

1.1. L'ESQUIZOFRÈNIA

1.1.1. ANTECEDENTS HISTÒRICS

La paraula esquizofrènia deriva dels termes grecs *skhizo* i *phren* que volen dir escissió i ment, respectivament. Tot i que el terme esquizofrènia va estar encunyat pel psiquiatre suís Eugen Bleuler l'any 1911, la condició a la que es refereix ha acompanyat a la humanitat des de fa molt de temps. Les descripcions més antigues d'aquest trastorn es troben en el papyrus d'Ebers (1550 aC). Les primeres teories sobre la base física de les malalties mentals, la integració del cos i la ment i els processos mentals no conscients van ser formulades per Plató (427-347 aC). Contemporàniament, Hipòcrates (470-? aC) proposà que els trastorns mentals tenen el seu origen en el cervell. Però no és fins a començaments del segle XVIII que es fan les primeres descripcions clíniques detallades de les malalties mentals.

La primera aproximació clínica a aquesta síndrome fou descrita pel psiquiatre alemany Emil Kraepelin (1856-1926) amb el nom de demència precoç. L'apreciació que en els malalts inclosos sota el diagnòstic de Kraepelin, predominaven més les característiques de dissociació i disgregació de la personalitat que les d'autèntica demència, va conduir a Bleuler a designar aquest grup de malalts com afectats d'esquizofrènia.

El criteri de diagnòstic i la lògica d'un diagnòstic diferencial pels trastorns mentals no apareix fins l'any 1948 amb la sisena edició de la Classificació Internacional de Malalties (International Classification of Disease, ICD-6) establert per l'Organització Mundial de la Salut. Paral·lelament, a l'any 1952, l'Associació Americana de Psiquiatria publicà la primera edició del Manual Diagnòstic i Estadístic de Trastorns Mentals (*Diagnostic and Statistical Manual of Mental Disorders*, DSM). Des d'aleshores, ambdós manuals han modificat i incorporat diagnòstics d'acord amb els avenços realitzats en psiquiatria. Aquests dos sistemes són els d'ús comú a tot el món. Les versions més recents ICD-10 (1992) i DSM-IV (1994) tenen definicions similars d'esquizofrènia i el diagnòstic que es realitza amb un i altre coincideix amb un elevat grau d'exactitud.

1.1.2. DEFINICIÓ

L'esquizofrènia es caracteritza per una alteració profunda del pensament i la percepció, afectant els atributs humans més fonamentals: llenguatge, pensament, percepció, afectivitat i sentit d'un mateix. La diversitat de símptomes freqüentment inclouen manifestacions psicòtiques com són sentir veus internes o experimentar sensacions no vinculades a cap font (al·lucinacions auditives i visuals) i tenir creences falses i persistents, com que el malalt és perseguit o es creu ser el centre de tot el passa al seu voltant (idees delirants).

No hi ha cap símptoma definitiu pel diagnòstic. Aquest es basa en la presentació d'una sèrie de signes i símptomes peculiars que han estat presents una part

significativa de temps durant el període d'un mes (o durant un període més breu si hi ha hagut tractament amb èxit) i amb alguns signes del trastorn que han persistit com a mínim 6 mesos. Aquests símptomes són: idees delirants, al·lucinacions, llenguatge desorganitzat, comportament greument desorganitzat o catatònic i símptomes negatius (pèrdua de les habilitats socials, falta d'interès, incapacitat d'experimentar plaer, incapacitat d'expressar emocions, pensament i llenguatge empobrits).

Es tracta d'un trastorn que sol ser crònic, recurrent i greu que té un gran impacte en cadascun dels aspectes de la vida de la persona que el pateix, així com també en l'entorn familiar i social. Les persones afectades pel trastorn se senten com si haguessin perdut la identitat, l'autonomia i la capacitat mental. El suïcidi és una causa de mort freqüent en aquests pacients: un 50% intenten suïcidar-se almenys una vegada a la vida i d'un 10 a un 15% moren per aquesta causa.

El tractament, per la majoria de malalts, combina l'administració de fàrmacs antipsicòtics amb la psicoteràpia. No existeix cap tractament curatiu per aquest trastorn. La tendència actual és la d'evitar l'hospitalització de llarga durada, per no allunyar el malalt del seu entorn social. L'hospitalització és indicada si l'agitació psicomotriu és incontrolable, hi ha perill d'autoagressió o heteroagressió i per establir la medicació.

1.1.3. EPIDEMIOLOGIA

L'esquizofrènia té una prevalença al voltant de l'1% en tots els àmbits geogràfics, culturals i socioeconòmics (Gottesman i Shields, 1992). Les primeres manifestacions d'aquesta malaltia apareixen entre el final de l'adolescència i el principi de l'edat adulta. En homes s'ha establert que l'edat d'inici oscil·la entre els 18 i els 25 anys, i en dones entre els 23 i 35 (Lewine, 1980).

Ja a l'any 1916, Ernst Rüdin va observar que l'esquizofrènia s'agregava en famílies (Rüdin, 1916). Els estudis de risc familiar duts a terme en els darrers 20 anys amb pacients amb esquizofrènia observen de forma consistent l'existència d'agregació familiar. Aquests estudis indiquen que els familiars de primer grau dels pacients tenen un risc d'afectació de 5 a 17 vegades major que els individus no relacionats (Kendler *et al*, 1985; Gershon *et al*, 1988; Gottesman i Shields, 1992). Aquesta agregació familiar és una indicació important d'una base genètica però no suficient, ja que altres components no genètics (ambientals, culturals) poden estar implicats en aquesta observació.

Els estudis de bessons han detectat una concordança del 5-15% i 40-50% en bessons dizigòtics i monozigòtics, respectivament (Farmer *et al*, 1987; Onstad *et al*, 1991; Gottesman i Shields, 1992). D'aquests estudis se'n desprèn que l'heretabilitat del trastorn és del 60-89% (Cannon *et al*, 1998; Cardno *et al*, 1999; Schultz i

Andreasen, 1999), recolzant la idea que hi ha factors genètics que contribueixen a la transmissió familiar del trastorn.

Els estudis de la descendència de bessons discordants per l'esquizofrènia indiquen que el risc d'afectació és el mateix pels fills de bessons monzigòtics discordants (17%). Mentre que en bessons dizigòtics, el risc és major pels fills del bessó malalt (17% versus 2%). Aquests resultats suggereixen que genotips que confereixen susceptibilitat a desenvolupar esquizofrènia poden no expressar-se (penetrança incompleta) i que la freqüència d'individus que no hereten la predisposició genètica però que desenvolupen la malaltia (fenocòpies) és baixa (Gottesman i Bertelsen, 1989). Aquestes dues característiques són importants a tenir en compte en fer les anàlisis dels estudis genètics ja que poden emascarar els resultats.

Els estudis de risc familiar, de bessons i d'adopció no són capaços de determinar l'existència d'un patró Mendelià de transmissió de la malaltia. El patró d'herència és complex i no està clar si l'agregació familiar en l'esquizofrènia és deguda a la contribució d'un gen major, varis gens o molts gens d'efecte menor. El model d'herència de la malaltia és encara desconegut.

L'esquizofrènia és sovint comòrbida amb altres condicions trastorn de pànic (10-20%) i trastorn obsessiu compulsiu (10-20%) (Craig *et al*, 2002). Un 50% dels pacients presenten algun trastorn d'abús de substàncies. I un 80-90% presenten dependència de nicotina.

1.1.4. ASPECTES NEUROBIOLÒGICS

El processament de la informació i, concretament, la tasca de correlacionar i integrar la informació emmagatzemada amb noves situacions en el context apropiat, sembla estar alterat en els malalts d'esquizofrènia. Les evidències d'un dèficit cognitiu en aquests malalts són múltiples (Elvevag i Goldberg, 2000), però el sistema o circuits neuronals subjacents són encara desconeguts. Els estudis de neuroimatge, neurofisiologia i anatomo-patològics *post-mortem* indiquen que l'esquizofrènia és una malaltia estructural i funcional del cervell que afecta varies regions corticals i subcorticals implicades en els aspectes cognitius, emocionals i de motivació de la conducta humana.

A nivell macroanatòmic, s'ha arribat al consens que els malalts d'esquizofrènia presenten una ampliació dels ventricles i una reducció del volum en àrees del sistema límbic (hipocamp i amígdala) i regions corticals (còrtex prefrontal) (Harrison, 1999 i 2004; McCarley *et al*, 1999; Pearlson i Marsh, 1999; Wright *et al*, 2000; Ho *et al*, 2003). Aquestes alteracions s'observen tant en malalts de llarga evolució com en primers episodis de malalts no medicats (Bogerts *et al*, 1991; DeLisi *et al*, 1991; Nopoulos *et al*, 1995). I, en un menor grau, també s'observen en familiars no afectats pel trastorn (Staal *et al*, 2000; Seidman *et al*, 2002). Alguns autors

consideren aquestes alteracions estructurals com a potencials 'marcadors biològics' de la malaltia.

Els estudis a nivell microanatòmic han detectat anomalies en la citoarquitectura neuronal, disminucions dels cossos neuronals, de la densitat i longitud de les espines dendrítiques i de proteïnes sinàptiques que serien, en part, responsables de les reduccions de les regions límbiques i corticals mencionades anteriorment (Harrison, 1999 i 2004; Honer *et al*, 1999; Glantz i Lewis, 2000; Chana *et al*, 2003). Estudis més recents convergeixen en la implicació de la glia (Cotter *et al*, 2001) i, en particular, de l'oligodendroglia i el procés de mielinització en l'esquizofrènia (Uranova *et al*, 2001 i 2004; Bartzokis, 2002; Hof *et al*, 2002 i 2003). Aquests resultats es veuen recolzats pels estudis d'expressió gènica, els quals es comenten més endavant, on es detecta una baixa expressió (o *downregulation*) de gens expressats per l'oligodendroglia i implicats en el procés de mielinització (Hakak *et al* 2001, Tkachev *et al*, 2003). En aquesta mateixa línia, s'han detectat nivells elevats de la proteïna S100B en sang (marcador de la integritat de les cèl.lules glials i dany cerebral) en malalts d'esquizofrènia no medicats (Lara *et al*, 2001; Rothermund *et al*, 2001). A més, les concentracions de S100B es mantenen elevades després del tractament i s'associen a una simptomatologia negativa (Rothermund *et al*, 2004).

Hi ha molta controvèrsia sobre si aquestes anomalies cerebrals fluctuen o són estàtiques amb l'aparició de la malaltia. S'accepta que el seu origen té lloc durant les etapes primerenques del neurodesenvolupament. Tot i que també es disposen de dades que apunten a l'existència d'un procés subtil de neurodegeneració amb la progressió de la malaltia. Les hipòtesis etiopatogèniques formulades més recentment per l'esquizofrènia integren ambdós processos, de neurodesenvolupament i neurodegeneració.

1.1.5. TEORIES CAUSALS

El procés patofisiològic i els factors etiològics causants de l'esquizofrènia són, doncs, desconeguts. L'heterogeneïtat que presenta l'esquizofrènia, manifestada a través del ventall de símptomes, resposta al tractament i evolució, complica la formulació de models teòrics per la mateixa (Andreasen i Carpenter, 1993). S'ha postulat que aquesta heterogeneïtat seria resultat de múltiples etiologies responsables de diferents mecanismes patofisiològics (Tsuang *et al*, 1990) o bé, d'un únic procés etiopatogènic (Crow, 1990), que donat el context individual donaria lloc a les diferents manifestacions del trastorn (Carpenter i Kirkpatrick, 1988). A més, com ja s'ha comentat prèviament, els factors ambientals contribueixen a la variabilitat de la manifestació i evolució de la malaltia.

Les hipòtesis etiopatogèniques per l'esquizofrènia van des de la concepció més organicista fins les psicològiques i socials. Les teories psicosocials postulen que són determinats successos quotidians estressants els que desborden la capacitat

defensiva de l'individu en el seu rol social i desencadenen la malaltia; un canvi ambiental, l'estrès o una tensió ambiental excessiva, o una capacitat defensiva i suport social escàs poden donar lloc a una crisi i exacerbació. Aquestes teories, que foren influents a la meitat del segle XX, han estat difícils d'avaluar científicament. Les teories actuals consideren que l'esquizofrènia és el resultat de la combinació d'uns factors genètics i uns factors ambientals. El gen o gens implicats en la malaltia conferirien una predisposició que acabaria en manifestació clínica després de l'exposició a determinats factors ambientals. Des d'aquest punt de vista, l'esquizofrènia s'inclou dins del grup de les malalties complexes o multifactorials (Lander i Schork, 1994).

Algunes de les hipòtesis més populars per l'esquizofrènia són les que impliquen un canvi en la neurotransmissió de dopamina, glutamat, acetilcolina, serotonina, GABA i/o noradrenalina (Owen i Simpson, 1994), la hipòtesi del neurodesenvolupament (Weinberger, 1986), la hipòtesi de la transmetilació (Smythies, 1983), la hipòtesi viral (Torrey i Peterson, 1976) i la hipòtesi dels fosfolípids de membrana (Pettegrew *et al*, 1991). Totes elles es recolzen de dades epidemiològiques i derivades d'estudis de resposta a tractament, farmacològics, de neuroimatge i de determinacions d'hormones i neurotransmissors. Les hipòtesis dopaminèrgica i del neurodesenvolupament són potser les més acceptades. Però ni aquestes dues ni la resta són incompatibles entre elles.

Les hipòtesis més recentment formulades impliquen una alteració en la glia i proporcionen una explicació heurística a les observacions neurobiològiques i genètiques realitzades en l'esquizofrènia. A continuació se'n fa una breu descripció.

Hipòtesi de la deficiència de factors de creixement glials i desestabilització sinàptica

Aquesta postula que una deficiència funcional dels factors de creixement de la glia i de factors de creixement sintetitzats per la glia comportaria una desestabilització de la sinapsi, que seria la causa principal de l'aparició de l'esquizofrènia (Moises 2001, Moises *et al*, 2002). La hipòtesi suggereix que les cèl.lules glials serien el punt de trobada de les interaccions gen-ambient implicades en l'esquizofrènia. La vulnerabilitat a desestabilitzar la sinapsi i, conseqüentment, desenvolupar l'esquizofrènia, es deuria per tenir uns nivells basals baixos de factors de creixement glials i d'estabilitat sinàptica determinats genèticament i epigenètica.

Hipòtesi de la disfunció de l'oligodendroglia i mielinització

Aquesta postula que una alteració en el procés de mielinització i/o el seu manteniment afectaria profundament la connexió neuronal que resultaria en l'aparició de l'esquizofrènia (Bartzokis, 2002; Hof *et al*, 2002). La mielinització és un procés dinàmic que s'estén fins l'edat adulta (Benes *et al*, 1994). En particular, la mielinització entre les estructures límbiques i el còrtex prefrontal dorsolateral sembla tenir lloc durant la segona i tercera dècada de vida, període de vulnerabilitat per

l'aparició de l'esquizofrènia (Benes, 1989; Paus *et al*, 1999). Totes aquelles interaccions gen-ambient que alteressin el funcionament dels oligodendròcits i el procés de mielinització durant el neurodesenvolupament i maduració del cervell desencadenarien el trastorn. La dispersa distribució de l'oligodendroglià explicaria, en part, les diferents troballes neurobiològiques i també el ventall de símptomes que presenten aquests malalats.

1.1.6. ESTUDIS GENÈTICS

Com se'n deriva de les dades epidemiològiques revisades anteriorment, el component genètic de l'esquizofrènia és considerable. Ara bé, no presenta un patró d'herència Mendeliana clàssica i, per això, és considerada una malaltia complexa. Les característiques que explicarien la dificultat en caracteritzar els gens en les malalties complexes són les següents:

- *Herència oligogènica o poligènica*. Presència de mutacions en varis o múltiples gens confereix susceptibilitat.
- *Penetrança incompleta*. Presència de la variant de susceptibilitat, però es requereix la interacció amb determinats factors ambientals (interaccions gen-ambient) o altres variants gèniques (interaccions gen-gen) per l'aparició dels símptomes.
- *Heterogeneïtat genètica*. Mutacions en qualsevol dels gens implicats poden resultar en el mateix fenotip.
- *Factors ambientals* comuns o compartits pels diferents membres familiars que faciliten l'aparició d'agregació familiar.
- *Fenocòpies*. Individus que no hereten la predisposició genètica desenvolupen la malaltia.
- *Elevada freqüència dels al·lels de susceptibilitat* de causar la malaltia en la població general.

Existeixen moltes evidències que qualsevol d'aquestes característiques pot estar operant en l'esquizofrènia. En són un clar exemple la manca de *LOD scores* positius en els estudis de lligament o la incapacitat de replicar-ne els resultats, junt amb la incapacitat de replicar els resultats positius de la gran majoria dels estudis d'associació.

S'han identificat varis mecanismes d'herència complexa o no mendeliana, i inclouen l'herència mitocondrial, l'empremta (*imprinting*), la disomia uniparental i l'anticipació. Fins el moment, no hi ha prou evidència de l'acció d'aquests mecanismes en l'esquizofrènia.

Els estudis d'investigació de *loci* de susceptibilitat per l'esquizofrènia es poden agrupar en sis categories (Taula I). Aquests no són mútuament excloents i la utilització d'un o altre mètode dependrà del tipus i la grandària de la mostra, el tret fenotípic estudiat i mecanisme o model genètic proposat. A continuació es fa una

breu revisió de l'estat actual de la investigació de gens de susceptibilitat per l'esquizofrènia.

Taula I. Resum de les aproximacions emprades per estudiar les regions cromosòmiques i/o gens implicats en l'esquizofrènia.

Mètode	Estudis d'anàlisi de lligament		Estudis d'associació		Estudis citogenètics i anomalies cromosòmiques	Estudis d'expressió gènica per microxips
	Model dependent (paramètric)	Model independent (no paramètric)	Basats en famílies	Basats en població		
Nom alternatiu	Anàlisi del <i>LOD score</i>	<i>Allele sharing method</i>	Test de desequilibri de la transmissió (TDT); Test de risc relatiu d'haplotip (HRR)	Estudi de casos i controls		Anàlisi per <i>microarray</i>
Objectiu	Localitzar locus implicat en la malaltia		Determinar si un al·lel o haplotip concret està associat o en desequilibri de lligament amb la malaltia		Localitzar alteracions cromosòmiques associades a la malaltia	Detectar si l'expressió determinada d'un gen s'associa a la malaltia
Mostra	Genealogies grans amb múltiples individus afectats	Parelles de germans afectats o pare o mare i fill afectats	Individu afectat i pares heterozigots pels marcadors estudiats	Individus afectats (casos) i individus no afectats (controls)	Casos i controls; genealogies	Casos i controls
Utilització	Anàlisi de tot el genoma (estudis globals del genoma) o d'un locus específic		Anàlisi d'un locus o gen candidat		Anàlisi d'una regió cromosòmica o gen candidat	Anàlisi de milers de gens simultàniament
Avantatge	Capaç de detectar l'efecte d'un locus en gran distàncies genètiques		Utilització dels genotips parentals com a control intern	Capaç de detectar gens d'efecte moderat o lleu	Capaç d'acotar la regió amb el gen o gens implicats	Capaç d'identificar nous gens candidats en absència d'hipòtesi preconcebuda
Desavantatge	Dificultat de trobar grans genealogies	Poc sensible al definir la localització del locus implicat	Dificultat de trobar ambdós pares	Estratificació de la mostra	Poc sensible	Falsos positius per manca de controls aparellats

Estudis de lligament

Els estudis globals del genoma en famílies amb esquizofrènia han trobat valors suggestius de lligament amb mètodes paramètrics (depenen d'un model genètic d'herència) i no paramètrics (no depenen d'un model genètic d'herència) en més de la meitat dels 23 cromosomes (Taula II). Molts d'aquests resultats positius no han estat replicats quan s'estudien cohorts més grans o les mateixes regions són estudiades per grups independents. Les explicacions a aquesta disparitat de resultats inclouen un poder insuficient de les mostres per a detectar lligament, heterogeneïtat genètica, manca d'una definició acurada del fenotip o bé existència de múltiples al·lells de susceptibilitat amb un efecte moderat. Si la predisposició a la malaltia és resultat de la combinació de diferents gens d'efecte moderat, és d'esperar que els resultats positius dels estudis de lligament siguin costosos de replicar.

Taula II. Principals estudis de lligament genètic en l'esquizofrènia.

Estudi ^a	Regió cromosòmica	Mostra	Diagnòstic
Blouin <i>et al</i> , 1998 (452 marcadors)	8p21-22 13q32 14q13 22q11	54 famílies (origen divers)	DSM-III-R
Levinson <i>et al</i> , 1998 (310 marcadors)	2q 4q 9 10q 11q	43 famílies (origen divers)	DSM-III-R
Hovatta <i>et al</i> , 1999 (351 marcadors)	1q32 4q34 9q21 Xp11	20 famílies de Finlàndia	DSM-IV
Williams <i>et al</i> , 1999 (229 marcadors)	4p 18q Xcen	196 parelles de germans afectes	DSM-IV
Gurling <i>et al</i> , 2001 (365 marcadors)	1q33.2 4q13-31 8p22.1-22 11q21 11q23.3-24	13 famílies	RDC ^b DSM-III
Paunio <i>et al</i> , 2001 (315 marcadors)	5q 2q	238 famílies de Finlàndia	DSM-IV
DeLisi <i>et al</i> , 2002 (396 marcadors)	2cen 10p15-p13 22q12	382 parelles de germans afectes	
Devlin <i>et al</i> , 2002 (546 marcadors)	3q 5q 9p	5 famílies de Palau	DSM-IV
Stefansson <i>et al</i> , 2002 (950 marcadors)	8p12-21	33 famílies d'Islàndia	RDC
Straub <i>et al</i> , 2002 (166 marcadors)	2p14-13 4q24-32 5q21-31 6p24-21 8p22-21 10p15-p11	270 famílies d'Irlanda	DSM-III-R

^a S'han inclòs aquells estudis amb valors de *LOD score* i *NPL* superiors a 1. Aquests són els estadístics que expressen la probabilitat de què un marcador cosegregui o estigui lligat amb el tret o malaltia d'estudi. Els valors que oscil·len entre 1 i 3 són considerats 'valors suggestius de lligament' i la seva rèplica en mostres independents és considerada una prova de confiança (Lander i Kruglyak, 1995). Per la confirmació de lligament d'una determinada regió cromosòmica o marcador es consideren valors superiors a 3. En el cas de les malalties complexes i en concret, l'esquizofrènia, la majoria dels estudis de lligament obtenen valors entre 1-3; ^bRDC, Criteris de Diagnòstic en Investigació (*Research Diagnostic Criteria*).

Estudis d'associació

S'ha postulat que els estudis d'associació són més efectius per detectar gens de susceptibilitat d'efecte moderat o lleu respecte els estudis de lligament (Risch i Merikangas, 1996). Ara bé, per l'estudi global del genoma per associació es requereix d'analitzar milers de marcadors, la qual cosa no és viable actualment. I, d'altra banda, tampoc és clar quin és el nombre de marcadors necessaris. És per això que

els investigadors s'han centrat en l'estudi de gens o regions candidates. Un gen o una regió es defineix com a candidat si es troba en un segment cromosòmic que ha estat prèviament implicat mitjançant el mapatge posicional i/o la seva funció és rellevant dins del marc d'una de les hipòtesis etiopatogèniques de la malaltia. Donat el nombre de regions cromosòmiques implicat en l'esquizofrènia i, que aquestes estan poc acotades, el número de gens candidats és extremadament elevat. Com ja s'ha mencionat, són moltes les hipòtesis etiopatogèniques formulades per l'esquizofrènia i la tria de gens candidats és igualment difícil en aquesta situació. La Taula III recull una llista parcial amb gens candidats per l'esquizofrènia.

Taula III. Principals gens estudiats com a candidats a participar en la etiologia de l'esquizofrènia.

Hipòtesi de treball	Gen	Proteïna
Hipòtesis que impliquen canvi en la neurotransmissió	<i>DR</i>	Receptors dopamina (D1, D2, D3, D4)
	<i>HTR</i>	Receptors serotonina (HTR1A, 1B, 2C)
	<i>SLC6A4</i>	Serotonin transporter
	<i>GRIN2B</i>	Glutamate receptor
	<i>COMT</i>	Cathecol-O-methyltransferase
	<i>MAO</i>	Monoamina oxidases A i B
Hipòtesi del neurodesenvolupament	<i>CNTF</i>	Ciliary neurotrophic factor
	<i>BDNF</i>	Brain derived neurotrophic factor
	<i>NT-3</i>	Neurotrophin 3
	<i>TRKs</i>	Receptors de neurotrofines (TRKA, TRKB, TRKC)
	<i>RELN</i>	Reelin
Hipòtesi de la disfunció de l'oligodendroglia i mielinització	<i>GSN</i>	Gelsolin
	<i>MAG</i>	Myelin associated glycoprotein
	<i>TF</i>	Transferrin
	<i>MAL</i>	T-cell differentiation protein
Hipòtesi de la deficiència de factors de creixement glials i desestabilització sinàptica	<i>NRG</i>	Neuregulin (NGR1, NGR2, NGR3, NGR4)
	<i>ErbB</i>	Receptors de NRG (ErbB1, ErbB2, ErbB3, ErbB4)
	<i>GDNF</i>	Glial cell line derived neurotrophic factor
	<i>IGF1</i>	Insuline-growth factor I
	<i>EGF</i>	Epidermal growth factor
Altres	<i>RGS4</i>	G-protein signaling 4
	<i>NOTCH4</i>	Notch homolog 4 (Drosophila)
	<i>IL</i>	Interleuquina (IL2, IL12)
	<i>APO</i>	Apolipoproteïnes (Apo E, Apo D, Apo L)
	<i>MTHFR</i>	Methylen tetrahydrofolate reductase
	<i>SCA1</i>	Ataxia spinocerebelar I

La majoria dels estudis d'associació de gens candidats s'han centrat en gens implicats en la neurotransmissió dopaminèrgica i serotoninèrgica i en la hipòtesi del neurodesenvolupament. La majoria d'aquests estudis han analitzat un únic polimorfisme per locus o gen candidat. Les interpretacions que deriven d'un resultat positiu en aquest tipus d'estudi poden ser dues: 1. l'al·lel estudiat comporta un canvi biològic i és la variant causant o de susceptibilitat (associació directa); 2. l'al·lel estudiat es troba tan proper a l'al·lel causant que ambdós s'hereten conjuntament, és a dir, hi ha associació mantinguda per un desequilibri de lligament entre *loci* (associació indirecta).

A l'igual que pels estudis de lligament, els resultats positius dels estudis d'associació tampoc s'han pogut replicar de forma consistent. Aquestes discrepàncies rauen sovint en una selecció inadequada de la mostra de casos i controls. Això pot donar lloc a:

- i. Falsos positius (error de tipus I) deguts a la presència d'estratificació en la mostra o un nivell de significació inadequada.
- ii. Falsos negatius (error de tipus II) deguts a una mida insuficient de la mostra, presència d'heterogeneïtat genètica o diferències en la definició del fenotip.

Malgrat aquesta disparitat de resultats, estudis recents han suggerit els gens *dysbindin* (*DTNBP1*) i *neuregulin 1* (*NRG1*) com a *loci* de susceptibilitat per l'esquizofrènia (Owen *et al*, 2004). Ambdós gens es localitzen en regions cromosòmiques les quals havien estat prèviament implicades mitjançant estudis de lligament. *DTNBP1* es localitza en la regió 6p22.3 i *NRG1* en la regió 8p22-p11. Malgrat la identificació de múltiples polimorfismes i haplotips específics de risc, els mecanismes patogènics són encara desconeguts.

Estudis de citogenètica i anomalies cromosòmiques

S'han descrit un gran nombre d'anomalies cromosòmiques en l'esquizofrènia; però tres s'han associat a aquest trastorn: les microdeleccions cromosòmiques de 22q11.2 o síndrome de la deleció de 22q (22qDS), la inversió pericèntrica del cromosoma 9 i la translocació (1;11)(q41;q14.3). Tots tres resultats es veuen recolzats per estudis de lligament.

Només la síndrome 22qDS, coneguda com a síndrome velo-cardio-facial (VCF) o de Digeorge, presenta un fenotip específic associat. És un trastorn molt variable que presenta una prevalença d'esquizofrènia del 24% (Murphy *et al*, 1999; Murphy, 2002). Tenint en compte que la prevalença de la síndrome VCF és de 0.03%, s'ha estimat que un 0.5–2% dels malalts d'esquizofrènia tindrien una microdeleció a 22q11 i que representaria un subtipus genètic de la malaltia d'inici en l'edat adulta (Bassett i Chow, 1999).

Estudis d'expressió gènica

Els estudis recents d'expressió gènica que utilitzen la tecnologia dels *microarrays* mesuren simultàniament els nivells d'expressió de milers de gens. Aquesta aproximació permet la identificació de gens amb una expressió diferencial entre teixits de malalts i individus controls. Aquests tipus d'estudi són una aproximació complementària als estudis d'anàlisi de lligament i associació, ja que analitzen milers de gens i poden suggerir nous mecanismes moleculars i gens candidats implicats en la malaltia sense la formulació preconcebuda d'hipòtesis etiopatogèniques.

Els estudis realitzats fins el moment s'han centrat principalment en l'estudi de l'expressió gènica diferencial del còrtex prefrontal dorsolateral de malalts

d'esquizofrènia (Mirnics *et al*, 2000; Hakak *et al*, 2001; Mimmack *et al*, 2002; Vawter *et al*, 2002). Donades les diferències metodològiques en aquests treballs (com per exemple, en els criteris de diagnòstic, preparació de la mostra i extracció del mRNA, tipus de plataforma o *microarray* i anàlisi de les dades), les conclusions que se'n deriven són diferents. Mirnics i col.laboradors detecten diferències en gens implicats en la funció secretora presinàptica, per la qual cosa proposen que l'esquizofrènia és una malaltia de la sinapsi. Hakak i col.laboradors observen una *downregulation* de gens implicats en el procés de mielinització i expressats de forma majoritària en oligodendròcits. Aquests resultats han estat recolzats per estudis més recents (Tkachev *et al*, 2003). Aquests autors suggereixen que part de les alteracions cerebrals observades en els malalts d'esquizofrènia són degudes a una alteració en la funció dels oligodendròcits i els processos de mielintzació. Mimmack i col.laboradors observen una sobre-expressió (*upregulation*) de tres membres de la família de les apoliproteïnes L (apo L) que localitzen en el cromosoma 22q12, regió propera a la de la síndrome VCF. De moment, es desconeix quina és la funció de les apo L en cervell.

Estudis d'endofenotips

Els endofenotips són trets o característiques mesurables que reflecteixen un genotip subjacent i que estarien més relacionats amb el genotip que amb la categoria diagnòstica del trastorn (Gottesman i Gould, 2003). Aquests trets poden implicar mecanismes neurofisiològics, bioquímics, endocrinològics, neuroanatòmics, o neuropsicològics i han de complir amb els criteris d'heretabilitat, estabilitat, especificitat i replicabilitat. Els endofenotips s'identifiquen amb més freqüència en els malalts i també sovint en familiars no afectes, indicant l'existència de penetrança incompleta. Els endofenotips identificats i lligats genèticament a l'esquizofrènia són els següents:

- i. Dèficit en els moviments de persecució ocular (*smooth pursuit eye tracking*) mapat al cromosoma 6p, regió implicada prèviament en l'esquizofrènia (Arolt *et al*, 1999).
- ii. Inhibició del potencial evocat P50 lligat als cromosomes 15 i 22 (Freedman *et al*, 1999; Myles-Worsley *et al*, 1999).
- iii. Disminució de l'amplitud potencial evocat P300 lligat al cromosoma 1q42 (Blackwood *et al*, 2001).
- iv. Dèficit en la memòria de treball associada al genotip Val158Met del gen *COMT* (Goldberg *et al*, 2003).

Models animals

L'esquizofrènia és una malaltia característica dels humans. Aquesta no s'observa o no es pot mesurar en animals. No obstant això, la utilització de models animals, en particular el múrid, pot ser molt informativa per identificar gens implicats en trets fenotípics considerats de rellevància per l'esquizofrènia. Com per exemple, l'estudi amb múrids de la inhibició prepuls (*prepulse inhibition*, PPI) de la resposta a

sobresalt (*startle response*, SR) (Joober *et al*, 2002; Palmer *et al*, 2003). La PPI és un comportament incondicionat que té lloc de forma natural en un ampli nombre d'espècies animals en les quals, la presentació d'un lleuger estímul sensorial (estímul prepols) unes mil·lèsimes de segons abans de la presentació d'un estímul més intens que indueix la SR, disminueix la magnitud d'aquesta resposta. S'hipotitza que la presentació del prepols activa un procés inhibitor que minva la SR. Malalts d'esquizofrènia així com rates tractades amb fàrmacs d'efecte psicòtic presenten un dèficit en la PPI, que és revertit per l'administració de fàrmacs antipsicòtics.

1.1.7. CROMOSOMA 6 I ESQUIZOFRÈNIA

Una de les regions cromosòmiques més estudiada en l'esquizofrènia és la regió 6p (locus SCZD3 [MIM 60051]). Els primers estudis de lligament amb resultats positius i replicats per grups independents són els de Wang *et al* (1995) i Straub *et al* (1995), involucrant la regió 6p24-p21. Evidències addicionals de lligament en aquesta regió deriven de diferents estudis (Moises *et al*, 1995; Schwab *et al*, 1995; Antonarakis *et al*, 1995; Bailer *et al*, 2000; Schwab *et al*, 2002; Straub *et al*, 2002b; Schwab *et al*, 2003). Tot i que d'altres no han pogut replicar els resultats (Pulver *et al*, 1995; Hovatta *et al*, 1998; Maziade *et al*, 2001; DeLisi *et al*, 2002).

Els marcadors amb valors suggestius de lligament a 6p per l'esquizofrènia es distribueixen per una regió àmplia de més de 25 Mb, entre els marcadors D6S309 a D6S282 (Taula IV). No obstant, sembla que aquests s'agrupen en 4 subregions: 6p24 a prop de D6S309-D6S296, 6p24 a prop de D6S940/D6S470, 6p23-22 a prop de D6260 i 6p21.3 on es troba el locus del sistema HLA (de l'anglès *Human Leukocyte Antigen*).

S'han dut a terme varis estudis d'associació amb gens candidats localitzats a 6p i algun d'ells amb resultats positius (Taula IV). Recentment, un estudi d'associació basat en famílies ha detectat desequilibri de lligament entre el gen *DTNBP1* i un diagnòstic estricte d'esquizofrènia (Straub *et al*, 2002). Aquests resultats han estat replicats en diverses poblacions mitjançant estudis basats en famílies (Schwab *et al*, 2003; Tang *et al*, 2003; van den Oord, *et al*, 2003; Kirov *et al*, 2004) i estudis de casos i controls (Van den Bogaert, *et al*, 2003; Williams *et al*, 2004). Es tracta de la primera replicació d'un estudi de desequilibri de lligament en esquizofrènia. El patró de desequilibri de lligament és consistent amb la presència de més d'un al·lel de susceptibilitat depenent de la població estudiada. De moment, no s'han identificat variants funcionals en les regions codificants de *DTNBP1*, però se sospita que les variants afectarien la seva expressió i processament proteic. En aquesta línia, s'han observat diferències subtils en l'expressió de *DTNBP1* en la regió de l'hipocamp en malalts d'esquizofrènia, però no s'han associat a cap dels polimorfismes identificats (Talbot *et al*, 2004).

Taula IV. Resultats més significatius dels estudis genètics (de lligament i associació) i gens candidats en la regió 6p24-21 per l'esquizofrènia.

Regió cromosòmica	Referència	Marcador genètic	Posició en el mapa ^a	Gen candidat ^b	Referència
6p24	Lindholm <i>et al</i> , 1999 Bailer <i>et al</i> , 2000	D6S309	14.07	<i>NRN1</i> ^c	Moises <i>et al</i> , 2002
	Antonarakis <i>et al</i> , 1995 Straub <i>et al</i> , 1995 Maziade <i>et al</i> , 1997 Hwu <i>et al</i> , 2000	D6S296	14.07	<i>TFAP2A</i> ^d	Kawanishi <i>et al</i> , 2000
	Maziade <i>et al</i> , 1997	D6S277	14.61		
	Straub <i>et al</i> , 2002a	D6S940	16.84		
	Schizophrenia Linkage Collaborative Group for Chromosomes 3, 6 and 8, 1996	D6S470	18.22		
6p23	Lindholm <i>et al</i> , 1999	D6S1578	26.71	<i>SCA1</i> ^e	Wang <i>et al</i> , 1996
	Straub <i>et al</i> , 2002a	D6S259	27.78	<i>DTNBP1</i> ^f	Straub <i>et al</i> , 2002b Schwab <i>et al</i> , 2003 Williams <i>et al</i> , 2004
	Wang <i>et al</i> , 1995 Schwab <i>et al</i> , 2000	D6S260	29.93		
6p22	Schwab <i>et al</i> , 1995 Moises <i>et al</i> , 1995 Schwab <i>et al</i> , 2000	D6S274	32.6		
	Straub <i>et al</i> , 1995 Straub <i>et al</i> , 2002a	D6S285	34.23		
6p21	Moises <i>et al</i> , 1995	D6S291	49.5		
	Arolt <i>et al</i> , 1996	D6S271	66.37	<i>NOTCH4</i>	Wei i Hemmings, 2000
	Arolt <i>et al</i> , 1996	D6S282	66.37	<i>TNF</i> ^g	Boin <i>et al</i> , 2001 Meira-Lima <i>et al</i> , 2003

^a Posició en el mapa dels marcadors en cM respecte pter (final del braç curt del cromosoma). Dades del mapa Mfd98 del Marshfield Medical Research Center (<http://research.marshfieldclinic.org/genetics/>); ^b Gens candidats de la regió 6p que han estat estudiats en l'esquizofrènia. A excepció de *NRN1* que ha estat recentement proposat (Moises *et al*, 2002); ^c *NRN1*, neuritin; ^d *TFAP2A*, transcription factor ap2-alpha; ^e *SCA1*, spinocerebellar ataxia 1; ^f *DTNBP1*, dystrobrevin-binding protein 1; ^g *TNF*, tumor necrosis factor alpha.

1.2. EL RECEPTOR DOMINI DISCOIDINA 1 (DDR1)

1.2.1. PER QUÈ ESTUDIAR *DDR1* COM A GEN CANDIDAT PER L'ESQUIZOFRÈNIA?

A finals de l'any 1997, en la recerca de nous gens candidats per l'esquizofrènia, diferents treballs ens van fer suggerir l'estudi del gen *DDR1* en el context de la hipòtesi del neurodesenvolupament per l'esquizofrènia. Aquestes troballes es poden resumir en:

- i. Localització de *DDR1* a 6p24-21 (Valent *et al*, 1996), regió candidata a contenir un o més gens de susceptibilitat per l'esquizofrènia, com s'ha comentat prèviament.
- ii. Caracterització de *DDR1* com a receptor tirosina quinasa i amb homologia elevada amb el domini tirosina quinasa dels receptors de neurotrofines (Di Marco *et al*, 1993).
- iii. Expressió de *DDR1* en sistema nerviós central (SNC) i sobretot durant el neurodesenvolupament en ratolí i rata (Zerlin *et al*, 1993; Sanchez *et al*, 1994).
- iv. Modulació de la proliferació i diferenciació de cèl·lules neuroepiteliales i astroglija per col·lagen, lligand fisiològic de *DDR1* (Goetschy *et al*, 1987; Shrivastava *et al*, 1997; Vogel *et al*, 1997; Ali *et al*, 1998).

1.2.2. FAMÍLIA DELS RECEPTORS DOMINI DISCOIDINA

Els receptors domini discoidina 1 i 2 (*DDR1* i *DDR2*) constitueixen una família de receptors tirosina quinasa (RPTK, de l'anglès *receptor protein-tyrosine kinase*) per les seves homologies estructurals i funcionals. Els RPTK són proteïnes transmembrana implicades en el control i regulació dels processos cel·lulars de proliferació, diferenciació, transformació, senescència i apoptosi durant el desenvolupament i en la vida adulta. Les aproximadament 20 famílies de RPTKs identificades es caracteritzen per tenir un domini extracel·lular d'unió al lligand, un únic domini transmembrana i un domini intracel·lular tirosina quinasa (Figura 1). La transmissió del senyal dels RPTKs requereix de l'oligomerització del receptor, com a resultat de la unió del lligand, la qual promou l'autofosforilació de residus tirosina de la regió citoplasmàtica del receptor (Heldin, 1995). La fosforilació específica dels residus tirosina estabilitza la conformació activa del receptor i permet l'ancoratge de molècules de transmissió del senyal com proteïnes amb dominis SH2 i PTB, les quals reclutaran proteïnes addicionals activant les cascades corresponents de transmissió del senyal cap a l'interior cel·lular (Blume-Jensen i Hunter, 2001).

Els DDRs es distingeixen de la resta de RPTKs per tenir en la regió extracel·lular un domini discoidina homòleg a la proteïna discoidina I de l'ameba *Dictyostelium discoideum*. Discoidina I és una lectina (d'unió a galactosa i N-acetil-galactosamina) amb un paper rellevant en el manteniment de la morfologia cel·lular, organització citoesquelètica i agregació cel·lular de *Dictyostelium* quan hi ha manca de nutrients al medi (Devine i Williams, 1982; Alexander *et al*, 1992). En humans, regions homòlogues al domini discoidina s'han identificat en proteïnes implicades en

adhesió cel.lular o interaccions cèl.lula-cèl.lula (Baumgartner *et al*, 1998) (Taula V). Aquestes proteïnes inclouen receptors de membrana (neuropilin 1 i 2, caspr/neurexin IV) i proteïnes de secreció (factors de coagulació V i VIII, AEBP1, RS1, MFGE8). A més, en ratolins, resultats recents suggereixen que els dominis discoidina de la proteïna SED1 estan implicats en la comunicació cèl.lula-matriu extracel.lular durant la unió de l'espermatozou a l'oòcit (Ensslin i Shur, 2003).

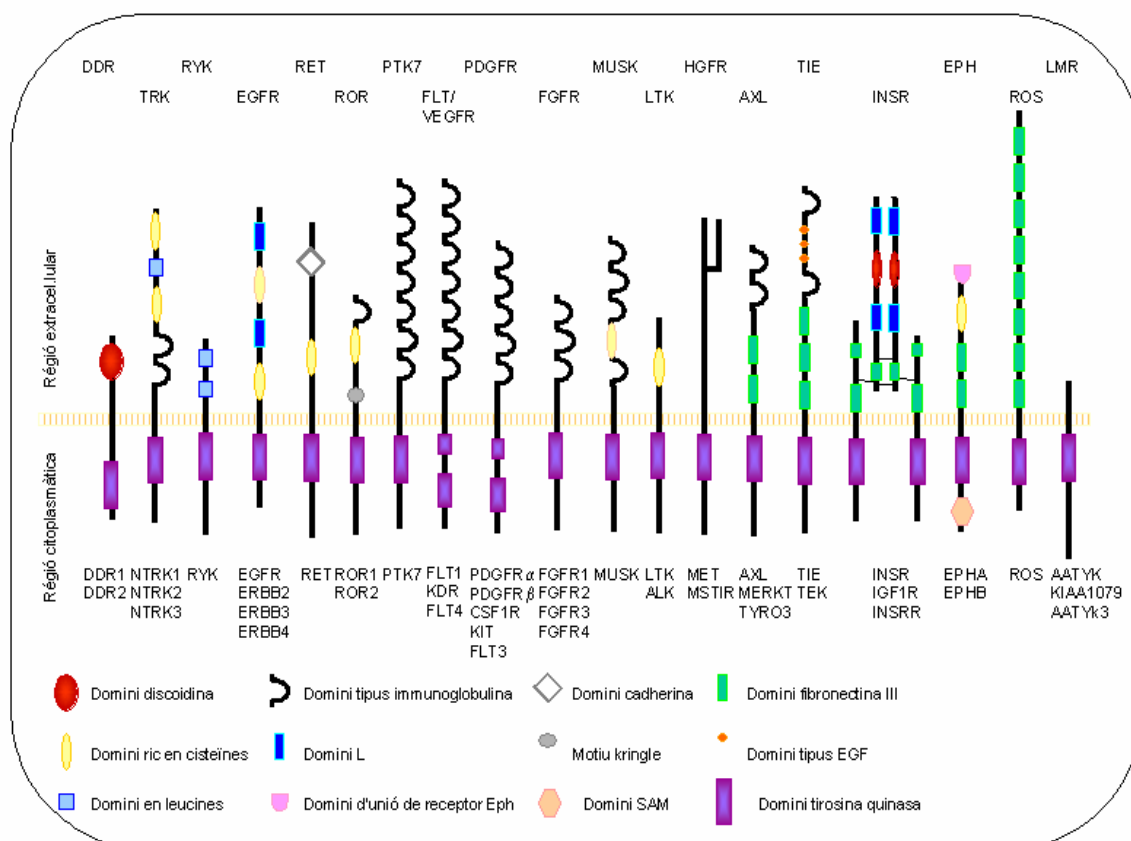


Figura 1. Esquema de les famílies estructurals de receptors tirosina quinasa en humans. El receptor prototípic de cada família s'indica a sobre el receptor i, a sota, els receptors representatius identificats. Els diferents dominis estructurals s'especifiquen a la dreta. Les abreviatures dels receptors prototípics són: DDR: receptor domini discoidina; TRK: receptor tirosina quinasa; RYK: receptor relacionat amb tirosina quinases; EGFR: receptor de creixement epidèrmic; ROR: receptor òrfan; PTK7: proteïna tirosina quinasa 7; FLT: receptor relacionat a fms (o receptor de factor de creixement endotelial); PDGFR: receptor de factor de creixement derivat de plaquetes; FGFR: receptor de creixement de fibroblasts; MUSK: receptor tirosina quinasa específica de múscul; LTK: tirosina quinasa de leucòcits; HGFR: receptor de creixement d'hepatòcits; TIE: receptor tirosina quinasa de cèl.lules endotel·lials; INSR: receptor d'insulina; EPH: receptor d'ephrin; ROS: RPTK expressat en determinades cèl.lules epitel·lials; LMR: lemur (modificada de Blume-Jensen i Hunter, 2001).

Taula V. Proteïnes amb domini discoidina en humans.

Gen	Proteïna	Núm. dominis ^a	Lligand	Funció	Referència
Receptors de membrana					
<i>CASPR4</i>	Proteïna de reconeixement cel.lular 4	1	Formació d'un complex trimèric amb contactina i neurofascina	Adhesió axó-glia en la regió paranodal	Peles <i>et al</i> , 1997
<i>DDR1</i>	Receptor domini discoidina 1	1	Col.lagen	Proliferació, migració, diferenciació cel.lular	Jonson <i>et al</i> , 1993
<i>DDR2</i>	Receptor domini discoidina 2	1	Col.lagen	Interacció cèl.lula-MEC ^b	Lai i Lemke, 1994
<i>NRP1</i>	Neuropilin-1	2	Semaforines classe 3 i VEGF ₁₆₅ ^c	Adhesió cel.lular	He i Tessier-Lavigne, 1993
<i>NRP2</i>	Neuropilin-2	2	Semaforines classe 3 i VEGF ₁₆₅	Adhesió cel.lular	Kolodkin <i>et al</i> , 1997
Proteïnes de secreció					
<i>AEBP1</i>	Proteïna 1 d'unió a AE (proteïna tipus carboxipeptidasa aòrtica)	1	Matriu extracel.lular (desconegut)	Desenvolupament embrionari i reparació de teixits	Layne <i>et al</i> , 1998
<i>F5</i>	Factor de coagulació V	2	Factor Xa	Cofactor del Factor Xa; unió a fosfolípids	Kane i Davie, 1986
<i>F8</i>	Factor de coagulació VIII	2	Factor IXa	Cofactor del Factor IXa; unió a fosfolípids	Kane i Davie, 1986
<i>MFGE8</i>	Proteïna grassa de la llet – factor de creixement epidèrmic 8	2	Aminofosfolípids (cèl.lules apoptòtiques)	Factor d'unió entre cèl.lules apoptòtiques i fagòcits	Hanayama <i>et al</i> , 2002
<i>RS1</i>	RS1	2	(desconegut)	Proteïna d'adhesió cel.lular	Sauer <i>et al</i> , 1997

^a Número de dominis discoidina; ^b MEC: matriu extracel.lular; ^c VEGF₁₆₅: isoforma 165 del factor de creixement de l'endoteli vascular.

Fins el present moment, el col.lagen ha estat identificat com a lligand fisiològic dels DDRs (Shrivastava *et al*, 1997; Vogel *et al*, 1997), tot i que no es descarta la presència d'altres lligands o cofactors. El DDR1 s'activa per la unió al col.lagen de tipus I a VI i també VIII. Mentre que DDR2 només és activat per col.lagens fibril·lars, en particular el tipus I i III.

No obstant, les funcions biològiques dels DDRs són encara desconegudes. La supressió de l'expressió de *DDR1* en ratolins (ratolins *DDR1 null*) és compatible amb el desenvolupament embrionari (Vogel *et al*, 2001). Els animals tenen una mida i pes menor fins la pubertat. En les femelles, aquesta disminució es manté en l'edat adulta

i presenten defectes en la implantació de blastòcits i en el desenvolupament de la glàndula mamària. També presenten problemes motors en les extremitats posteriors i una malformació de l'oida externa (Vogel, comunicació personal). L'eliminació del gen *DDR2* en ratolins també dona lloc a animals viables, però amb clars defectes en el creixement postnatal dels ossos degut a una disminució en la proliferació dels condrocits (Labrador *et al*, 2001).

DDR1 és particularment abundant en cèl.lules epitelials. En humans, la seva expressió s'ha detectat en cervell, ronyó, pulmó, tracte gastrointestinal, glàndula mamària, illots de Langerhans, la tiroide i en fibroblasts de la còrnia i dermis (Alves *et al*, 1995; Mohan *et al*, 2001; Sakamoto *et al*, 2001). Mentre que l'expressió de *DDR2* es detecta en múscul esquelètic, cardíac i teixit connectiu (Vogel, 1999; Olaso *et al*, 2001).

Ambdós receptors presenten una expressió elevada en determinats tumors malignes. En concret, s'ha detectat una sobreexpressió de *DDR1* en càncer de mama, ovari, d'esòfag, pulmó, gliomes i tumors cerebrals infantils (Laval *et al*, 1994; Alves *et al*, 1995; Barker *et al*, 1995; Perez *et al*, 1996; Weiner *et al*, 1996; Nemoto *et al*, 1997; Weiner *et al*, 2000; Garber *et al*, 2001; Ji *et al*, 2002; Pomeroy *et al*, 2002).

Mitjançant estudis d'hibridació *in situ* de seccions de tumors d'ovari i pulmó en humans, s'ha observat una expressió diferencial dels *DDRs* (Alves *et al*, 1995). *DDR1* es localitza en les cèl.lules tumorals invasives, mentre que *DDR2* en les cèl.lules circumdants al tumor, fet que indica una expressió coordinada entre *DDR1* i *DDR2* en la progressió del tumor. L'activació d'aquests receptors per col.lagen estimula la síntesi de metal.loproteases de matriu extracel.lular (MMPs) (Vogel *et al*, 1997; Hou *et al*, 2001; Hou *et al*, 2002). Aquestes són endopeptidases que degraden els components de la matriu extracel.lular (MEC). Les MMPs estan implicades en la remodelació tissular, processos de cicatrització, angiogènesi i invasió tumoral. Per això, s'ha suggerit que els *DDRs* actuarien com a sensors al col.lagen en la superfície cel.lular i estarien implicats en processos de diferenciació, proliferació i migració cel.lular (Vogel *et al*, 1997).

1.2.3. ESTRUCTURA GÈNICA I PROTEICA DE *DDR1*

DDR1 es va clonar per primera vegada a partir d'una genoteca de cDNA de placenta humana utilitzant una sonda amb una seqüència consens del domini tirosina quinasa dels RPTK i va ser anomenat *DDR* (Johnson *et al*, 1993). Simultàniament, va ser identificat per altres grups i anomenat amb els noms *NEP* (Zerlin *et al*, 1993), *trkE* (Di Marco *et al*, 1993), *Ptk-3* (Sanchez *et al*, 1994), *Cak* (Perez *et al*, 1994), *RTK6* (Laval *et al*, 1994) i *MCK-10* (Alves *et al*, 1995). Gens homòlegs a *DDR1* s'han identificat en ratolí (Zerlin *et al*, 1993), rata (Sanchez *et al*, 1994), i *Caenorhabditis elegans* (Ruvkun i Hobert, 1998).

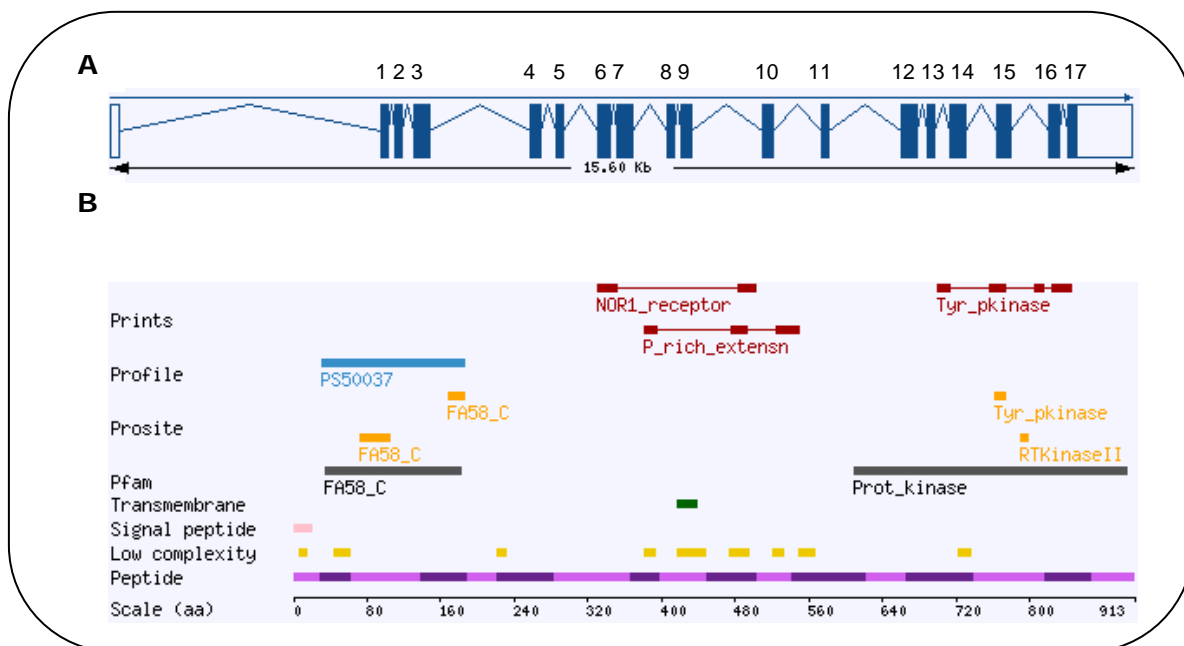


Figura 3. Estructura gènica i característiques proteiques del DDR1. A. Estructura exònica del gen *DDR1*. Les barres blaves corresponen als exons. Les barres blanques a les regions 5' i 3' UTR. B. Característiques proteiques d'acord amb la base de dades Vertebrate Genome Annotation (VEGA) (http://vega.sanger.ac.uk/Homo_sapiens/geneview?gene=OTTHUMG00000016356&db=core).

Les isoformes, ordenades de més llarga a més curta, són les següents:

- i. *isoforma c*. Codifica per una proteïna de 919 aminoàcids, essent la isoforma més llarga i l'única que conté tota la regió codificant.
- ii. *isoforma b*. Codifica per una proteïna de 913 aminoàcids. L'extrem 5' de l'exó 14 conté un lloc acceptor de *splicing* alternatiu críptic, el qual dona lloc a l'eliminació de les primeres 18 pb de l'exó 14 (codifiquen per 6 aminoàcids SFSLFS). Es desconeix quina és la conseqüència funcional de l'eliminació d'aquests 6 aminoàcids. Atès que l'expressió d'aquesta isoforma és la més abundant durant l'embriogènesi i la vida adulta, en aquest treball s'ha considerat la seva seqüència com la de referència per la posició de nucleòtids i aminoàcids.
- iii. *isoforma a*. Codifica per una proteïna de 875 aminoàcids. Eliminació per *splicing* alternatiu de les 18 pb entre els exons 13 i 14 i de l'exó 11 (37 aminoàcids). Com es descriu més endavant, l'exó 11 conté un motiu d'unió a dominis PTB i SH2.
- iv. *isoforma e*. Codifica per una proteïna de 767 aminoàcids. La regió exó-intró 5' de l'exó 10 té 2 llocs acceptors de *splicing* alternatiu. El més distal a 5' dona lloc a l'eliminació d'una part de l'exó 10, 11 i 12. En aquest cas la seqüència jxtamembrana està alterada en 28 aminoàcids i hi ha una pèrdua del lloc d'unió a ATP (seqüència d'aminoàcids GEGAFG). Sense aquest lloc d'unió a ATP, DDR1 molt probablement perd l'activitat intrínseca tirosina quinasa.

v. *isoforma d*. Codifica per una proteïna de 508 aminoàcids. Els exons 11 i 12 són eliminats per *splicing* alternatiu. La unió de l'exó 10 amb el 13 provoca un canvi en el marc de lectura i una terminació prematura de la traducció. Dóna lloc a una proteïna ancorada a la membrana, però sense el domini tirosina quinasa i la cua C-terminal.

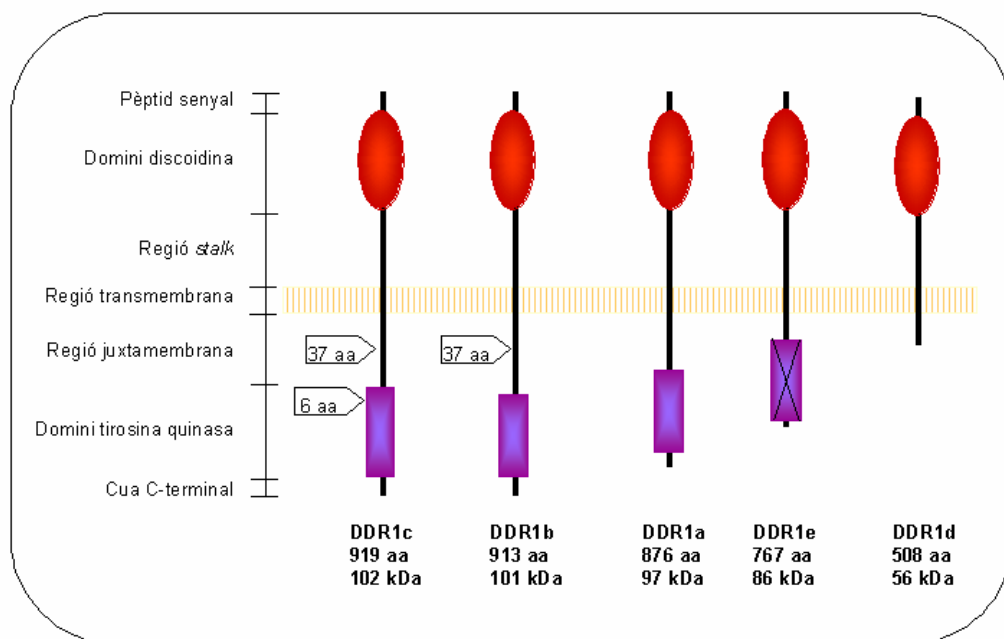


Figura 4. Esquema de les isoformes de DDR1. Isoformes ordenades de major a menor pes molecular. La isoforma b és la que es pren com referència.

L'expressió diferencial d'aquestes isoformes, tant a nivell gènic com proteic, i el seu significat biològic no ha estat estudiat en detall. S'ha observat, per exemple, que l'expressió de DDR1b augmenta amb el desenvolupament postnatal en comparació amb la isoforma DDR1a en rata (Sakuma *et al*, 1996). La isoforma DDR1a se sobre-exprés en determinats tumors mamaris (Perez *et al*, 1994) i macròfags infiltrant-se en tumors (Kamohara *et al*, 2001). Les isoformes DDR1d i DDR1e només s'han observat en línies cel·lulars de càncer de còlon (Alves *et al*, 2001).

Recentment s'ha observat que l'expressió i activació de DDR1 són induïdes per p53 en resposta a estrès genotòxic o dany del DNA (Ongusaha *et al*, 2003). Els mateixos autors han demostrat que la seqüència consens o element de resposta a p53 es localitza a l'intró 3 de *DDR1*, i no en la regió promotora tal i com havien proposat d'altres autors (Sakuma *et al*, 1996). No obstant, els mecanismes de regulació de l'expressió diferencial de les diferents isoformes i l'estabilitat dels mRNA estan encara per caracteritzar.

A continuació es fa una breu descripció dels dominis funcionals i estructurals i seqüències de rellevància per la funcionalitat del DDR1.

Regió extracel.lular: domini discoidina i regió *stalk*

El domini discoidina està constituït pels 160 aminoàcids N-terminals. La regió que va des del final del domini discoidina fins la regió transmembrana és anomenada regió *stalk* (Curat *et al*, 2001). Aquesta té 215 aminoàcids i no presenta homologia amb d'altres proteïnes. Ambdós, domini discoidina i regió *stalk*, són essencials per la unió a col.lagen, la dimerització i activació del receptor (Curat *et al*, 2001).

Els residus d'unió a col.lagen es troben en el domini discoidina i corresponen a aminoàcids conservats entre els dominis discoidina de les proteïnes incloses a la Taula VI (Abdulhussein *et al*, 2004). Curiosament, mutacions en aquests residus en la proteïna RS1 provoquen retinosis juvenil lligada al cromosoma X. Un 80% de les mutacions identificades en aquesta malaltia es localitzen en el domini discoidina del gen *RS1*, suggerint que aquest domini és crucial per la funcionalitat de la proteïna (The Retinoschisis Consortium, 1998).

La regió *stalk* conté dues seqüències rellevants per la funcionalitat del receptor:

- i. Una seqüència de glicosilació. Aquesta és necessària per la funcionalitat del receptor (Curat *et al*, 2001).
- ii. Una seqüència diana per una proteasa de superfície cel.lular. El DDR1, en resposta a la interacció amb el col.lagen, és processat per una proteasa de superfície de membrana en un fragment soluble, el qual pot difondre per la MEC, i un fragment ancorat a la membrana i que roman fosforilat durant dies (Vogel, 2002). El mecanisme i la implicació biològica del processament del DDR1 són encara desconeguts. Se sospita que, a l'igual que en d'altres receptors d'adhesió (VCAM-1, L-selectin, syndecan), receptors acoblats a proteïnes G (receptor de vasopressina V2) o precursors de lligands ancorats a membrana (TNF- α , Fas ligand), aquest processament proteolític és un mecanisme de regulació de l'activitat de DDR1.

Regió juxtamembrana

La regió juxtamembrana està codificada pels exons 10, 11 i 12 i té 176 aminoàcids. Aquesta regió és la més afectada pels mecanismes de *splicing* alternatiu i conté diverses seqüències consens de fosforilació de tirosines que indica que és una regió important per la transducció del senyal.

L'exó 11, present en les isoformes DDR1b i DDR1c, conté el motiu LLXNPXY que és fosforilat en el residu tirosina per l'activació del domini tirosina quinasa en resposta a la unió a col.lagen. Aquest motiu correspon al motiu consens ψ XNPXpY (essent ψ un aminoàcid hidrofòbic) d'unió al domini PTB de la proteïna adaptora Shc. Recentment, s'ha demostrat que l'activació de la isoforma DDR1b té un paper important en la

diferenciació de monòcits a macròfags i aquesta depèn de la unió de Shc a DDR1, activant la via de la p38 MAPK (Matsuyama *et al*, 2003).

Hi ha d'altres tirosines conservades en l'exó 12 (tirosina 543 (YSGD) i tirosina 547 (YMEP)). Aquestes són potencials llocs d'unió a proteïnes amb un domini SH2 o PTB i, per tant, serien importants en la via de senyalització de DDR1. Fins el moment no s'ha demostrat la seva fosforilació o interacció amb d'altres proteïnes adaptadores. No obstant, s'ha observat que la proteïna FRS2, amb un domini PTB, interacciona amb la regió juxtamembrana de la isoforma DDR1a (Foehr *et al*, 2000).

L'homologia de la regió juxtamembrana entre DDR1 i DDR2 és baixa, però hi ha una seqüència de 25 aminoàcids (aminoàcids 566-590 del DDR1) localitzats en l'exó 12, que es manté pràcticament igual en ambdós receptors, suggerint que pot ser una regió crítica per la transducció del senyal d'ambdós DDRs (Alves *et al*, 2001).

Domini tirosina quinasa i cua C-terminal

El domini tirosina quinasa presenta l'homologia més elevada (del 55-58%) amb els RPTK de neurotrofines (NTRK1, 2 i 3 i NGFR). No obstant, l'autofosforilació del receptor per unió a col.lagen té una cinètica completament diferent comparada amb la dels RPTKs. La fosforilació s'observa al cap d'unes hores en cèl.lules adherents i d'uns 30 minuts en cèl.lules en suspensió (Vogel *et al*, 1997; L'Hote *et al*, 2002) i es manté durant dies (Vogel, 2002). Si el DDR1 actua com a sensor al col.lagen en la superfície cel.lular, aquesta cinètica donaria el temps suficient per permetre l'establiment de les connexions cèl.lula-MEC necessàries pel posicionament cel.lular.

La funció de les isoformes sense activitat catalítica (DDR1d i DDR1e) són desconegudes. S'ha postulat que tindrien una funció reguladora del mateix receptor DDR1 per analogia a les isoformes truncades (sense domini tirosina quinasa) dels receptors NTRK2 i NTRK3 (Baxter *et al*, 1997; Palko *et al*, 1999; Alves *et al*, 2001).

En l'extrem C-terminal de les isoformes DDR1a, DDR1b, DDR1c hi ha un motiu LNTV. Aquest motiu pot actuar com a lloc d'unió a les proteïnes amb un o varis dominis PDZ. Els dominis PDZ són punts d'interacció entre proteïnes i faciliten l'ancoratge i localització de complexos proteics (Garner *et al*, 2000). Per exemple, s'ha demostrat que els dominis PDZ són essencials per l'agrupació i localització de receptors eph i HER2 (Borg *et al*, 2000). Curiosament, el canvi en el marc de lectura de la isoforma DDR1d dona lloc a l'aparició del motiu GAPV que també és lloc d'unió a PDZ (Alves *et al*, 2001).

La presència de mutacions en qualsevol dels dominis mencionats, així com en les seqüències de *splicing* alternatiu, podrien afectar la funcionalitat del receptor. De moment, però, no s'han descrit mutacions en el gen *DDR1* associades a cap malaltia o tret fenotípic.

1.2.4. IMPLICACIÓ DE DDR1 EN EL SISTEMA NERVIÓS CENTRAL

En sistema nerviós central (SNC), l'expressió de DDR1 ha estat estudiada en els models múrids de ratolí (Zerlin *et al*, 1993; Bhatt *et al*, 2000) i rata (Sanchez *et al*, 1994) mitjançant les tècniques d'hibridació *in situ* i RT-PCR. En aquests treballs s'observa una expressió de DDR1 en zones proliferatives durant el desenvolupament embrionari i en zones de neurogènesi secundària en estadis adults. Sanchez i col.laboradors també detecten expressió de DDR1 en substància blanca.

El nostre grup ha estudiat l'expressió de DDR1 durant el neurodesenvolupament i en l'adult mitjançant la tècnica de la hibridació *in situ* en talls de cervell de ratolí (estudi realitzat per la doctoranda Neus Franco). El patró d'expressió que observem és comparable amb el dels estudis anteriors. No obstant, en l'animal postnatal i en l'adult detectem una expressió en una subpoblació cel.lular compatible amb oligodendròcits. Actualment estem treballant per esbrinar de quina subpoblació es tracta.

A finals del 2000, es va publicar el primer i únic estudi que posa de manifest la importància de DDR1 en el desenvolupament del SNC (Bhatt *et al*, 2000). Aquest estudi demostra que la interacció de DDR1 amb el col.lagen i la corresponent via de senyalització activada és essencial per la formació axonal de les cèl.lules granulars durant el desenvolupament del cerebel. El fet que DDR1 s'expressi també durant l'edat adulta, va fer suggerir a Bhatt i col.laboradors que DDR1 estaria implicat en processos de manteniment d'estructures neuronals, regulació de la formació de sinapsis i en els processos de *sprouting*. Precisament, en el SNC de l'adult, aquestes tasques són realitzades per l'astroglia i oligodendroglia (Fields i Stevens-Graham, 2002).

D'altra banda, hi ha varis treballs puntuals i independents que demostren la modulació de la proliferació i diferenciació de cèl.lules neuroepitelials i astroglia per col.lagen (Goetschy *et al*, 1987; Ali *et al*, 1998). Però cap d'aquests estudis explora la possible participació del DDR1.

1.2.5. IMPLICACIÓ DE DDR1 EN PROCESSOS D'ADHESIÓ, MIGRACIÓ, DIFERENCIACIÓ I PROLIFERACIÓ CEL.LULAR

El conjunt de macromolècules que constitueixen la MEC no només té un paper important en el manteniment de la integritat dels teixits i els òrgans, sinó també en l'adhesió, la proliferació, la migració i la diferenciació de les cèl.lules que envolten. És més, les interaccions cèl.lula-MEC tenen implicacions molt importants durant el desenvolupament i morfogènesi.

El component majoritari de la MEC és el col.lagen. S'han descrit 3 tipus de receptors de col.lagen: les integrines, els DDRs i la glicoproteïna VI (Vogel, 2001). Els tres

activen diverses vies de transmissió del senyal implicades en processos d'adhesió, proliferació, migració i diferenciació cel·lulars. La glicoproteïna VI s'expressa exclusivament en plaquetes, mentre que integrines i DDRs es troben expressats en diferents tipus cel·lulars i tenen la capacitat de monitoratge i modulació de la integritat de la MEC mitjançant l'activació de mecanismes de degradació o síntesi de col·lagen. En concret, l'expressió de DDR1 i de col·lagen de tipus I, III i VIII té lloc en paral·lel en cèl·lules musculars de rata (Hou *et al*, 2001). Per altra banda, en cèl·lules de ratolí DDR1 *null* s'observa una inhibició de l'expressió de metal·loproteases (MMP-2 i MMP-9) (Hou *et al*, 2001). Ambdues observacions són clars indicadors del paper de DDR1 en la migració i adhesió cel·lular.

Fins el present moment, s'ha demostrat que la unió de col·lagen a DDR1 promou la diferenciació de cèl·lules musculars esquelètiques (Vogel *et al*, 2000), l'adhesió, proliferació i migració de cèl·lules musculars llises en el procés de reparació de la paret arterial (Hou *et al*, 2001), l'adhesió cel·lular i inhibició de la proliferació cel·lular de cèl·lules mesangials i epitel·lials (Curat i Vogel, 2002), la diferenciació de monòcits a macròfags (Matsuyama *et al*, 2003b) i alliberació de citokines per macròfags (Matsuyama *et al*, 2004). I, com ja s'ha mencionat, també promou l'elongació dels axons en cèl·lules granulars de cerebel (Bhatt *et al*, 2000).

No obstant, en diverses línies cel·lulars, no es detecta una activació de DDR1 per col·lagen, suggerint que aquesta activació requereix de la presència de cofactors o lligands addicionals (Foehr *et al*, 2000; L'hote *et al*, 2002). Recentment, s'ha demostrat que la proteïna Wnt-5a és co-factor necessari per l'activació del DDR1 per col·lagen en l'adhesió i migració de cèl·lules epitel·lials de mama HB2 (Jonsson i Andersson, 2002; Dejmek *et al*, 2003), implicant el DDR1 en la via de senyalització de les proteïnes Wnt. Aquestes són glicoproteïnes de secreció autocrina o paracrina que modulen gran varietat de processos cel·lulars (diferenciació, polarització, migració i proliferació) tant durant el desenvolupament com en l'adult (Miller, 2001).

1.3. MÈTODES DE GENOTIPATGE DE POLIMORFISMES D'UN ÚNIC NUCLEÒTID (SNPs)

1.3.1. POLIMORFISMES D'UN ÚNIC NUCLEÒTID

Els SNPs són substitucions d'un sol nucleòtid en la seqüència de DNA que resulten en la generació de variants de seqüència (o al·lels), les quals han de coexistir en la població i la freqüència de l'al·lel menys representat ha de ser superior a l'1%. En principi, els SNPs podrien ser bi, tri o tetra-al·lèlics. No obstant, en humans, tan sols s'han detectat SNPs bi-al·lèlics en DNA genòmic. Els SNP tri i tetra-al·lèlics tan sols s'han identificat en molt baixa freqüència en el genoma mitocondrial (Kogelnik *et al*, 1998). Les freqüències de les diferents substitucions nucleotídiques possibles no són equivalents. Així, unes 2/3 parts dels SNPs impliquen la transició C/T (o G/A, en la cadena complementària), mentre que les transversions s'observen en igual proporció. L'elevat percentatge de SNPs C/T (o G/A) es deu probablement a la reacció de deaminació de la 5-metilcitosina resultant en timina (Holliday i Grigg, 1993).

El 90% dels polimorfismes observats en el genoma humà són SNPs. S'ha estimat que la seva distribució és d'un SNP per cada 1000 pb de DNA (Sachidanandam *et al*, 2001). El genoma haploide humà consta d'unes 3.3×10^9 pb, de manera que, un individu qualsevol tindrà més de 3 milions de SNPs en el seu genoma. Fins el present moment, hi ha registrats més de 4 milions de SNPs candidats en la base de dades d'accés públic dbSNP mantinguda pel National Center for Biotechnology Information (NCBI, <http://www.ncbi.nlm.nih.gov/SNP/>). Molts d'aquests s'han identificat *in silico* o en mostres aïllades i estan essent verificats i les seves freqüències estudiades en diverses poblacions.

A finals dels 90, els SNPs es van proposar com a marcadors genètics per la identificació de *loci* associats a malalties complexes i a la farmacogenètica. L'estudi de la variabilitat en la resposta a fàrmacs determinada genèticament, o farmacogenètica, és anàloga a l'estudi de malalties complexes en quant a la implicació de factors genètics i ambientals. Fins aquell moment, els polimorfismes més estudiats havien estat els microsatèl·lits o STRPs (*Short Tandem Repeat Polymorphisms*). Aquests són repeticions en tàndem d'una unitat bàsica de seqüència de DNA, d'entre dos i sis nucleòtids. Són polimorfismes abundants i presenten un elevat nombre d'al·lels. Hom estima que hi ha un STRP per cada 30-60 kb de DNA. El fet de què els STRPs fossin més informatius que els SNPs (*loci* multi-al·lèlics versus bi-al·lèlics, respectivament), tinguessin una distribució homogènia en el genoma i es poguessin analitzar fent ús de la tècnica de la PCR, va promoure el desenvolupament de sistemes semi-automàtics d'anàlisi de STRP (Livak *et al*, 1995; Hall *et al*, 1996). Els STRPs van esdevenir els marcadors genètics ideals en els estudis de lligament per la identificació de gens responsables de malalties monogèniques.

Ara bé, el viratge a finals dels 90 per l'estudi de malalties complexes i farmacogenètica, va comportar un canvi en els dissenys experimentals. En ambdós casos, com ja s'ha comentat anteriorment, els estudis d'associació poden ser més efectius per detectar gens d'efecte moderat o lleu que els estudis de lligament (Risch i Merikangas, 1996). En aquests estudis, els SNPs són els marcadors genètics més adequats a estudiar pels següents motius:

1. Menor taxa de mutació que els STRPs ($\sim 10^{-8}$ i $\sim 10^{-5}$ - 10^{-2} , per SNPs i STRPs, respectivament (Weber i Wong, 1993; Chakraborty *et al*, 1997; Li *et al*, 1996)). L'elevada variabilitat o múltiples al·lels observats en els STRPs reflecteix, en part, una elevada taxa de mutació, la qual pot ser un factor confusionari en els estudis d'associació.
2. Elevada densitat en el genoma humà (1 cada 1000 pb) que permet la caracterització d'un mapa d'haplotips per la identificació de gens de susceptibilitat per malalties complexes (Sachidanandam *et al*, 2001).
3. Disponibilitat d'un gran ventall de tècniques d'identificació i rastreig d'SNPs.
4. Implicació directa en la malaltia o tret fenotípic d'aquells SNPs amb conseqüències funcionals (afectant les regions codificants o reguladores del gen).

L'interès en l'anàlisi d'SNPs rau en l'esperança que aquests seran una eina poderosa en el procés d'identificació de gens de susceptibilitat de malalties complexes, en la farmacogenètica i en d'altres àrees de la biomedicina. No obstant, hi ha una gran disparitat entre el ritme amb què es descobreixen nous SNPs i el desenvolupament de noves tècniques que permetin l'anàlisi d'un gran nombre d'aquests. Malgrat l'existència de múltiples mètodes de genotipatge, hi ha una demanda de nous mètodes capaços d'analitzar gran nombre d'SNPs de forma fiable, ràpida i amb un baix cost. En aquest context, la doctoranda ha estat treballant en el desenvolupament nous protocols de genotipatge fent ús de la tecnologia dels *eTag*TM *reporters* a ACLARA BioSciences Inc.

Els múltiples mètodes de genotipatge d'SNPs actuals han evolucionat a partir de diferents principis bàsics molts dels quals es coneixien des de fa més de 20 anys. Com per exemple, la hibridació al·lel específica d'oligonucleòtids o la restricció enzimàtica al·lel específica (Smith i Wilcox, 1970; Kafatos *et al*, 1979). No obstant, no ha estat fins el descobriment de la PCR que el genotipatge s'ha pogut realitzar a mitjana i gran escala. La majoria dels mètodes actuals requereixen de l'amplificació prèvia per PCR de les seqüències que contenen els SNPs d'interès. Hi ha mètodes alternatius (que no depenen de PCR), però requereixen de quantitats més grans de DNA, el qual és un factor limitant quan es treballa amb mostres humanes.

A continuació es revisen els principis bàsics per la discriminació al·lèlica i es mencionen les tècniques de genotipatge d'SNPs d'ús actual.

1.3.2. TÈCNIQUES DE GENOTIPATGE D'SNPS: REQUERIMENTS

S'han desenvolupat un gran nombre de mètodes de genotipatge. La tria d'un o altre mètode depèn dels objectius experimentals i recursos disponibles de l'investigador o grup de recerca. Per exemple, un projecte pot requerir el genotipatge d'un nombre limitat d'SNPs en una mostra de mida gran o bé de molts SNPs en un o pocs individus. La flexibilitat en el disseny, l'habilitat de quantificar freqüències al·lèliques en *pools* de mostres de DNA i d'analitzar més d'un SNP per reacció, així com la semi-automatització del procés, són alguns dels aspectes a considerar en l'avaluació d'una tècnica de genotipatge. En resum, les propietats del mètode de genotipatge ideal capaç d'acomplir amb totes les necessitats actuals serien les següents:

- Específic
- Robust, fins i tot en presència de poca quantitat de DNA
- Disseny senzill i flexible
- Assaig homogeni, en un sol tub o pouet
- Mínima optimització de la reacció
- Quantitatiu, permet l'anàlisi en *pools* de mostres
- Multiplicador, permet l'anàlisi de més d'un SNP per reacció
- Assaig de fàcil automatització
- Automatització de l'anàlisi de resultats
- Assaig escalable i adaptable (processament des de baix a elevat nombre de mostres)
- Cost baix per genotip (reactius, instruments, manteniment i mà d'obra inclosos)

1.3.3. ESTRATÈGIES DE DISCRIMINACIÓ AL·LÈLICA

Les estratègies bàsiques emprades en la discriminació al·lèlica d'SNPs es poden agrupar en quatre categories. Aquestes són:

1. Hibridació al·lèl específica
2. Discriminació al·lèl específica per una endonucleasa
3. Extensió al·lèl específica d'un oligonucleòtid o *primer*
4. Ligació al·lèl específica d'oligonucleòtids

Cadascuna d'aquestes estratègies pot acoblar-se a diferents mètodes de detecció els quals detecten canvis en la massa, la conductivitat elèctrica o l'emissió de llum de molècules presents en els productes al·lèl específics. La detecció de fluorescència és potser el mètode de detecció més utilitzat. El marcatge fluorescent pot acoblar-se fàcilment a la majoria de tècniques i disposa de múltiples plataformes de mesura (lectors de plaques, electroforesi de gel i capil·lar i *microarrays* de DNA). L'espectrofotometria de masses i la quimioluminiscència (on l'emissió de llum és resultat d'una reacció bioquímica, com és el cas particular de la tècnica Pyrosequencing[®]) són dos sistemes de detecció que estan guanyant acceptació en els últims tres anys. Tots tres mètodes de detecció són compatibles amb el desenvolupament de protocols que permeten el rastreig semi- i automatitzat d'SNPs a mitjana i gran escala.

El format d'aquestes estratègies és també variable i molts dels mètodes poden realitzar-se en solució o bé en fase sòlida (immobilitzant part dels components de la reacció en membranes de niló, *microarrays*, microesferes o plaques de 96 o 384 pouets).

1. Hibridació al.llel específica

Es basa en les diferències en l'estabilitat entre dos cadenes de DNA perfectament complementàries o amb un nucleòtid desparellat. Sota unes condicions determinades d'hibridació (temperatura, tampó de reacció), una sonda o oligonucleòtid de seqüència curta (entre 10 i 30 nucleòtids) formarà un dúplex amb el DNA diana només si són perfectament complementaris (Figura 5.A).

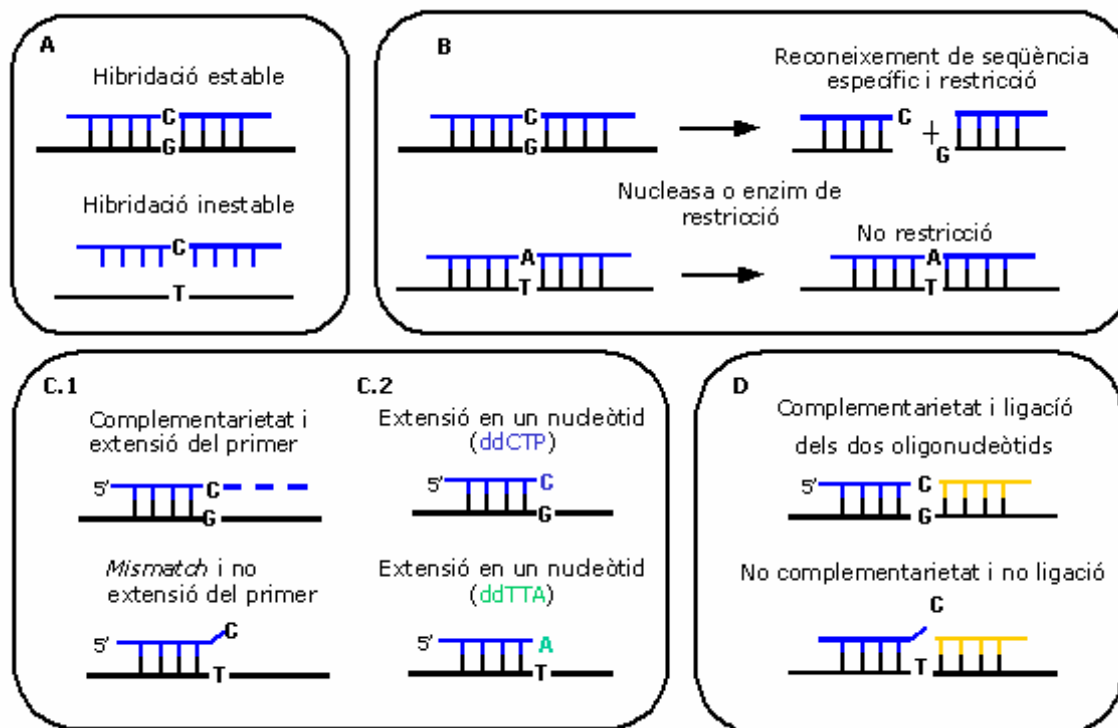


Figura 5. Estratègies bàsiques emprades en la discriminació al·lèlica d'SNPs. A. Hibridació al.llel específica; B. Discriminació al.llel específica per una endonucleasa; C.1. Extensió al.llel específica d'un oligonucleòtid o *primer*; C.2. Incorporació específica d'un dideoxinucleòtid; D. Ligació d'oligonucleòtids al.llel específica.

Els primers mètodes d'hibridació al.llel específica van sorgir a partir de les tècniques del *dot blot* (Conner *et al*, 1983) i *reverse dot blot* (Saiki *et al*, 1989). Aquesta última pot considerar-se com la precursora dels *microarrays* pel genotipatge simultani de centenars d'SNPs. Les tècniques basades en la hibridació al.llel específica es resumeixen en la Taula VI. Es diferencien per com té lloc la discriminació al·lèlica, el sistema de detecció i el format de la reacció. Són tècniques que poden automatitzar-se fàcilment i permeten l'anàlisi de múltiples mostres simultàniament. Les dues tècniques més populars són TaqMan® (Applied Biosystems, FosterCity, CA, EUA)

(Livak *et al*, 1995; 1999) i Molecular Beacons (Tyagi i Kramer, 1996; Tyagi *et al*, 1998). En aquestes, la discriminació al·lèlica té lloc durant l'amplificació del DNA per PCR. Es tracta, doncs, d'assaigs homogenis i els resultats poden obtenir-se en menys de 2 hores. No obstant, no permeten l'anàlisi simultània de múltiples SNPs ja que les condicions òptimes d'hibridació són diferents per cada SNP. I, d'altra banda, no hi ha fluorocroms o molècules de marcatge suficients com per analitzar més d'1 o 2 SNPs per reacció. Malgrat la utilització dels *microarrays* (com per exemple, el GeneChip d'Affymetrix, Santa Clara, CA, EUA) pot semblar una solució a aquests dos problemes, varis estudis han demostrat que aquests no poden diferenciar els genotips homozigots dels heterozigots per un percentatge d'SNPs considerable (2–12%) (Wang *et al*, 1998; Cargill *et al*, 1999; Halushka *et al*, 1999).

Taula VI. Tècniques de discriminació al·lèlica basades en la hibridació al·lèl específica.

Nom (empresa) ^a	Mètode de detecció	Format	Avantatge	Desavantatge	Referència
Reverse dot blot	F ^b , Q ^c , C ^d	FS ^e	Multiplicatiu	Soroll de fons per hibridació inespecífica	Saiki <i>et al</i> , 1989
TaqMan (Applied Biosystems)	FR ^f	S ^g	Homogeni, simple, ràpid, automatitzable	Elevat cost	Livak <i>et al</i> , 1995
Molecular Beacons	FR	S	Homogeni, simple i ràpid	Elevat cost	Tyagi i Kramer, 1996; Tyagi <i>et al</i> , 1998
Microarray GeneChip (Affymetrix)	F	FS	Multiplicatiu	Elevat cost i percentatge elevat de resultats erronis per hibridació inespecífica	Wang <i>et al</i> , 1998; Cargill <i>et al</i> , 1999
DASH ^h (ThermoHybaid)	F	S, FS	Baix cost	Disseny complex	Prince <i>et al</i> , 2001

^aIndica l'empresa que comercialitza la tècnica; ^bF, fluorescència; ^cQ, quimioluminiscència; ^dC, colorimetria; ^eFS, fase sòlida; ^fFR, transferència d'energia per ressonància de fluorescència; ^gS, fase en solució; ^hDASH, *dynamic allele-specific hybridization*.

Per tal de millorar la discriminació al·lèlica, s'han utilitzat anàlegs de nucleòtids com els àcids nucleics peptídics (PNA, *peptide nucleic acid*) o bé els àcids nucleics tancats (LNA, *locked nucleic acid*) en la síntesi de les sondes (Ross *et al*, 1997; Orum *et al*, 1999). De la mateixa manera, les sondes TaqMan també han estat modificades afegint en l'extrem 3' una molècula d'unió al solc petit del DNA (MGB, *minor groove-binder molecule*) (Kuimelis *et al*, 1997). Aquestes modificacions incrementen la temperatura de fusió (T_m) de les sondes i, per tant, milloren la capacitat discriminatòria de les mateixes, però encareixen considerablement el cost de la reacció.

2. Discriminació al.lél específica per una endonucleasa

Es basen en la capacitat que tenen algunes endonucleases en reconèixer i tallar seqüències específiques de DNA (Figura 5.B, Taula VII). La tècnica més emprada en els laboratoris de genotipatge ha estat la que combina l'amplificació per PCR de la regió de DNA que conté l'SNP amb la posterior digestió amb un enzim de restricció (PCR-RFLP). No obstant, aquesta està essent substituïda per d'altres ja que no sempre es disposa d'una endonucleasa o enzim de restricció per la seqüència d'interès, és un procés llarg (la restricció enzimàtica requereix d'incubacions de 2 a 10 h) i de difícil automatització.

La tècnica Invader o de digestió invasiva per una 5'endonucleasa es descriu en detall més endavant (vegeu apartat *eTag™ Multiplex Invader® SNP Assay*: un nou protocol per l'anàlisi simultània d'SNPs).

Taula VII. Tècniques de discriminació al.lèlica basades en l'acció d'una nucleasa específica de seqüència.

Nom (empresa) ^a	Mètode de detecció	Format	Avantatge	Desavantatge	Referència
PCR-RFLP ^b	F ^c , Q ^d , C ^e	S ^f	Barat; no requereix d'equips sofisticats	Laboriós	Hixson i Vernier, 1990
Invader (ThirdWave Technologies Inc)	FR ^g	S	Homogeni; no requereix d'amplificació del DNA d'interès	Requereix gran quantitat de DNA genòmic	Lyamichev <i>et al</i> , 1999
McSNP ^h	F	S	Homogeni; barat	Laboriós	Akey <i>et al</i> , 2001

^aIndica l'empresa que comercialitza la tècnica; ^bPCR-RFLP, *polymerase chain reaction-restriction length polymorphism*; ^cF, fluorescència; ^dQ, quimioluminiscència; ^eC, colorimetria; ^fS, fase en solució; ^gFR, transferència d'energia per ressonància de fluorescència; ^hMcSNP, *Melting Curve Single Nucleotide Polymorphism*.

3. Extensió al.lél específica d'un oligonucleòtid o primer

Es basa en la precisió de la incorporació de nucleòtids complementaris a una seqüència diana de DNA per una DNA polimerasa (Sylvänen *et al*, 1990). Hi ha dos tipus de mètodes els quals es descriuen a continuació (Figura 5.C, Taula VIII).

3.1. Amplificació al.lél específica per PCR

Aquesta també es coneix amb el nom de ARMS (*Amplification Refractory Mutation System*) (Newton *et al*, 1989), ASA (*Allele Specific Amplification*) (Okayama *et al*, 1989), ASPCR (*Allele-Specific PCR*) (Wu *et al*, 1989) i PASA (*PCR Amplification of Specific Alleles*) (Sommer *et al*, 1989). Es basa en què l'extensió de l'encebador o oligonucleòtid durant la PCR depèn de la complementarietat del seu extrem 3' a la seqüència diana (Figura 7.C.1). Malgrat la simplicitat i baix cost d'aquest mètode, ja que no cal fer cap tractament previ o posterior a la PCR, no ha estat gaire utilitzat ja

que les DNA polimerases sovint donen lloc a falsos positius. El disseny dels oligonucleòtids al·lel específics no és simple i cal, sovint, optimitzar el protocol de PCR per cadascun dels SNPs a analitzar. Tot i que no és objectiu del present treball, mencionar que la doctoranda ha desenvolupat un protocol pel genotipatge simultani d'SNPs combinant els mètodes TaqMan i ARMS amb el marcatge dels *eTag reporters* (Virgos i Cronin, PCT/US02/32867).

3.2. Incorporació al·lel específica de nucleòtids

En aquest disseny, l'oligonucleòtid híbrida amb el seu extrem 3' en un nucleòtid *upstream* respecte la posició de l'SNP i es detecta el tipus de nucleòtid incorporat (Figura 5.C.2). Per tal d'evitar l'elongació de l'oligonucleòtid en més d'una base, s'utilitzen dideoxinucleòtids en la barreja de PCR (Syvänen *et al*, 1990). Cadascun dels quatre dideoxinucleòtids està marcat amb un fluorocrom diferent, de manera que el genotip es dedueix per anàlisi de la llum emesa pel dideoxinucleòtid incorporat (Pastinen *et al*, 1997).

Taula VIII. Tècniques de discriminació al·lèlica basades en l'extensió d'un oligonucleòtid per una DNA polimerasa. Incloses només aquelles tècniques amb més capacitat multiplicadora.

Nom (empresa) ^a	Mètode de detecció	Format	Avantatge	Desavantatge	Referència
Amplificació al·lel específica					
ARMS ^b	F ^c , FR ^d , C ^e	S ^f	Baix cost, homogeni	Disseny complexe	Newton <i>et al</i> , 1989
Incorporació al·lel específica de nucleòtids					
SnaPshot (Applied Biosystems)	F	S ^g	Multiplicador	Purificació i quantificació del producte de PCR	www.appliedbiosystems.com
Pyrosequencing (Pyrosequencing)	LI ^h	S	Multiplicador, capaç de seqüenciar fins 50 bases	Elevat cost i disseny complexe	Nyren <i>et al</i> , 1993
SNP-it, SNPstream (Orchid Biosciences, Luminex)	C	FS	Multiplicador, barat	Múltiples reaccions	Nikiforov <i>et al</i> , 1994
MassExtend, MassArray (Sequenom)	MS ⁱ	FS	Multiplicador	Elevat cost (instrumentació)	Braun <i>et al</i> , 1997
Microarray universal GenFlex (Affymetrix)	F	S	Multiplicador, disseny flexible	Elevat cost (reacció i instrumentació)	Fan <i>et al</i> , 2000

^aIndica l'empresa que comercialitza la tècnica; ^bARMS, *Amplification refractory mutation system*; ^cF, Fluorescència; ^dFR, Transferència d'energia per ressonància de fluorescència; ^eC, Colorimetria; ^fS, Fase en solució; ^gFS, Fase sòlida; ^hLI, llum; ⁱMS, Espectrofotometria de masses.

Actualment, el mètode d'extensió al·lel específica d'un oligonucleòtid és el mètode més emprat pels grups de genotipatge d'SNPs ja que el disseny és simple i ofereix una gran diversitat de sistemes de detecció (Syvänen, 1999). Es tracta d'un mètode altament específic, quantitatiu, multiplicador i escalable, però que requereix de la

purificació de la reacció de PCR i, molt sovint també, de la reacció d'incorporació al.llel específica de nucleòtids prèviament a la detecció i lectura dels genotips.

4. Ligació d'oligonucleòtids al.llel específica

Es basa en la unió covalent de dos oligonucleòtids adjacents per una DNA ligasa, si aquests són perfectament complementaris al DNA diana en la regió de ligació (Wu i Wallace, 1989) (Figura 5.D). El disseny consta de dos oligonucleòtids, un dels quals és al.llel específic en el seu extrem 3' o 5'. La detecció del producte resultat de la ligació és indicatiu de la presència de l'al.llel complementari a l'oligonucleòtid al.llel específic. Els mètodes basats en la ligació al.llel específica d'oligonucleòtids es poden classificar en dues categories:

4.1. Assaig de la ligació d'oligonucleòtids (Landegren *et al*, 1988) i que generalment inclou l'amplificació posterior dels oligonucleòtids ligats. Recentment, s'han desenvolupat dos mètodes basats en aquesta aproximació i que permeten l'anàlisi de múltiples SNPs per reacció. Aquests són l'assaig SNPlex™ (<http://www.appliedbiosystems.com>) i SNPWave™ (van Eijk *et al*, 2004; <http://www.keygene.com>).

4.2. Assaig de ligació i amplificació per gir-volta del cercle (Nilsson *et al*, 1994; Lizardi *et al*, 1998), que no depèn de la PCR. El problema que presenten aquests mètodes és el cost de la DNA ligasa i del marcatge dels múltiples oligonucleòtids.

1.3.4. eTAG™ MULTIPLEX INVADER® SNP ASSAY: UN NOU PROTOCOL PER L'ANÀLISI SIMULTÀNIA D'SNPs

Què són els eTag reporters?

Els *eTag reporters* són molècules fluorescents de baix pes molecular (150-5.000 da) amb una relació càrrega/massa única que permet la seva separació per electroforesi capil·lar (EC) (Figura 6). Els *eTag reporters* es poden acoblar a sondes, siguin aquestes oligonucleòtids, pèptids o proteïnes. Quan la sonda interacciona amb la corresponent diana, l'*eTag reporter* és alliberat a la solució i separat per EC (Singh, USPTO 6,514,700). Així, doncs, la identificació d'un *eTag reporter* concret és indicatiu de la presència d'un determinat component en la reacció. ACLARA BioSciences Inc. (<http://www.aclara.com>) ha desenvolupat una llibreria de més d'un centenar d'*eTag reporters*. Aquests es poden acoblar a múltiples sondes i permeten la identificació simultània de diferents components en una mateixa reacció ja sigui per l'anàlisi d'àcids nucleics (DNA i RNA) o proteïnes.

La separació i quantificació dels *eTag reporters* té lloc mitjançant la utilització d'aparells d'electroforesi capil·lar (EC) o seqüenciadors automàtics (com per exemple, ABI PRISM 3100 Genetic Analyzer (Applied Biosystems), MegaBACE1000(APBiotech)). L'anàlisi dels resultats pot fer-se de forma semi-

automàtica utilitzant el programa eTag Informer™ desenvolupat a ACLARA BioSciences Inc.

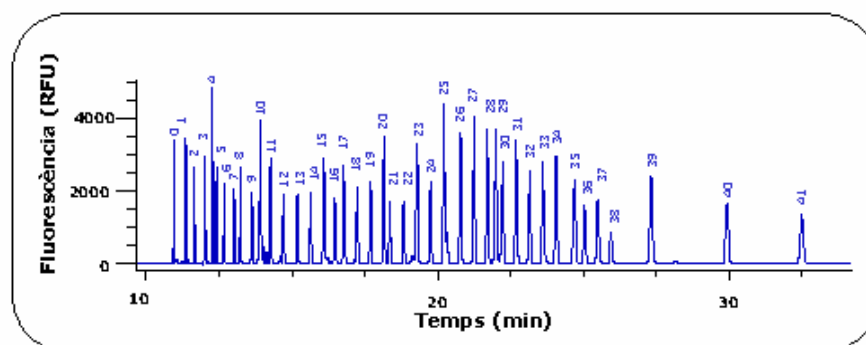


Figura 6. Electroforograma amb 42 *eTag* reporters. Separació electroforètica realitzada en un ABI3100 Genetic Analyzer (Applied Biosystems).

Combinació dels *eTag* reporter amb la tècnica Invader per l'anàlisi d'SNPs

En la tècnica Invader (ThirdWave Technologies Inc, Madison, WI, EUA), dos oligonucleòtids (les sondes Invader i al.l.el específica), hibriden en tàndem a una cadena de DNA complementària i formen una estructura triple com s'indica en la Figura 7. En aquesta estructura, l'extrem 3' de la sonda Invader se solapa en un nucleòtid amb la sonda al.l.el específica. Aquesta última té una cua de nucleòtids en l'extrem 5' (anomenada *flap*), la qual no és complementària a la mostra de DNA a analitzar. Les 5'endonucleases, conegudes amb el nom de Cleavases®, reconeixen aquesta estructura triple i tallen la sonda al.l.el específica en el nucleòtid superposat (Lyamichev *et al*, 1999). Aquest tall allibera la *flap* a la solució i es detecta en una segona reacció Invader la qual és comuna per tots els SNPs (Invader Square). La *flap* hibrida amb un *cassette* de transferència d'energia per ressonància de fluorescència (FRET, *fluorescence resonance energy transfer*) i forma una estructura triple que serà reconeguda i tallada per la 5'endonucleasa. La digestió del *cassette* FRET separa el fluorocrom del *quencher* i genera un senyal fluorescent que és detectat en un lector de fluorescència (Figura 7.A).

Les Cleavases són enzims de la família de les 5' *flap* endonucleases (FEN), les quals tenen un paper clau en la replicació, recombinació i reparació del DNA (Lieber, 1997). L'especificitat d'aquesta tècnica està determinada per l'habilitat d'aquests enzims de discriminar entre substrats on les dues sondes se superposen de les que no. La substitució d'un únic nucleòtid en el DNA diana en el punt de superposició entre sondes impedeix la formació de l'estructura triple, la qual és necessària per l'activitat de l'enzim. Aquesta forta dependència entre l'activitat de l'enzim i la formació de l'estructura triple és la que permet que aquesta tècnica sigui molt adequada per la detecció d'SNPs (Ryan *et al*, 1999; Mein *et al*, 2000). De fet, Third Wave

Technologies disposa d'assajos de detecció d'SNPs d'aplicació clínica (<http://www.twt.com>).

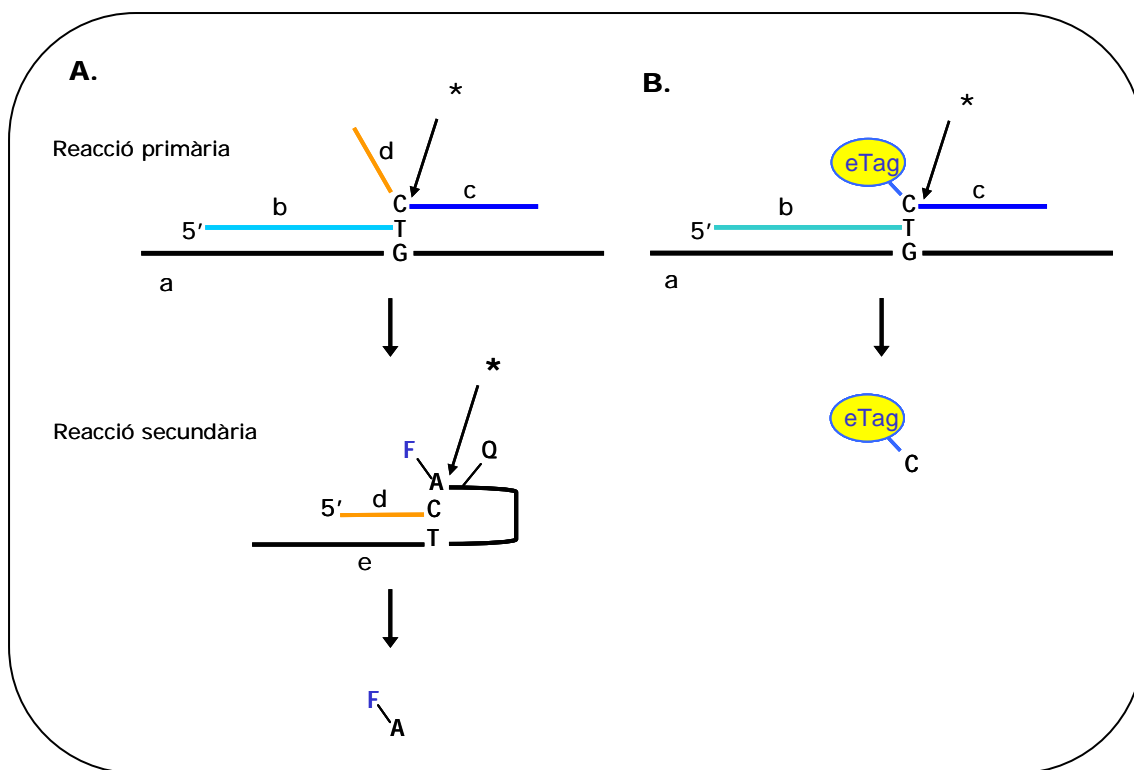


Figura 7. Discriminació al·lelica per la tècnica Invader. A. Tècnica Invader Square. B. Combinació de la tècnica Invader amb els *eTag reporters*. a, DNA diana; b, sonda Invader; c, sonda al·lel específica; d, *flap*; e, *cassette* FRET (transferència d'energia per ressonància de fluorescència); * punt de tall per l'enzim 5'nucleasa o Cleavase.

Per la identificació d'un SNP es requereix, doncs, del disseny d'una sonda Invader i dues al·lel específiques, una per cada al·lel. El nucleòtid de l'extrem 3' de la sonda Invader no cal que sigui complementari al DNA diana i no hi ha preferència per cap dels quatre nucleòtids (Lyamichev *et al*, 1999). Les sondes al·lel específiques es marcaran amb una *flap* diferent en el seu extrem 3' la qual hibrida de forma específica a un *cassette* FRET. L'amplificació del senyal i el tall de múltiples *flaps* per molècula de DNA, té lloc per incubació de la reacció a una temperatura propera a la T_m de les sondes al·lel específiques, les quals estan en excés. Sota aquestes condicions, la sonda tallada es dissocia ràpidament del DNA diana i és reemplaçada per una sonda intacta, iniciant-se un nou cicle. El nombre de sondes tallades per molècula diana i minut és aproximadament 30 min^{-1} , que resulta amb una amplificació del senyal de 10^{e3} a 10^{e4} en una reacció de 2 a 4 hores (Lyamichev *et al*, 1999; 2000). La sonda Invader es dissenya que tingui una T_m més elevada que les sondes al·lel específiques i es mantingui hibridada al DNA diana a la temperatura de la reacció. Així, doncs, sota unes condicions isoterms, múltiples sondes són tallades

per molècula de DNA i dóna lloc a una amplificació lineal de les *flap* respecte la concentració de DNA inicial.

El disseny Invader original només permet la detecció d'un SNP per reacció donat el nombre limitat de fluorocroms disponibles. Les principals avantatges d'aquesta tècnica són tres: *i.* Assaig homogeni; *ii.* Assaig quantitatiu (amplificació lineal); *iii.* No dependent de PCR (detecció directa en DNA genòmic).

No obstant, requereix d'una quantitat considerable de DNA genòmic (>100 ng per reacció). És per això que, molts dels grups que utilitzen aquesta tècnica amplifiquen les regions genòmiques que contenen els SNPs d'interès per PCR simple o multiplicadora (Mein *et al*, 2000; Ohnishi *et al*, 2001; Nevilie *et al*, 2002).

La combinació de la tècnica Invader amb els *eTag reporter* es basa en la substitució de la *flap* per un *eTag reporter* específic (Figura 7.B). Aquesta nova aproximació rep el nom *eTag Multiplex Invader SNP Assay* ja que permet l'anàlisi de múltiples SNPs en una mateixa reacció. El desenvolupament, optimització i validació d'aquest mètode ha estat realitzat per la doctoranda a l'empresa ACLARA BioSciences Inc (vegeu Estudi 3).