unable to find unambiguous derived features for some major clades, resulting in some polytomies. However, the analysis supported some conclusions, such as the basal situation of *Spathula* (which included the genera *Reynoldsonia* and *Eviella*, probably merely being aberrant species of *Spathula*). *Romankenkius* and *Neppia* constituted a monophyletic group sister to a clade that includes *Girardia*, *Schmidtea*, *Cura* and *Dugesia* (the last one being the first offshoot of the group). This is in contrast with a recent molecular study of the Continenticola (Álvarez-Presas *et al.,* 2008), in which the genera *Spathula* and *Romankenkius* were not grouped within the Dugesiidae but within the Geoplanidae (*Eviella, Weissius* and *Reynoldsonia* were not included in the analysis). Within the monophyletic Dugesiidae, *Girardia* was most basal, and *Dugesia* and *Schmidtea* constituted a sister clade to *Cura* (Fig. 4).

Schmidtea *(Ball 1974)*

The genus *Schmidtea* was originally known as the *Dugesia lugubris-polychroa* group or *Dugesia lugubris* s.l. (Benazzi 1957; Reynoldson, Bellamy 1970) and later as the subgenus *Dugesia (Schmidtea)* (Ball, 1974). Finally, it was raised to the genus level, together with the other two subgenera of *Dugesia* (*Dugesia (Dugesia)* and *Dugesia (Girardia)*); De Vries, Sluys 1991), based on morphological differences. Later, molecular data supported this new taxonomical status (Riutort *et al.,* 1992). Seven biotypes (named with letters A,B,C,D,E,F and G) were recognized within the genus, differing in their karyotype (chromosome morphology) and ploidy level (cf. Benazzi, Benazzi-Lentati, 1976). Benazzi recognised the existence of 3 species: *D. polychroa* (including biotypes A,B,C and D), *D. lugubris* (E,F) and *D. mediterranea* (biotype G) (Benazzi *et al.,* 1975). Later, the new species *D. nova* was described for biotype F (Benazzi 1982). Within each species, either amphimictic (diploids) or parthenogenetic (polyploids) modes of reproduction can be found. *S. mediterranea* presents a third type of reproduction, fissiparity, in diploid populations presenting a heteromorphic translocation. This type of reproduction is common in other genera of planarians, such as *Dugesia;* however, this is the only known case in *Schmidtea*. While there is no specific study analysing the relationships among the *Schmidtea* species, they have been included in some molecular studies. Such analyses have shown a closer relationship between *S. polychroa* and *S. mediterranea*, whereas the relationship between *S. lugubris* and *S. nova* is not clear due to the lack of good molecular information for the latter (Álvarez-Presas *et al.,* 2008; Lazaro *et al.,* 2011). It is worth noting that the molecular analyses revealed a low level of diversification of the group as compared to other Dugesiidae genera of similar age, particularly *Dugesia*. It is possible that these four species have a recent origin, but molecular trees seem to show an old diversification, which could give support to the idea that they are the remnants of a more widespread diversity (Lazaro *et al.,* 2011).

Dugesia *(Girard, 1850)*

Dugesia is a species-rich genus, in stark contrast with *Schmidtea*. It includes approximately 75 described species with a wide distribution, viz. the Afrotropical, Palearctic, Oriental, and Australian biogeographic regions. Of these 75 species, more than 20 occur in Europe and in the Mediterranean area (cf. Sluys *et al.,* 1998), indicating a wide radiation of the genus in this area. However, several factors render the number and distribution of *Dugesia* species in the Mediterranean uncertain. First, they are externally very similar. Second, many of their populations are triploid and reproduce asexually by fission (fissiparous forms). Such forms do not develop a reproductive system or copulatory apparatus, the only source of diagnostic taxonomic characters, thus making proper species assignment impossible. The net result for *Dugesia* in the Mediterranean is the presence of several sexually reproducing species occurring together with a much larger number of asexual triploid populations that have been known as *Dugesia gonocephala* s.l. or, in more recent studies, as *Dugesia* sp. The broadest phylogenetic study based on morphological data to date (Sluys *et al.,* 1998) found two monophyletic clades, defined by the course of the ejaculatory duct. Unfortunately, the number of characters was insufficient to obtain a fully resolved phylogeny for the whole genus, thus resulting in a highly polytomous phylogenetic tree.

The use of molecular data for species identification in asexual populations and to resolve phylogenies for species in the western Mediterranean has been a successful strategy (Baguñà *et al.,* 1999; Lazaro *et al.,* 2009). These studies have demonstrated that the mitochondrial gene Cytochrome Oxidase I (COI) is an excellent barcoding tool that allows the assignment of the asexual populations to species and at the same time, together with the nuclear marker ITS-1, has resulted in a well-resolved phylogeny. The results of these studies yielded many interesting points:

1. *Dugesia* is divided in two main molecular groups in the western Mediterranean. There are some differences with the groups defined on morphological grounds. One of the groups includes only two species (*D. sicula* and *D. aethiopica*) at present, and all the rest belong to a clade that we will here call the European clade.

2. Most of the triploid asexual populations found in the Mediterranean basin belong to the species *D. sicula*. Although other species have asexual populations, in general, they have an endemic distribution, and only a few populations show that type of reproduction.

3. The *sicula-aethiopica* clade presents two outstanding features: almost molecular identity between the two species and low genetic diversity among *D. sicula* populations geographically distant as Greece, Italy, Tunisia, Spain and the Canary Islands.

4. *D. gonocephala*, the North European representative of the genus, presents low genetic diversity, is buried deeply within the tree and is closely related to some Italian species, suggesting that European colonisation proceeded from South to North.

5. All central European *D. gonocephala* populations studied are grouped into clades, with nearly no genetic diversity within them and with low diversity among them. Nonetheless, they are older than the last glacial maximum, suggesting the presence of various glacial refugia in Central Europe from which the species spread after the Ice Age.

6. *D. subtentaculata*, the only species of the European clade known to be present in the Iberian Peninsula, shows a high genetic

differentiation between the only two sexual populations studied molecularly thus far (both in Mallorca). This differentiation is even stronger when compared to the asexual populations analysed. This suggests a highly structured species or even the existence of more than one species. *D. subtentaculata* has in common with *D. sicula* the fact that most of its populations are asexual but is different in that geographically close populations are highly differentiated. However, a more detailed molecular analysis of the populations present in Spain may reveal the existence of more than one species and change this impression.

7. The rest of the species of the group are primarily endemic to in small continental areas or islands. In Greece, from where 9 endemic species have been described (De Vries 1984), studies in progress (Solà *et al.,* in preparation) indicate that molecular clades coincide with morphologically described species, or else molecular and morphological data point to the existence of new species (Sluys *et al.,* in preparation). Also, the relationships found among the species mostly correlate with the complex geological history of the region, which will allow the calibration of a molecular clock.

Perspectives

In the genomic era, the ease of acquiring massive amounts of molecular data, even from single individuals, is beginning to open new possibilities for systematic studies. In Platyhelminthes in general and planarians in particular, the new era could mean, on the one hand, finally solving the unsettled position of the phylum within the Metazoa. On the other hand, it will make possible the extraction of genetic data at the population level and, hence, allow their use in fine-scale phylogeographical and demographic studies. This will give planarians a role that was previously closed to non-model organisms, enabling their use in studies on the origin and maintenance of biodiversity and its conservation.

Acknowledgements

We would like to acknowledge all people who provided material and data that made possible the work described in this review. We also will like to thank all the collaborators in the different projects that are reviewed here, specially Tim Littlewood, Ronald Sluys, Maria Pala, Giacinta Stocchino, Saïda Tekaya, Abdul Halim Harrath, Fernando Carbayo. We are grateful to the past PhD students of the Molecular phylogeny group, Salvador Carranza and Iñaki Ruiz-Trillo, now with a successful career of their own, for having contributed to the advance of the studies here presented. Special thanks are due to Jaume Baguñà, who initiated this line of research at the end of the 80s when he introduced MR into the world of phylogenetics. None of this would have been possible without his perseverance in arising interesting questions on the evolution of animals, and in trying to find the best and most innovative ways to answer them. This work was funded by grants from the Ministerio de Educación y Ciencia PB90-0477, PB97-0937, BOS2002-02097 to Jaume Baguñà and CGL2005-00371/BOS, CGL2008-00378/BOS to MR, and from the Generalitat de Catalunya Nos. 1999SGR-00026 and 2001SGR-00102 to Jaume Baguñà.

References

ABOUHEIF, E., ZARDOYA, R. and MEYER, A. (1998). Limitations of metazoan 18S rRNA sequence data: implications for reconstructing a phylogeny of the animal kingdom and inferring the reality of the Cambrian explosion. *J Mol Evol* 47: 394-405.

ÁLVAREZ-PRESAS, M., BAGUÑÀ, J. and RIUTORT, M. (2008). Molecular phylogeny of land and freshwater planarians (Tricladida, Platyhelminthes): From freshwater to land and back. *Mol Phylogenet Evol* 47: 555-568.

ÁLVAREZ-PRESAS, M., CARBAYO, F., ROZAS, J. and RIUTORT, M. (2011). Land planarians (Platyhelminthes) as a model organism for fine-scale phylogeographic studies: understanding patterns of biodiversity in the Brazilian Atlantic Forest hotspot. *J Evolution Biol* 24: 887-896.

ANDERSON, F.E., CORDOBA, A.J. and THOLLESSON, M. (2004). Bilaterian phylogeny based on analyses of a region of the sodium-potassium ATPase beta-subunit gene. *J Mol Evol* 58: 252-68.

BAGUÑÀ, J., CARRANZA, S., PALA, M., RIBERA, C., GIRIBET, G., ARNEDO, M.A., RIBAS, M. and RIUTORT, M. (1999). From morphology and karyology to molecules. New methods for taxonomical identification of asexual populations of freshwater planarians. A tribute to Professor Mario Benazzi. *Ital. J. Zool.* 66: 207-214.

BAGUÑÀ, J., CARRANZA, S., PAPS, J., RUIZ-TRILLO, I. and RIUTORT, M. (2001). Molecular taxonomy and phylogeny of Tricladida. In I*nterrelationships of the Platyhelminthes* (Eds. D.T.J. Littlewood and R.D. Bray). Taylor and Francis, London, England, pp: 49-56.

BAGUÑÀ, J. and RIUTORT, M. (2004). Molecular phylogeny of the Platyhelminthes. *Can. J. Zool*, no. 82: 168-193.

BALL, I.R. (1981). The phyletic status of the Paludicola. *Hydrobiologia* 84: 7-12.

BALL, I.R. (1977). On the phylogenetic classi¢cation of aquatic planarians. *Acta Zool. Fenn.* 154: 21-35.

BALL, I.R. (1974). A contribution to the phylogeny and biogeography of the freshwater Triclads (Platyhelminthes: Turbellaria). *Biology of the Turbellaria*, (Eds. N.W. Riser and M.P. Morse). McGraw-Hill, NY.

BENAZZI, M. (1982). Speciation events evidenced in Turbellaria. In *Mechanisms of Speciation* (Ed. C. Barigozzi, Alan R. Liss). New York, pp. 307-344.

BENAZZI, M. (1957). Cariologia di Dugesia lugubris (0. Schmidt) (Tricladida, Paludicola). *Cariologia* 10: 276-303.

BENAZZI, M., BAGUÑÀ, J., BALLESTER, R., PUCCINELLI, I. and DEL PAPA, R. (1975). Further contribution to the taxonomy of the *Dugesia lugubris-polychroa* group- with description of *Dugesia mediterranea* n.sp. (Tricladida, Paludicola). *Boll. Zool.* 42: 81-89.

BENAZZI, M. and BENAZZI-LENTATI, G. (1976). *Animal Cytogenetics. Vol. I Platyhehninthes*, Gebruder Borntraeger, Berlin.

BOURLAT, S.J., NIELSEN, C., ECONOMOU, A.D. and TELFORD, M.J. (2008). Testing the new animal phylogeny: A phylum level molecular analysis of the animal kingdom. *Mol Phylogenet Evol* 49: 23-31.

CARBAYO, F., LEAL-ZANCHET, A.M. and VIEIRA, E.M. (2002). Terrestrial flatworm (Platyhelminthes: Tricladida: Terricola) diversity versus man-induced disturbance in an ombrophilous forest in southern Brazil. *Biodiversity* and *Conservation* 11: 1091-1104.

CARRANZA, S., BAGUÑÀ, J. and RIUTORT, M. (1997). Are the Platyhelminthes a monophyletic primitive group? An assessment using 18S rDNA sequences. *Mol Biol Evol* 14: 485-497.

CARRANZA, S., LITTLEWOOD, D.T.J., CLOUGH, K.A., RUIZ-TRILLO, I., BAGUÑÀ, J. and RIUTORT, M. (1998). A robust molecular phylogeny of the Tricladida (Platyhelminthes: Seriata) with a discussion on morphological synapomorphies *P Roy Soc B-Biol Sci* 265: 631-640.

CARRANZA, S., BAGUÑÀ, J. and RIUTORT, M. (1999). Origin and Evolution of Paralogous rRNA Gene Clusters Within the Flatworm Family Dugesiidae (Platyhelminthes, Tricladida). *J Mol Evol.* 49: 250-259.

CAVALIER-SMITH, T. (1998). A revised six-kingdom system of life. *Biol Rev* 73: 203.

CHARBAGI-BARBIROU, K., ÁLVAREZ-PRESAS, M., VILA, M., GAMMOUDI, M., TEKAYA, S. and RIUTORT, M. (2011). Marine triclads (Platyhelminthes, Tricladida, Maricola). a preliminary molecular approach to their phylogeny. *Cah. Biol. Mar.* 52: 303-311.

DE BEAUCHAMP, P. (1961). Classe des Turbellariés. In *Traité de Zoologie* (Ed. P. Grassé), Masson et Cie, Paris. pp. 35-212.

DE VRIES, E.J. (1984). On the species of the *Dugesia gonocephala* group (Platyhelminthes, Turbellaria, Tricladida) from Greece. *Bijdragen tot de Dierkunde* 54: 101-126.

DE VRIES, E.J. and SLUYS, R. (1991). Phylogenetic relationships of the genus Dugesia (Platyhelminthes, Tricladida, Paludicola). *J Zool (Lond)* 223: 103-116.

DIESING, K.M. (1850). Systema Helminthum, Vindobonae.

DINDAL, D.L. (1970). Feeding behaivor of a terrestrial turbellarian *Bipalium adventitium. Am. Midl. Nat.* 83: 635-637.

DUNN, C.W., HEJNOL, A., MATUS, D.Q., PANG, K., BROWNE, W.E., SMITH, S.A.,

SEAVER, E., ROUSE, G.W., OBST, M., EDGECOMBE, G.D., SORENSEN, M.V., HADDOCK, S.H.D., SCHMIDT-RHAESA, A., OKUSU, A., KRISTENSEN, R.M., WHEELER, W.C., MARTINDALE, M.Q. and GIRIBET, G. (2008). Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* 452: 745-749.

EHLERS, U. (1985). Das Phylogenetische System der Plathelminthes. G. Fischer. Stuttgart and New York.

EHRENBERG, C.G. (1831). Animalia evertebrata exdusis Insectis in Symbolae physicae (Ed. Fredericus Guilelmus Hemprich and Christianus Godofredus Ehrenberg). *Symbolae physicae*. Volume 4, series prima cum tabularum decade prima. Berolini.

FALLENI, A., LUCCHESI, P., GHEZZANI, C., MCDONALD, J.C. and JONES, H.D. (2009). The female gonad in two species of *Microplana* (Platyhelminthes, Tricladida, Rhynchodemidae): Ultrastructural and cytochemical investigations. *J Morphol.* 270: 1042-1054.

FIELD, K.G., OLSEN, G.J., LANE, D.J., GIOVANNONI, S.J., GHISELIN, M.T., RAFF, E.C., PACE, N.R. and RAFF, R.A. (1988) Molecular phylogeny of the animal kingdom. *Science* 239: 748-753.

GAREY, J.R. (2002). The Lesser-Known Protostome Taxa: An Introduction and a Tribute to Robert P. Higgins. *Int. Comp. Biol.* 42: 611-618.

GEGENBAUR, C. (1859). Grundzuge der vergleichenden Anatomie. W. Engelmann, Leipzig, Germany.

GEGENBAUR, C., JEFFREY BELL, F. and RAY LANKESTER, E. (1878). Elements of Comparative Anatomy, MacMillan and Co, London.

GIRIBET, G., DISTEL, D.L., POLZ, M., STERRER, W. and WHEELER, W.C. (2000). Triploblastic relationships with emphasis on the acoelomates and the position of Gnathostomulida, Cycliophora, Plathelminthes, and Chaetognatha: a combined approach of 18S rDNA sequences and morphology. *Syst Biol* 49: 539-62.

GRAFF, L.V. (1882). Monographie der Turbellarien. I. Rhabdocoela.

HAECKEL, E. (1866). Generelle Morphologie der Organismen. Allgemeine Grundzüge der organischen Formen-Wissenschaft, mecanisch begründet durch die von Charles Darwin reformirte Descendenz-Theorie. 2 vols. Verlag von Georg Reimer, Berlin.

HAECKEL, E. (1876). The History of Creation (6th New English Edition 1914). D. Appelton and Company, New York.

HALANYCH, K.M. (2004). The new view of animal phylogeny. Annual Review of Ecology, Evolution, and Systematics 35: 229-256.

HALLEZ, P. (1894). Catalogue des Rhabdocoelides, Triclades and Polyclades du Nord de la France. 2nd ed. edn, L. Daniel, Lille (France).

HASZPRUNAR, G. (1996a). The Mollusca: coelomate turbellarians or mesenchymate annelids? In *Origin an Evolutionary Radiation of the Mollusca*. (Ed. J. Taylor, Oxford University Press, Oxford, pp. 1-28.

HASZPRUNAR, G. (1996b). Plathelminthes and Plathelminthomorpha - paraphyletic taxa. *J Zool Syst Evol Res* 34: 41-48.

HEJNOL, A., OBST, M., STAMATAKIS, A., OTT, M., ROUSE, G.W., EDGECOMBE, G.D., MARTINEZ, P., BAGUÑÀ, J., BAILLY, X., JONDELIUS, U., WIENS, M., MÜLLER, W.E.G., SEAVER, E., WHEELER, W.C., MARTINDALE, M.Q., GIRIBET, G. and DUNN, C.W. (2009). Assessing the root of bilaterian animals with scalable phylogenomic methods. *P Roy Soc B-Biol Sci* 276: 4261-4270.

HYMAN, L.H. (1951). The Invertebrates: Platyhelminthes and Rhynchocoela the acoelomate Bilateria, McGraw-Hill, London and New York.

HYMAN, L.H. (1940). The invertebrates: Protozoa through Ctenophora. Vol 1. New York.

HYMAN, L.H. (1937). Studies on the morphology, Taxonomy, and distribution of North American Triclad Turbellaria. VIII. Some cave Planarians of the United States. *Transactions of the American Microscopical Society* 56: 457-477.

JONDELIUS, U., RUIZ-TRILLO, I., BAGUÑÀ, J. and RIUTORT, M. (2002). The Nemertodermatida are basal bilaterians and not members of the Platyhelminthes. *Zool Scr* 31: 201-215.

JONES, H.D. and BOAG, B. (1996). The distribution of New Zealand and Australian terrestrial flatworms (Platyhelminthes: Turbellaria: Tricladida:Terricola) in the British Isles—the Scotish survey and Megalab worms. *J. Nat. Hist.* 30: 955-975.

KARLING, T.G. (1974). On the anatomy and affinities of the Turbellarian orders. In *Biology of the Tureballaria.* McGraw-Hill., NewYork: 1-16.

KENK, R. (1975). Fresh-water triclads (Turbellaria) of North America. VII. The genus Macrocotyla. *Trans Am Microsc Soc* 94: 324-339.

KENK, R. (1974). Index of the genera and species of the freshwater Triclads (Turbellaria) of the World. *Smithsonian Contrib. Zool.* 183: 1-90.

KENK, R. (1930). Beiträge zum System der Probursalier (Tricladida Paludicola). III. *Zool. Anz.* 89: 289-302.

LAZARO, E.M., SLUYS, R., PALA, M., STOCCHINO, G.A., BAGUÑÀ, J. and RIUTORT, M. (2009). Molecular barcoding and phylogeography of sexual and asexual freshwater planarians of the genus *Dugesia* in the Western Mediterranean (Platyhelminthes, Tricladida, Dugesiidae). *Mol Phylogenet Evol* 52: 835-845.

LAZARO, E., HARRATH, A.H., STOCCHINO, G., PALA, M., BAGUNA, J. and RIUTORT, M. (2011). *Schmidtea mediterranea* phylogeography: an old species surviving on a few Mediterranean islands? *BMC Evol Biol* 11: 274.

LITTLEWOOD, D.T.J. and OLSON, P.D. (2001). Small subunit rDNA and the Platyhelminthes: signal, noise and compromise. In *The interrelationships of the Platyhelminthes* (Ed. D.T.J.a.B. Littlewood R.D.) Taylor and Francis, London, England. pp: 262-278.

LITTLEWOOD, D.T.J., ROHDE, K., BRAY, R.A. and HERNIOU, E.A. (1999a). Phylogeny of the Platyhelminthes and the evolution of parasitism. *Biol J Linn Soc* 68 1-2: 257-287.

LITTLEWOOD, D.T.J., ROHDE, K. and CLOUGH, K.A. (1999b). The interrelationships of all major groups of Platyhelminthes: phylogenetic evidence from morphology and molecules. *Biol J Linn Soc* 66: 75-114.

MALLATT, J. and WINCHELL, C.J. (2002). Testing the new animal phylogeny: first use of combined large-subunit and small-subunit rRNA gene sequences to classify the protostomes. *Mol Biol Evol* 19: 289-301.

MALLATT, J., CRAIG, C.W. and YODER, M.J. (2010) Nearly complete rRNA genes assembled from across the metazoan animals: Effects of more taxa, a structure-based alignment, and paired-sites evolutionary models on phylogeny reconstruction. *Mol Phylogenet Evol* 55: 1-17.

MEDINA, M., COLLINS, A.G., SILBERMAN, J.D. and SOGIN, M.L. (2001). Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. *P Natl Acad Sci USA.* 98: 9707-12.

MITCHELL, R.W. (1968). New species of Sphalloplana (Turbellaria; Paludicola) from the caves of Texas and a reexamination of the genus Speophila and the family Kenkiidae. *Ann. Spéléol.* 23: 597-620.

MÜLLER, O.F. (1776). Zoologiae danicae prodromus, seu Animalium Daniae et Norvegiae indigenarum characteres, nomina, et synonyma imprimis popularium, Havniae.

NIELSEN, C. (2010). After all: Xenoturbella is an acoelomorph!. *Evol Dev* 12: 241-243.

NOREN, M. and JONDELIUS, U. (2002). The phylogenetic position of the Prolecithophora (Rhabditophora, ?Platyhelminthes?). *Zool Scr* 31: 403-414.

PAPS, J., BAGUÑÀ, J. and RIUTORT, M. (2009a). Bilaterian Phylogeny: A Broad Sampling of 13 Nuclear Genes Provides a New Lophotrochozoa Phylogeny and Supports a Paraphyletic Basal Acoelomorpha. *Mol Biol Evol* 26: 2397-2406.

PAPS, J., BAGUÑÀ, J. and RIUTORT, M. (2009b). Lophotrochozoa internal phylogeny: new insights from an up-to-date analysis of nuclear ribosomal genes. *P Roy Soc B-Biol Sci* 276: 1245-1254.

PHILIPPE, H., BRINKMANN, H., COPLEY, R.R., MOROZ, L.L., NAKANO, H., POUSTKA, A.J., WALLBERG, A., PETERSON, K.J. and TELFORD, M.J. (2011). Acoelomorph flatworms are deuterostomes related to Xenoturbella. *Nature* 470: 255-258.

REYNOLDSON, T.B. and BELLAMY, L.S. (1970). The status of Dugesia lugubris and D. polychroa (Turbellaria, Tricladida) in Britain. *J Zool Lond.* 162: 157-177.

RIUTORT, M., FIELD, K.G., TURBEVILLE, J.M., RAFF, R.R. and BAGUÑÀ, J. (1992) Enzyme electrophoresis, 18S rRNA sequences, and levels of phylogenetic resolution among several species of freshwater planarians (Platyhelminthes, Tricladida, Paludicola). *Can. J. Zool* 70: 1425-1439.

ROHDE, K. (1990) Phylogeny of platyhelminthes, with special reference to parasitic groups. *Int J Parasito.l* 20: 979-1007.

RUIZ-TRILLO, I., PAPS, J., LOUKOTA, M., RIBERA, C., JONDELIUS, U., BAGUÑÀ, J. and RIUTORT, M. (2002). A phylogenetic analysis of myosin heavy chain type II sequences corroborates that Acoela and Nemertodermatida are basal bilaterians. *P Natl Acad Sci USA.* 99: 11246-51.

RUIZ-TRILLO, I., RIUTORT, M., LITTLEWOOD, D.T.J., HERNIOU, E.A. and BAGUÑÀ, J. (1999). Acoel flatworms: earliest extant bilaterian Metazoans, not members of Platyhelminthes. *Science* 283: 1919-23.

RUPPERT, E.E., FOX, R. and BARNES, R.D. (2003). Invertebrate Zoology: A Functional Evolutionary Approach. Seventh Edition, Brooks/Cole Thompson Learning, Belmont, California.

SLUYS, R. (2001). Towards a phylogenetic classification and characterization of dugesiid genera (Platyhelminthes, Tricladida, Dugesiidae): a morphological perspective. In *The interrelationships of the Platyhelminthes* (Ed. D.T.J.a.B. Littlewood R.D.) Taylor and Francis, London, England. pp: 57-73.

SLUYS, R. (1999). Global diversity of land planarians (Platyhelminthes, Tricladida, Terricola): a new indicator-taxon in *Biodivers Conserv* studies. *Biodivers Conserv* 8: 1663-1681.

SLUYS, R. (1990). A monograph of the Dimarcusidae (Platyhelminthes, Seriata, Tricladida). *Zool Scr* 19: 13-29.

SLUYS, R. (1989a). A monograph of the marine triclads. A. A. Balkema, Rotterdam.

SLUYS, R. (1989b). Phylogenetic relationships of the triclads (Platyhelminthes, Seriata, Tricladida). *Bijdr. Dierkd.* 59: 3-25.

SLUYS, R., GRANT, L. and BLAIR, D. (2007). Freshwater planarians from artesian springs in Queensland, Australia (Platyhelminthes, Tricladida, Paludicola). *Contrib Zool* 76: 9-19.

SLUYS, R. and KAWAKATSU, M. (2006). Towards a phylogenetic classification of dendrocoelid freshwater planarians (Platyhelminthes): a morphological and eclectic approach.. 44: 274-284.

SLUYS, R., KAWAKATSU, M., RIUTORT, M. and BAGUÑÀ, J. (2009) A new higher classification of planarian flatworms (Platyhelminthes, Tricladida). *J Nat Hist* 43 29-30: 1763-1777.

SLUYS, R., KAWAKATSU, M. and WINSOR, L. (1998). The genus Dugesia in Australia, with its phylogenetic analysis and historical biogeography (Platyhelminthes, Tricladida, Dugesiidae). *Zool Scr* 27: 273-290.

SMITH J.P.S., TYLER S., RIEGER R.M. (1986). Is the Turbellaria polyphyletic?. *Hydrobiologia* 132: 13-21.

SOPOTT-EHLERS, B. (1985). The phylogenetic relationships within the Seriata (Platyhelminthes). In *The origins and relationships of lower invertebrates.* (Eds. S. Conway-Morris, J.D. George, R. Gibson and H.M. Platt) Clarendon Press, Oxford: 159-167.

SPERLING, E.A., PETERSON, K.J. and PISANI, D. (2009). Phylogenetic-Signal Dissection of Nuclear Housekeeping Genes Supports the Paraphyly of Sponges and the Monophyly of Eumetazoa. *Mol Biol Evol* 26: 2261-2274.

STEINBÖCK, O. (1925). Zur Systematik der Turbellaria metamerata, zugleich ein Beitrag zur Morphologie des Tricladen-Nervensystems. Zool. Anz. 64: 165-192.

SUGIURA, S. and YAMAMURA, Y. (2009). Potential impacts of the invasive flatworm Platydemus manokwari on arboreal snails. *Biol.Invasions* 11: 737-742.

SUNNUCKS, P., BLACKET, M.J., TAYLOR, J.M., SANDS, C.J., CIAVAGLIA, S.A., GARRICK, R.C., TAIT, N.N., ROWELL, D.M. and PAVLOVA, A. (2006). A tale of two flatties: different responses of two terrestrial flatworms to past environmental climatic fluctuations at Tallaganda in montane southeastern Australia. 15: 4513-4531.

TELFORD, M.J., LOCKYER, A.E., CARTWRIGHT-FINCH, C. and LITTLEWOOD, D.T.J. (2003). Combined large and small subunit ribosomal RNA phylogenies support a basal position of the acoelomorph flatworms. *P Roy Soc B-Biol Sci* 270: 1077-83.

TURBEVILLE, J.M. (2002). Progress in Nemertean Biology: Development and Phylogeny. *Integrative and Comparative Biology* 42: 692-703.

VALENTINE, J.W. (2004). On the Origin of Phyla. University Of Chicago Press.

WINSOR, L., JOHNS, P.M. and YEATES, G.W. (1998). Introduction, and ecological and systematic background, to the Terricola (Tricladida). *Pedobiologia* 42: 389-404.

Systematic Biology Advance Access published January 29, 2014

Point of View

Syst. Biol. 0(0):1–6, 2014

DOI:10.1093/sysbio/syt106

Upstream Analyses Create Problems with DNA-Based Species Delimitation

MELISA OLAVE[1], EDUARD SOLÀ[2], AND L. LACEY KNOWLES[3]

[1]*Centro Nacional Patagónico – Consejo Nacional de Investigaciones Científicas y Técnicas (CENPAT-CONICET), Puerto Madryn, Chubut U 9120 ACD, Argentina,* [2]*Department de Genètica, Facultat de Biologia and Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Av. Diagonal, 643, 08028, Barcelona, Catalonia, Spain and* [3]*Department of Ecology and Evolutionary Biology, The University of Michigan, Ann Arbor, MI 41809-1029, USA*
**Correspondence to be sent to: Department of Ecology and Evolutionary Biology, The University of Michigan, Ann Arbor, MI 41809-1029, USA; E-mail: knowlesl@umich.edu.*

Received 30 July 2013; reviews returned 18 September 2013; accepted 10 December 2013
Associate Editor: Robb Brumfield

Molecular data are expanding rapidly as a primary data source for species delimitation owing to both the availability of DNA sequences and recent analytical developments based upon the multispecies coalescent (Rannala and Yang 2003; Degnan and Rosenberg 2009). With such methodologies, species can be recognized despite genealogical discord across loci and incomplete lineage sorting (i.e., before reciprocal monophyly has been achieved) (Knowles and Carstens 2007). Nevertheless, we show that some of the zeal bestowed by theoretical ideals needs to be tempered by the practical problems associated with the implementation of coalescent-based approaches to species delimitation because of the potential for errors to be compounded across the multiple steps involved with analyzing DNA sequences.

Genetic approaches to species delimitation generally involve three separate steps: 1) assigning individuals to species, 2) estimating species relationships, and 3) in the case of Bayesian approaches to species delimitation (e.g., Yang and Rannala 2010), estimating the posterior probability that assigned groups are distinct (see O'Meara 2010 for a heuristic approach that does not require a priori assignment of individuals to species). The accuracy of approaches used for delimiting species in the latter two portions of this framework has received considerable attention (e.g., Liu 2008; Knowles 2009; Kubatko et al. 2009; Heled and Drummond 2010; Yang and Rannala 2010; Huang et al. 2010; Leaché and Rannala 2011; Camargo et al. 2012a; Knowles et al. 2012; Rannala and Yang 2013). In contrast, the assignment of individuals to putative species—the first step in species delimitation and pre-requisite in the increasingly popular Bayesian method implemented in the program bpp (Yang and Rannala 2010)—and how it impacts the accuracy of systematic studies that rely exclusively on genetic data for species delimitation has not been studied. Here, we specifically examine how the accuracy of assigning individuals to putative species impacts the downstream delimitation of species from the Bayesian program bpp (Yang and Rannala 2010).

The aim of this study is not an evaluation of bpp *per se*. In fact, previous studies have shown very good performance of bpp when the correct guide tree is provided, even with small datasets (Yang and Rannala 2010; Zhang et al. 2011; Camargo et al. 2012b; Rannala and Yang 2013). Our focus is on the input to bpp, and specifically, how errors and uncertainty with the assignment of individuals to species (i.e., determining individual-species associations) affect the accuracy of species delimitation. Our study focuses on the accuracy of delimited species from the Bayesian program bpp (Yang and Rannala 2010) when using the program STRUCTURAMA (Huelsenbeck and Andolfatto 2007), which like the program STRUCTURE (Pritchard et al. 2000; Falush et al. 2003), is advocated and typically used for the required a priori assignment of individuals to species in bpp, including for datasets with as few as six to eight loci (e.g., Leaché and Fujita 2010; Burbrink et al. 2011; Fujita et al. 2012).

Using simulated data, we chose a small set of conditions that differ with respect to the level of incomplete lineage sorting and conducted analyses with the goal of identifying which factors are driving the errors in the delimitation of species in downstream analyses with bpp (as opposed to characterizing the probability of errors in delimited species by simulating datasets across a broad range of divergence histories and sampling efforts). Nevertheless, the results are directly relevant to empiricists. For example, we focus on the first steps in the DNA sequence-based species delimitation process because of a mismatch between the theoretical recommendations for sampling (e.g., number of loci and individuals) for each of the separate components of

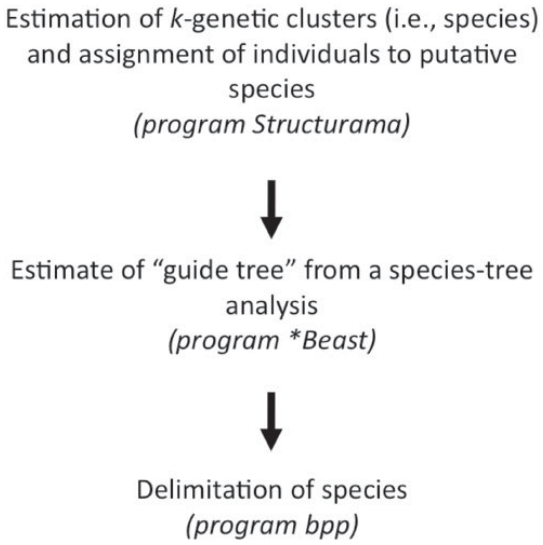


FIGURE 1. Steps involved in genetic-based species delimitation, which involve a series of analyses using different programs (which in this study involved STRUCTURAMA, *BEAST, and bpp). Note that bpp analyses were run with the following conditions: 1) set k = 8 with individuals assigned to species a priori (as opposed to estimating them), 2) set k = 16 for datasets with two individual sampled per species (i.e., assume that each individual is potentially a different species), and 3) set k = 8 and k =10, and estimate individual-species associations with STRUCTURAMA.

analysis, which have gone largely overlooked in practice (Fig. 1). In particular, although bpp may provide accurate estimates of the number of species with a sample of fewer than 10 sequenced loci (Yang and Rannala 2010), estimates of putative species numbers (i.e., k genetic clusters) and assignments of individuals to species with programs like STRUCTURE and STRUCTURAMA (Pritchard et al. 2000; Huelsenbeck and Andolfatto 2007) may not be accurate without large datasets (i.e., datasets approaching 100 independent loci; Rittmeyer and Austin 2012). This raises the concern that the results from empirical studies may be compromised by errors incurred during the estimation of the number of putative species and/or assignments of individuals to species in upstream analyses, even when the practices advanced for users of programs like bpp are followed (see Fujita et al. 2012). Moreover, by using simulated datasets that mirror empirical data collected for species delimitation and species-tree analysis, and in this specific case, a group of South American lizards (genus *Liolaemus*), the estimates can be compared with the known history to assess accuracy using sample sizes currently advocated as best practices. Not only do our results confirm that errors in the upstream analyses used to estimate individual-species association have a significant impact on the accuracy of delimited species but they also call into question current practices with species delimitation based solely on DNA sequences, despite the potential of such approaches in theory.

MATERIALS AND METHODS

Datasets

Simulated datasets, with respect to both the number of taxa and loci, correspond to many representative empirical datasets (reviewed in Fujita et al. 2012). Specifically, eight-taxon symmetric and asymmetric species trees were generated in Mesquite v2.74 (Maddison and Maddison 2010) under total tree depths of 0.4N and 4.0N, representing more and less difficult conditions for species delimitation, respectively (e.g., Knowles and Carstens 2007; Yang and Rannala 2010; Rittmeyer and Austin 2012). Note that we do not consider older species divergences because such scenarios are not particularly challenging and such data would not typically be analyzed with the coalescent-based approaches used here. Coalescent genealogies were generated for five individuals per species for each species tree using the program *ms* (Hudson 2002) under a model of constant population size, no migration, and no recombination within loci. DNA sequences were simulated with the program Seq-Gen (Rambaut and Grassly 1997). All nucleotide datasets were simulated under an HKY model of nucleotide substitution, with a transition–transversion ratio of 3.0, a gamma distribution with shape parameter of 0.8, and nucleotide frequencies of A = 0.3, C = 0.2, T = 0.3, and G = 0.2. Specifically, 1000 base pairs were generated, with θ = 0.07, which was estimated from an actual empirical lizard dataset (genus *Liolaemus*) (Olave et al. in review) using Lamarc v2.1.8 (Kuhner 2006). Similar results were observed with smaller theta values for simulating nucleotide datasets (results not shown). Datasets were simulated with 4, 8, and 14 loci, which cover the range of loci used in the majority of published datasets that apply this approach to delimit species (Fig. 1; reviewed in Fujita et al. 2012).

We also analyzed an empirical dataset with eight *Liolaemus* species of the *boulengeri* and *rothi* complexes (five individuals per species, 14 loci; details of the markers are shown in Supplementary Table S1; doi:10.5061/dryad.3hc8s). These taxa are a subset of those used in a large phylogenetic study of the genus (Olave et al. in review).

Analyses

For each species tree and sample design, 50 replicate datasets were analyzed (following the three steps summarized in Fig. 1; these are the same steps that an empiricist would follow). A total of 2400 bpp analyses were conducted across the 50 replicates of each simulated dataset under the different scenarios.

Individual-species associations.—During the standard practice of species delimitation (Fig. 1), the number of genetic groups (or putative species in this case, and hereafter referred to as species) and individual-species associations would be estimated, for example, using the software STRUCTURAMA 2.0 (Huelsenbeck and

Andolfatto 2007). However, because of difficulties with accurately estimating the number of k genetic clusters using STRUCTURAMA (i.e., the number of species was significantly underestimated under a Dirichlet process prior, with an average of $k = 4$ across datasets), the total number of distinct species was not estimated. Instead of estimating the number of k genetic clusters (i.e., species), individual-group associations were determined assuming eight distinct genetic groups (i.e., a $k = 8$, which corresponded to the actual conditions used to simulate the data). Because k was set to the known value, issues over how to estimate k (see Evanno et al. 2005) do not confound the interpretations of our results from the analyses. However, note that our results are conservative with respect to the errors introduced to downstream analyses involved in delimiting species because we set the number of putative species to the actual value k, as opposed to estimating k. To confirm that the difficulties with estimating the number of k genetic clusters reflect limited amounts of sequence data (see also Rittmeyer and Austin 2012), rather than a sensitivity to the number of taxa used in the simulations, we also simulated and analyzed 50 replicate datasets for species trees with two and four taxa, instead of eight, with five individuals per species under the same parameter settings described earlier. Inaccuracy of the estimated number of clusters was also observed for these datasets, with k significantly overestimated (e.g., $k > 10$ with the four-taxon datasets). Hence, only the eight-taxon datasets were considered for further analyses and discussion. All STRUCTURAMA analyses were run for a total of one million generations for each diploid dataset, sampling every 100 generations; 10% of the data were discarded as burn-in.

Because we are interested in ways that might improve the accuracy of DNA sequence-based species delimitation, we also used a slightly larger number than the actual number of species (i.e., set $k = 10$) for estimating individual-group associations. This decision was made because the maximum number of species a program like bpp can identify is set by the user based on the input of the guide tree. By using a larger number of genetic groups (e.g., $k = 10$ when the data were simulated under a $k = 8$), we can therefore evaluate whether downstream analyses of the species delimitation process are robust to divisions of genetic groupings that are slightly finer than the actual species boundaries. This issue has never been investigated in bpp.

We also considered an alternative approach in which each individual is treated as a potential species, thereby skipping the first step of estimating the number of putative species and assigning individuals to putative species with a program like STRUCTURAMA (and likewise, by passing the potential errors in individual-species associations). For these analyses, only two individuals per species (for a total of 16 potential species) were considered because of computational constraints with bpp; the input into bpp was the tree estimated from *BEAST (Heled and Drummond 2010).

Generating a guide tree of the relationships among putative species.—A species tree was estimated for each dataset using *BEAST (Heled and Drummond 2010) for individual-species assignments based on estimates made with either $k = 8$ or $k = 10$ in STRUCTURAMA, or considering each individual as a potential species (i.e., $k = 16$ in this case with two individuals sampled per species). Each *BEAST analysis was run for 50 million generations with samples taken every 5000 generations and 10% of the data discarded as burn-in, with a model of nucleotide evolution that matched the simulated data (as detailed earlier). Effective sample size (ESS) values were checked and for the few cases where ESS were <200, we ran the Markov chain Monte Carlo (MCMC) until every ESS parameter was >200.

Species delimitation with the program bpp.—The program bpp samples from the posterior distribution of models of species limits using reversible-jump MCMC. That is, given a starting guide tree, the program sequentially collapses internal nodes in the guide tree, evaluating the posterior distribution for each of fewer and fewer putative species. The program assumes no recombination within a locus, free recombination between loci, no gene flow between species, and that the DNA sequences evolved neutrally.

The simulated data were analyzed with bpp v2.0 (Yang and Rannala 2010) with algorithm 1 and the finetune parameter ε set to 15. For species trees with a depth of 4N, we set θ and τ (the timing of species divergence) priors to values that encompass those used to simulate the data, specifically G (7, 100), which results in a mean = 0.07. For the more recent divergence history of 0.4N, the priors on θ and τ were adjusted accordingly to G (0.7, 100), which results in a mean = 0.007. The step lengths for proposals in the MCMC were automatically adjusted to obtain optimal acceptance rates during the analysis that consisted of a burn-in phase of 10 000 steps and 100 000 posterior samples sampled every two steps. Runs were checked to make sure values were between 0.15 and 0.7, as well as ESS values were >200 to assure convergence.

Accuracy of analyses

The accuracy of species delimitation at different steps in the process was evaluated using a number of metrics. This included measures of errors associated with upstream analyses that might impact the downstream bpp analysis (see Fig. 1).

Accuracy of individual-species associations.—A simple index (I_s) was used to examine errors in upstream analyses involving the assignment of individuals to their respective putative species (i.e., individual-species associations). This index, (I_s), measures how many times actual species lineages (i.e., known species lineages) were split as

$$\frac{\sum_{i=1}^{k} \frac{n_{g_i}-1}{n_{r_i}-1}}{k_r}$$

where the numbers of different species (or genetic clusters) that software recognized within the *i*th actual species minus one ($n_{gi} - 1$), is calculated relative to the maximum number of splits possible, which is the actual number of individuals that are part of the *i*th species minus one ($n_{ri} - 1$). The index ranges from zero (perfect assignment) to one (species maximally oversplit). Additionally, a mean was calculated among *k* species.

Number of putative species recovered by bpp.—A mean and standard deviation of number of species delimited per dataset among the 50 replicated analyses were calculated for each scenario and combination of different number of loci.

Type I error estimation (failure to reject the wrong hypothesis).—We calculate the proportion of analyses that led to well supported, but nonetheless incorrect inferences about the number of putative species (i.e., under- or overestimates of the number of species with posterior probabilities >0.95). We also used the R statistical software environment (R Core Team 2013) to test for an association between datasets with incorrectly delimited numbers of species (i.e., |the actual number of species – the estimated number of species from bpp|) and the estimated posterior probabilities from the bpp analyses, using linear regression analyses and correlation tests.

Results and Discussion

For the number of loci considered here (which span those typically used in empirical studies that delimited species with bpp to date), there were frequent errors in the delimitation of species for both divergence times (Fig. 2). Particularly disconcerting is the high error rates in the delimitation of species even when the correct number of species is set in STRUCTURAMA (i.e., *k* = 8), with almost all datasets showing errors in the delimitation of species with four loci (i.e., >90% of datasets) and most datasets showing errors with eight loci (i.e., >60% of datasets). Surprisingly in some cases (i.e., when the number of putative species is incorrectly set at *k* = 10), the support for the wrong number of species actually gets stronger with the addition of loci (i.e., the frequency of species delimited with posterior probability of >0.95, shown in black, increases disproportionately relative to the total frequency of errors).

These errors in the delimited species do not reflect the inherent difficultly (i.e., the recency of diversification) of the scenarios represented in the simulations, such that the datasets are simply intractable with respect to analysis with bpp. In almost every case where the number of putative species and individual-species associations were input into bpp (i.e., when they are not estimated with STRUCTURAMA), the number of species was accurately delimited (see Fig. 2 for the few exceptions), which is consistent with studies investigating the performance of bpp by itself (e.g., Yang and Rannala 2010; Zhang et al. 2011; Camargo et al. 2012b). Likewise, the high errors in delimitation do not apparently reflect recalcitrant species-tree estimates. If this was the case, when each individual sampled was treated as a putative species (i.e., *k* = 16 in this case, where two individuals per species were simulated), we would expect pervasive high error rates in the delimitation of species with bpp because of errors in the guide tree. Yet, instead much lower error rates were observed when each individual was treated as a putative species (Fig. 2).

Considered together, these results highlight that a primary source of error in the upstream analysis involves the assignment of individuals to putative species (discussed below). Moreover, the large impact of upstream analyses on the accuracy of downstream analyses used to delimit species (Fig. 2) not only

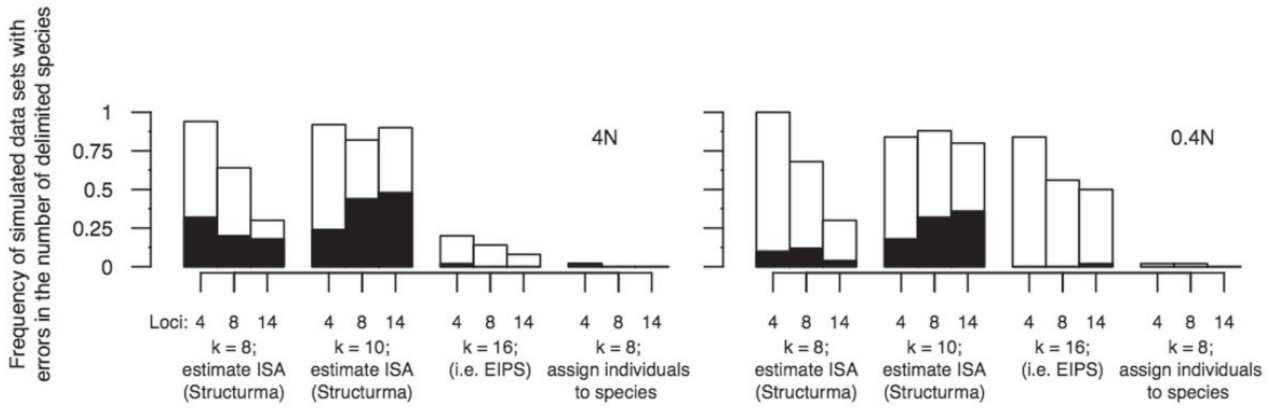


FIGURE 2. The frequency of incorrect inferences with bpp about the number of species delimited across simulated datasets for different sampling efforts and when individual-species associations (ISA) were estimated with STRUCTURAMA with different settings for numbers of putative species (i.e., *k* = 8, *k* = 10), or when the species were correctly assigned to the known species, or when each individual was treated as a potential species in bpp (i.e., EIPS). In some cases, support for the wrong number of species gets stronger with additional loci (i.e., the number of species delimited with posterior probability of > 0.95, shown in black, increases disproportionately). Only the results for the simulations under an asymmetric species tree are shown, and see Supplementary Figure S6 for similar results under a symmetric species tree.

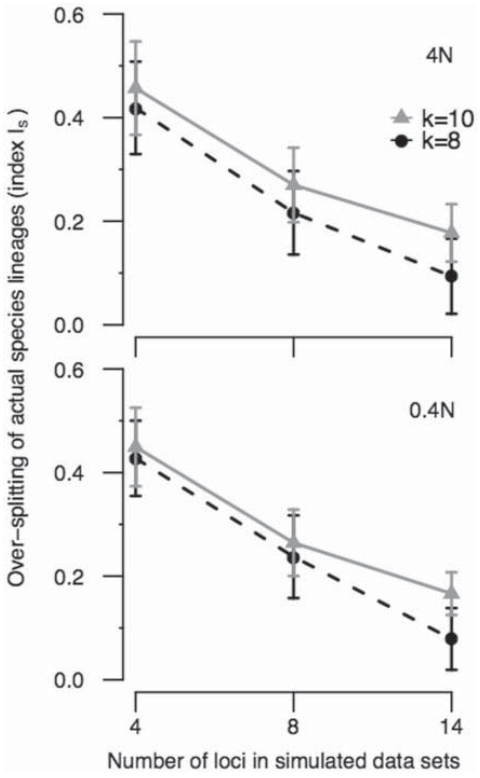


FIGURE 3. Measure of the over-splitting of actual species lineages by the index I_s for different numbers of putative species, k, used for assigning individuals to species and for estimating the guide tree for bpp, for simulated datasets with four, eight, or 14 loci; only the results for the simulations under an asymmetric species tree (at 4N and 0.4N total tree depth) are shown, and see Supplementary Figure S7 for similar results under a symmetric species tree. The index ranges from zero (perfect assignment) to one (species maximally over split).

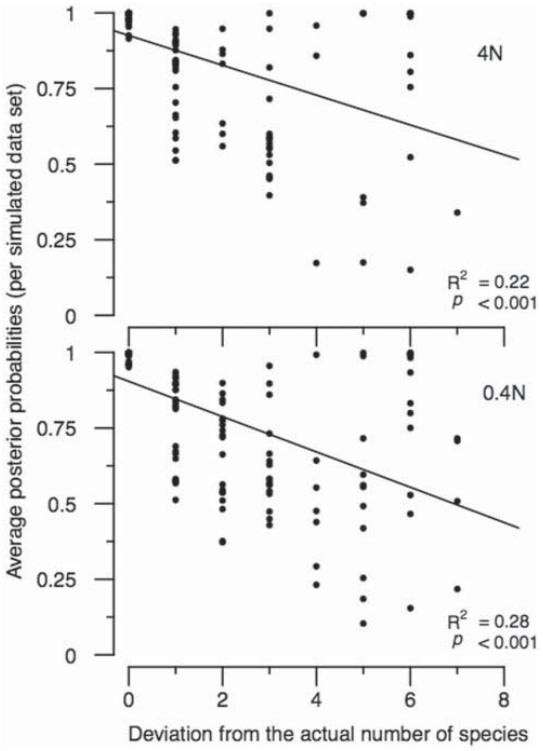


FIGURE 4. Negative association between the posterior probabilities of species delimited with bpp and the deviation from the actual number of species (i.e., 8) when the putative number of species, k, is set to 8, and individual-species associations are estimated with STRUCTURAMA. Note that the correlation was only significant when the putative number of species k is set as 8 (see Supplementary Fig. S9 for k set as 10). Only the results for the simulations under an asymmetric species tree (at 4N and 0.4N total tree depth) are shown given the similarity of results under a symmetric species tree (see Supplementary Fig. S8).

highlights a significant problem in current practices (summarized in Fig. 1) but also suggests an alternative approach for delimiting species with genetic data that may prove more accurate (discussed below).

Impact of errors with upstream analyses on the accuracy of bpp output..—The analyses show that there were always errors with the assignment of individuals to species (Fig. 3), even when the correct number of species (i.e., k = 8) and largest number of loci were used (14 loci, which is consistent with the sample sizes used in empirical datasets; see Fujita et al. 2012). Larger numbers of loci can certainly reduce the errors with upstream STRUCTURAMA (or STRUCTURE) analyses (see Rittmeyer and Austin 2012), as might lower haplotype diversity within loci, given that information about coancestry among individuals from k putative species are characterized by a set of allele frequencies at each locus with these programs (Huelsenbeck and Andolfatto 2007). Nevertheless, our results highlight the problems that can arise because of the mismatch in the data types required at different steps in the delimitation process (Fig. 1) and the high error rates that may accompany studies that rely exclusively on limited numbers of DNA sequences to delimit species (Fig. 2) without some additional data for improving the accuracy of assigning individuals to putative species for recently diverged taxa.

Because we used simulations with a known history, we are able to explore the cause of errors in the downstream analyses (i.e., we can show it is not a function of an intractable history with respect to estimating a guide tree or the delimitation process implemented in bpp, as discussed earlier; Figs. 2 and 3). We can also show that although the posterior probabilities from bpp analyses may be negatively correlated with the number of incorrectly delimited species (Fig. 4), the high variance among replicate datasets (at both k = 8 and k = 10) means that it is possible to get strong support for incorrect estimates of the number of putative species (see also Fig. 2).

These findings have direct relevance to observations from empirical studies regarding the delimitation of species using DNA sequences exclusively. For example, consistent with the high errors in the detection of putative species and assignment of individuals to taxa observed in the simulations here, different empirical studies have also shown that genetic data alone did not detect the same number of putative species recognized in traditional taxonomic treatments in upstream analyses

TABLE 1. Results of the analysis of empirical lizard dataset (genus *Liolaemus*) with properties that corresponded to the simulated datasets (with respect to sampling effort and models of nucleotide variation; see Supplementary Table S1) when the number of species and individual-species associations are set according to traditional taxonomic criteria (primarily morphological features), as opposed to using estimates from STRUCTURAMA

Number of putative species, k	Individual-species associations	I_s	Delimited species with bpp	Posterior probabilities
Set at $k = 8$ according to traditional taxonomy	Set according to traditional taxonomy	*na*	8	0.995584
Estimated with STRUCTURAMA, $k = 9$	Estimated with STRUCTURAMA	0.196	7	0.551412
Set at $k = 8$ according to traditional taxonomy	Estimated with STRUCTURAMA	0.053	7	0.987816
Set to at $k = 10$	Estimated with STRUCTURAMA	0.071	7	0.478044

Notes: The I_s-index is a measure of the over-splitting of actual species lineages; I_s is not applicable (na) when the number of putative species is set at $k = 8$ according to traditional taxonomy.

(e.g., Harrington and Near 2012; Edwards and Knowles 2014). With the simulated datasets analyzed here, a much lower number of species was estimated with STRUCTURAMA as well, with an average of $k = 4$ (as noted in the section Materials and Methods, this is why we set k, rather than estimated k). As a consequence, an underestimation of taxa compared with traditional taxonomic treatments would result from estimates with bpp without alternative approaches for establishing individual-species associations. Because the number of estimated species can only decrease, and not increase, from the number of putative species identified in the guide tree used by bpp (Yang and Rannala 2010), underestimates of number of putative species in upstream analyses will always have a significant impact. Such underestimates are certainly not obvious based on an examination of the support values accompanying delimited species. High posterior probability support is associated with many simulated datasets in which the number of species is underestimated with bpp (Fig. 4; and Supplementary Fig. S8). Likewise, when analyzed following the standard protocol advocated for species delimitation (Fig. 1), the actual DNA sequences collected in the *Liolaemus* lizards also provide what appears to be an underestimate of the putative species with high posterior support compared with recognized taxa based on morphology (Table 1).

Alternative procedures in the delimitation of species

Interestingly, treating each individual as a possible species—that is, bypassing the steps of estimating putative species and assigning individuals to these taxa with STRUCTURAMA—produced fewer errors in the delimitation of species with bpp (Fig. 2). Although the correct number of species was frequently delimited with this approach, unfortunately the support for the delimited species was consistently quite low (Fig. 5). This means it is unlikely that the analysis would be interpreted as supporting the correct number of species (which in this case was 8). Note that when the empirical data from *Liolaemus* were analyzed using this strategy, indeed very low posterior probabilities were observed (an average of 0.3411). The low posterior probabilities from the bpp analyses probably reflect the limited information contained in the data about the effective population size of species, a key parameter in bpp, considering that only two individuals were sampled per species (i.e., setting $k = 16$). Adding more individuals sampled per species would provide more information for estimating population parameters (see Yang and Rannala 2010). However, this strategy of considering each individual as a putative species would also have the undesirable effect of increasing the number of parameters to be estimated in bpp, as well as introducing additional errors in the guide tree because of incomplete lineage sorting.

Of course the approach discussed here (Fig. 1), and the program bpp in particular, is just one of many different methods available for species delimitation based on genetic data (reviewed in Carstens et al. 2013). Moreover, despite the failure to accurately delimit species for the set of conditions simulated here, we are not suggesting that researchers should avoid bpp and adopt a different program. Given differences in the assumptions and algorithms employed across methods, the accuracy of the delimited species from the simulated datasets could very well differ depending upon the method used. Instead, our aim is to draw attention to what are potentially compounded problems when the properties of the genetic datasets are sufficient for one, but not all steps in the practice of species delimitation.

Applying multiple genetic markers, such as single nucleotide polymorphisms or microsatellites across multiple loci for the estimation individual-species associations for k putative species and multilocus DNA sequence data for downstream bpp analyses, could provide one obvious potential solution. Another alternative and efficient approach, and perhaps the most cost effective, would be to use more than one data type for species delimitation. For example, traditional taxonomic boundaries might be used in cases where such information is available to determine the number of putative species and assign individuals to species, thereby bypassing the high errors associated with using DNA sequences from a limited number of loci to perform such steps. This alone should greatly enhance the accuracy of species delimitation (i.e., compare the results when putative species and individual-species

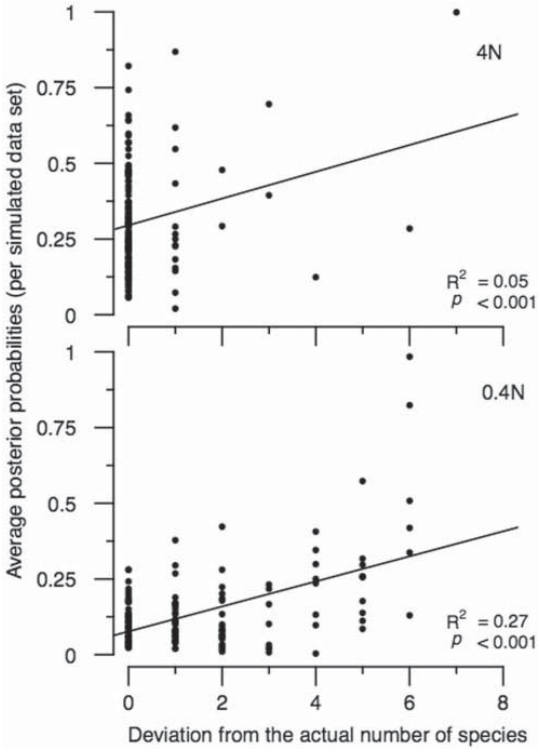


FIGURE 5. Positive association between the posterior probabilities of species delimited with bpp and the deviation from the actual number of species when each individual is treated as potentially a different species (i.e., k = 16, given that two individuals were sampled per species); only results for the simulations under an asymmetric species tree are shown; see Supplementary Figure S10 for similar results under a symmetric species tree.

associations are estimated from genetic data to when they are set, Fig. 2). Morphological and geographic data can also provide valuable information in delimiting species (e.g., Zapata and Jiménez 2012), especially in the identification of putative species and establishing individual-species associations needed for downstream analyses with bpp, even in cases with cryptic species are involved (e.g., Barley et al. 2013).

There is also arguably inherent merit in incorporating multiple data types when delimiting species, which extends beyond the aim of avoiding potential errors in upstream analyses that impact DNA-based estimates of putative species. These pertain to the interpretation of our DNA-based putative species. Depending upon the genetic markers and sampling strategy employed, there is no theoretical reason why the "minimal diagnostic genetic unit" would not extend below species boundaries. As such, it is important to recognize that the issues surrounding DNA-based species delimitation are certainly broader than decisions about what particular analytical approach to use to analyze the genetic data or whether different approaches produce congruent results (see discussion in Carstens et al. 2013). Efforts toward developing methods to accommodate multiple data types in a single quantitative framework, as opposed to the sequential analyses used to integrate information from different data types, are critically needed (Yeates et al. 2011). If such model-based approaches could be extended to multiple data types, such as morphology (i.e., a program that considers not only neutral markers but also morphological characters, including those undergoing selective divergence, for evaluating hypotheses about putative species), we could accommodate taxa where divergence might be more evident along axes of differentiation other than neutral genetic divergence. Moreover, it would bring the field of species delimitation one step closer to identifying boundaries that reflect the accumulation of differences associated with reproductive isolation, as opposed to the ephemeral boundaries only evident in the patterns of neutral genetic markers (i.e., differentiation below the species level).

CONCLUSIONS

Our study highlights how errors in upstream analyses, and specifically, the estimation of individual-species associations, impact the accuracy of downstream analyses with the program bpp. Contrasts in the accuracy of delimited species when individual-species associations are estimated versus setting them to conditions used in the simulations demonstrate that the errors encountered in the bpp analyses are not simply a byproduct of recalcitrant species histories. The errors associated with assignment of individuals to species reflect the mismatch in data requirements at different steps in the process—in fact, the frequency of error estimates reported here is an underestimate given that we set the number of putative species when estimating individual-species associations (e.g., k = 8), rather than estimating both the number of putative species and individual-species associations with the program STRUCTURAMA (see Evanno et al. 2005). Interestingly, higher accuracy of delimitation with bpp was achieved when treating each individual sampled as a putative species, but the low posterior probabilities from such analyses mean it is unlikely that this alternative approach of bypassing the errors in upstream analyses will be useful practically. Overall, these results raise significant questions about current advocated practices for DNA sequence-based species delimitation (note the number of loci used in the simulations, albeit limited, covers the range from the majority of published papers to date).

We suggest that complementing DNA-based approaches for delimitation with other data types, such as morphology, especially for the assignment of individuals to putative species, may be one of the best ways to increase the accuracy of species delimited with programs like bpp (as also noted by Yang and Rannala 2010), which by themselves are accurate with limited genetic data. Moreover, the integration of data types might be necessary given that increasing the number of loci, for example, by applying next-generation sequencing technologies, is unlikely to provide a simple

solution because here too lies a mismatch between data requirements for the programs used in the delimitation process. That is, the short sequence reads from next-generation sequencing platforms (e.g., those from Illumina), while compatible for estimating individual-species associations based on allele frequencies at each locus with programs like STRUCTURAMA, are not ideal for gene-tree based approaches like bpp. Finally, without integrating across data types, interpreting what our DNA-based approaches actually delimit (i.e., putative species, populations, or kin groups) will remain ambiguous, reflecting the resolution of the genetic markers and sampling strategy of the researcher.

Supplementary Material

Data files and/or other supplementary information related to this paper have been deposited at Dryad under doi:10.5061/dryad.3hc8s.

Funding

This work was supported by a doctoral fellowship from Consejo Nacional de Investigaciones Científicas y Técnicas from Argentina (CONICET) (to M.O.); a Fulbright-Bunge y Born fellowship (to M.O.); doctoral fellowships BES-2009-022530 and EEBB-1-2012-05462 from the Ministerio de Ciencia e Innovación from Spain (to E.S.); and a NSF grant [DEB 11-18815] (to L.L.K.).

Acknowledgments

Thanks to members of the Knowles lab, especially to D. Alvarado, C. Muñoz, Q. He, and H. Lanier for helpful comments and suggestions. We also thank all members of Grupo de Herpetología Patagónica (CENPAT-CONICET, Argentina), especially Dr. M. Morando and Dr. L.J. Avila, Dr. M. Riutort group (Facultat de Biologia and IRBio, Universitat de Barcelona), and Dr. J.W. Sites Jr. for generously allowing us to use the marylou5 supercomputer cluster in the Fulton Supercomputing Lab at Brigham Young University (BYU). This research benefitted from valuable comments from three anonymous reviewers.

References

Barley A.J., White J., Diesmos A.C., Brown R.M. 2013. The challenge of species delimitation at the extremes: diversification without morphological change in Philippine sun skinks. Evolution 67:3556–3572.

Burbrink F.T., Yao H., Ingrasci M., Bryson R.W. Jr., Guiher T.J., Ruane S. 2011. Speciation at the Mogollon Rim in the Arizona Mountain Kingsnake (*Lampropeltis pyromelana*). Mol. Phylogenet. Evol. 60:445–454.

Camargo A., Avila L.J., Morando M., Sites J.W. Jr. 2012a. Accuracy and precision of species trees: effects of locus, individual, and base pair sampling on inference of species trees in lizards of the *Liolaemus darwinii* group (Squamata, Liolaemidae). Syst. Biol. 61:272–288.

Camargo A., Avila L.J., Morando M., Sites J.W. Jr. 2012b. Species delimitation with abc and othercoalescent-based methods: a test of accuracy with simulations and anempirical example with lizards of the *Liolaemus darwinii* complex (Squamata:Liolaemidae). Evolution 66:2834–2849.

Carstens B.C., Pelletier T.A., Reid N.M., Salter J.D. 2013. How to fail at species delimitation. Mol. Ecol. 22:4369–4383.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference, and the multispecies coalescent. Trends Ecol. Evol. 24:332–340.

Edwards D.L, Knowles L.L. 2014. Species detection and individual assignment in species delimitation: can integrative data increase efficacy? Proc. R. Soc. B 20132765. http://dx.doi.org/10.1098/rspb.2013.2765

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14:2611–2620.

Falush D., Stephens M., Pritchard J.K. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587.

Fujita M.K., Leaché A.D, Burbrink F.T., McGuire J.A., Mortiz C. 2012. Coalescent-based species delimitation in an integrative taxonomy. Trends Ecol. Evol. 27:480–488.

Harrington R.C., Near T.J., 2012. Phylogenetic and coalescent strategies of species delimitation in Snubnose Darters (Percidae: *Etheostoma*). Syst. Biol. 61:63–79.

Heled J., Drummond A.J. 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27:570–580.

Huang H., He Q., Kubatko L.S., Knowles L.L. 2010. Sources of error for species-tree estimation: impact of mutational and coalescent effects on accuracy and implications for choosing among different methods. Syst. Biol. 59:573–583.

Hudson R.R. 2002. Generating samples under a Wright-Fisher neutral model. Bioinformatics 18:337–338.

Huelsenbeck J.P., Andolfatto P. 2007. Inference of population structure under a Dirichlet process model. Genetics 175:1787–1802.

Knowles L.L. 2009. Statistical phylogeography. Annu. Rev. Ecol. Syst. 40:593–612.

Knowles L.L., Carstens B.C. 2007. Delimiting species without monophyletic gene trees. Syst. Biol. 56:887–895.

Knowles L.L., Lanier H., Klimov P.B., He Q. 2012. Method choice and species-tree accuracy: full modeling versus summarizing gene-tree uncertainty Mol. Phylogenet. Evol. 65:501–509.

Kubatko L.S., Carstens B.C., Knowles L.L. 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25:971–973.

Kuhner M.K. 2006. LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22: 768–770.

Leaché A.D., Fujita M.K. 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proc. R. Soc. Lond. B Biol. Sci. 277:3071–3077.

Leaché A.D., Rannala B. 2011. The accuracy of species tree estimation under simulation: a comparison of methods. Syst. Biol. 60:126–137.

Liu L. 2008. BEST: Bayesian estimation of species trees under the coalescent model. Bioinformatics 24:2542–2543.

Maddison W.P., Maddison D.R. 2010. Mesquite: a modular system for evolutionary analysis. Version 2.74. Available from: http://mesquiteproject.org.

Olave M., Avila L.J., Sites J.W. Jr., Morando M. Model-based approach to test hard polytomies in the Eulaemus clade of the most diverse South American lizard genus Liolaemus (Liolaemini, Squamata). J Evol. Biol. (in review)

O'Meara B. 2010. New heuristic methods for joint species delimitation and species tree inference. Syst. Biol. 59:59–73.

Pritchard J.K., Stephens M., Donnelly P.J. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

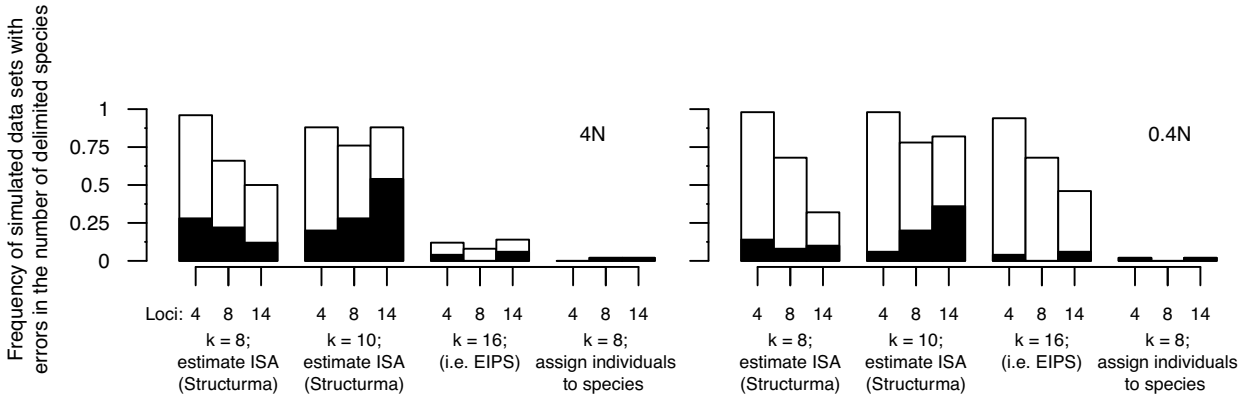
R Core Team. 2013. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: http://www.R-project.org/.

Rambaut A., Grassly N.C. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput. Appl. Biosci. 13:235–238.

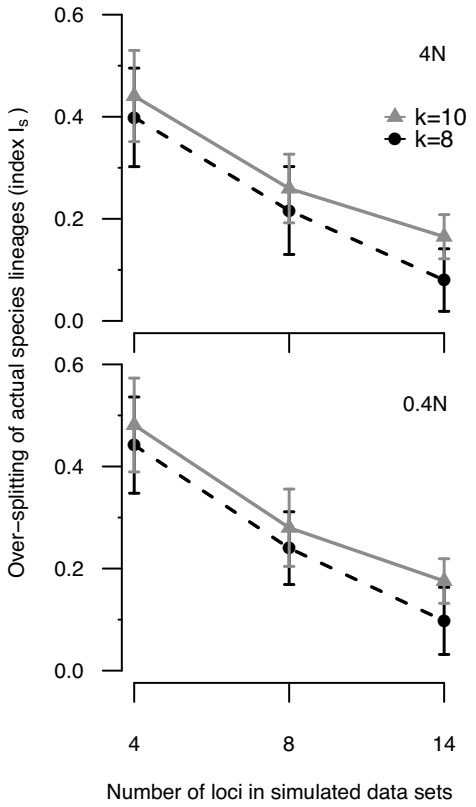
Rannala B., Yang Z. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164:1645–1656.
Rannala B., Yang Z. 2013. Improved reversible jump algorithms for Bayesian species delimitation. Genetics 194:245–253.
Rittmeyer E.N., Austin C.C. 2012. The effects of sampling on delimiting species from multi-locus sequence data. Mol. Phylogenet. Evol. 65:451–463.
Yang Z., Rannala B. 2010. Bayesian species delimitation using multilocus sequence data. Proc. Natl. Acad. Sci. USA. 107:9264–9269.
Yeates D.K., Seago A., Nelson L., Cameron S.L., Joseph L., Trueman J.W.H. 2011. Integrative taxonomy, or iterative taxonomy? Syst. Entomol. 36:209–217.
Zapata F., Jiménez I. 2012. Species delimitation: inferring gaps in morphology across geography. Syst. Biol. 61:179–194.
Zhang C., Zhang D.X., Zhu T., Yang Z. 2011. Evaluation of a Bayesia coalescent method of species delimitation. Syst. Biol. 60: 747–761.

SUPPLEMENTARY FIGURES

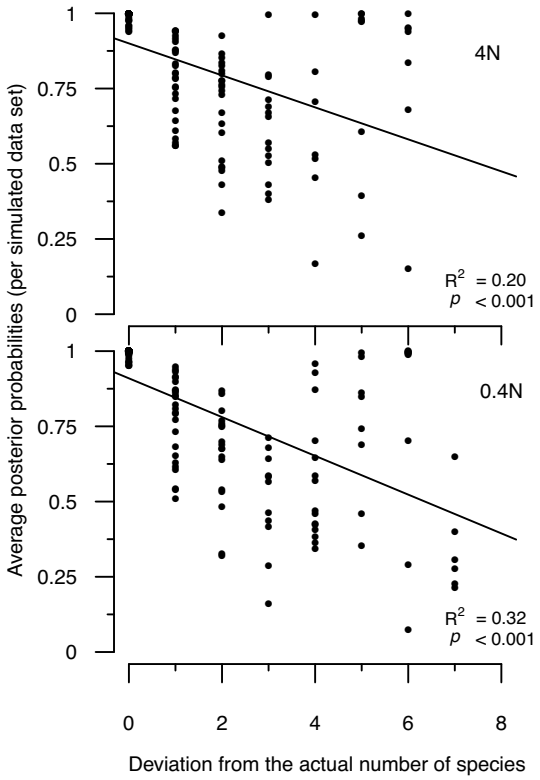
SUPPLEMENTARY FIGURE 6: The frequency of incorrect inferences with bpp about the number of species delimited across simulated data sets for a symmetric species tree with different sampling efforts and when individual-species associations were estimated with Structurama with different settings for numbers of putative species (i.e., $k = 8$, $k = 10$), or when each individual was treated as a potential species in bpp (i.e., EIPS), or when the species were correctly assigned to the known species (TT). In some cases, support for the wrong number of species gets stronger with additional loci (i.e., the number of species delimited with posterior probability of > 0.95, shown in black, increases disproportionately).



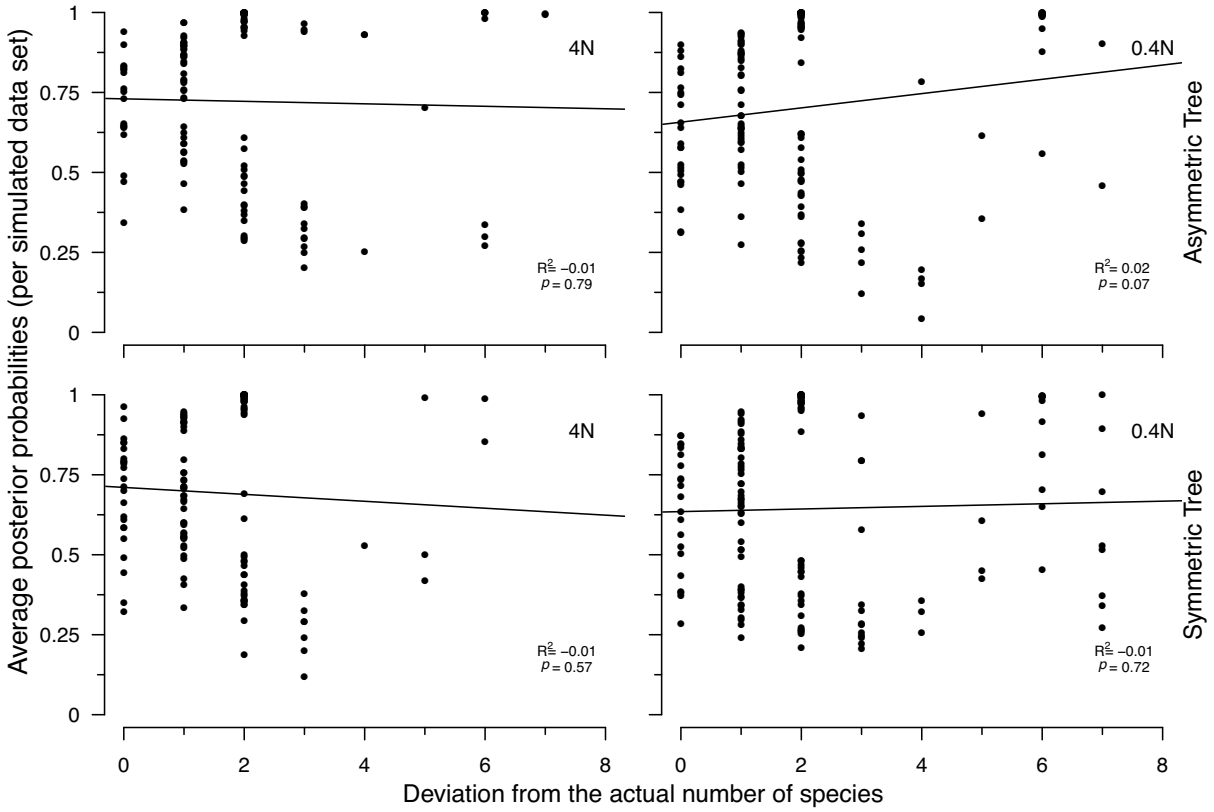
SUPPLEMENTARY FIGURE 7: Measure of the over-splitting of actual species lineages by the index I_s for different numbers of putative species, *k*, used for assigning individuals to species and for estimating the guide tree for bpp, for simulated datasets under a symmetric species tree with four, eight, or 14 loci; the index ranges from zero (perfect assignment) to one (species maximally over split).



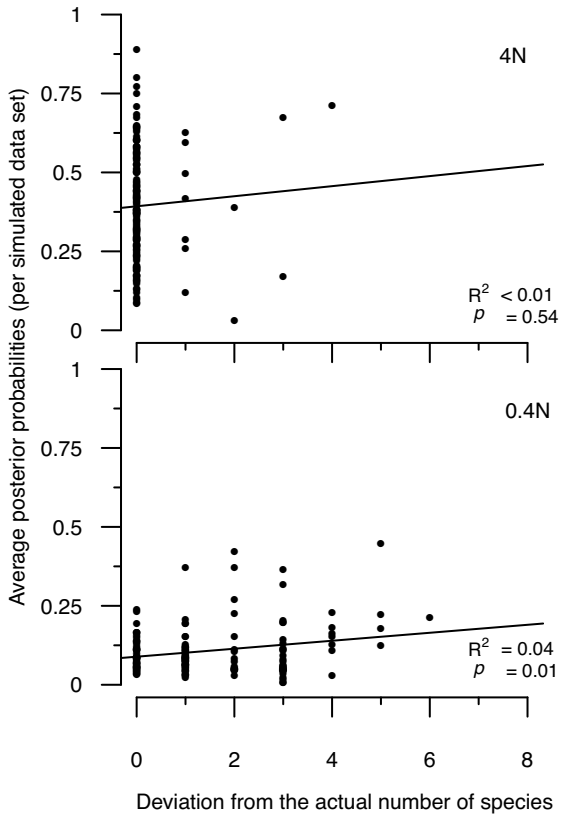
SUPPLEMENTARY FIGURE 8: Correlation between the posterior probabilities of species delimited with bpp and the deviation from the actual number of species (i.e., 8) when the putative number of species, *k*, is set to 8, and individual-species associations are estimated with Structurama; results are for the simulations under a symmetric species tree (at 4N and 0.4N total tree depth).



SUPPLEMENTARY FIGURE 9: Correlation between the posterior probabilities of species delimited with bpp and the deviation from the actual number of species (i.e., 8) when the putative number of species, *k*, is set to 10, and individual-species associations are estimated with Structurama; results are for simulations under an asymmetric and symmetric species tree (at 4N and 0.4N total tree depth).



SUPPLEMENTARY FIGURE 10: Correlation between posterior probabilities of species delimited with bpp and the deviation from the actual number of species when each individual is treated as potentially a different species (i.e., $k = 16$, given that 2 individuals were sampled per species); results are for simulations under an asymmetric and symmetric species tree (at 4N and 0.4N total tree depth).



SUPPLEMENTARY TABLE 1. List of molecular markers with information about number of base pairs and estimated model of nucleotide evolution (see references for details) for empirical dataset of *Liolaemus* lizard genus (*boulengeri* and *rothi* complexes) used to inform the simulations for this study; ANL = anonymous nuclear locus and NPCL = nuclear protein-coding locus.

Locus	Nature	Length (bp)	Model	Reference
12S	mitochondrial	897	GTR+I+G	Morando et al. (2004)
cyt-b	mitochondrial	804	HKY+I+G	Morando et al. (2004)
A12D	ANL	736	GTR	Camargo et al. (2012)
A1D	ANL	938	HKY+G	Camargo et al. (2012)
A4B	ANL	508	HKY+G	Camargo et al. (2012)
A9C	ANL	739	HKY+G	Camargo et al. (2012)
CMOS	NPCL	526	HKY	Saint et al. (1998)
DNAH3	NPCL	725	HKY+G	Townsend et al. (2008)
EXPH5	NPCL	866	HKY+G	Portik et al. (2011)
KIF24	NPCL	533	HKY+I	Portik et al. (2011)
MXRA5	NPCL	866	HKY+I+G	Portik et al. (2011)
PNN	NPCL	950	HKY+G	Townsend et al. (2008)
PRLR	NPCL	530	HKY+G	Townsend et al. (2008)
SNCAIP	NPCL	474	HKY+I	Townsend et al. (2008)