

**POLIPOSI ADENOMATOSA FAMILIAR I CÀNCER
COLORECTAL**
Estudi genòmic i anàlisi d'alteracions de la via de Wnt

Memòria presentada per

Antònia Obrador Hevia

Per optar al Grau de

Doctor

Tesi realitzada sota la direcció del
Dr. Gabriel Capellà i Munar
a l'Institut Català d'Oncologia

Tesi adscrita al departament de Genètica de la
Facultat de Biologia
Universitat de Barcelona
Programa de Genètica (Bienni 2002-2004)
Tutora: Dra. Marta Pascual

Gabriel Capellà

Marta Pascual

Antònia Obrador

Barcelona, Abril de 2007

Resultats

Una impossibilitat probable és sempre preferible a una possibilitat no convincent
Aristòtil

*We wish to suggest a structure for the salt of deoxyribose nucleic acid (DNA).
This structure has novel features which are of considerable biological interest*
James Dewey Watson and Francis Crick

Capítol 1

Caracterització molecular de la
poliposi adenomatosa familiar

Les mutacions germinals en el gen *APC* són les responsables de la majoria de casos de poliposi adenomatosa familiar. S'ha demostrat que *APC* participa en diferents processos cel·lulars que inclouen la regulació del cicle cel·lular, l'apoptosi, l'adhesió cel·lular, la migració cel·lular, la transducció de senyals, l'acoblament de microtúbuls i la segregació dels cromosomes. Encara que tots aquests processos estan potencialment lligats amb el càncer, sembla que la funció supressora tumoral del gen resideix en la seva capacitat per regular de manera acurada els nivells intracel·lulars de la β -catenina. Per tant, el paper d'*APC* en la via de Wnt sembla que té una implicació clau en la tumorigènesi. Aquesta via és important en el desenvolupament embrionari per determinar llinatges cel·lulars específics i en l'homeòstasi cel·lular de determinats teixits com la mucosa intestinal. Les proteïnes que participen en la via han de ser finament regulades, ja que la cèl·lula ha de respondre a diferents situacions a través de variacions en aquestes proteïnes. Per la mateixa raó, si aquestes proteïnes es veuen alterades per mutacions o canvis en els seus nivells de transcripció, la seva funció es pot veure compromesa i afectar la correcta senyalització.

1. Alteracions somàtiques d'*APC* en la poliposi adenomatosa familiar

L'estudi es va centrar principalment en la identificació de les alteracions en el gen *APC* en adenomes de pacients de FAP, ja que aquestes alteracions són la causa principal de la malaltia i representa una informació clau per a tot estudi posterior amb aquestes mostres.

Es varen triar sis pacients de FAP provinents de sis famílies diferents. El criteri de selecció va ser que la mutació germinal en el gen *APC* fos coneguda (prèviament identificada per l'equip de diagnòstic del laboratori) i que el diagnòstic de la malaltia fos clar. En moltes ocasions poden aparèixer casos que es classifiquen com a FAP, però que en realitat no ho són. Alguns casos que no presenten més de 100 adenomes són a vegades classificats com a FAP, quan es poden tractar de poliposi adenomatosa familiar atenuada o d'altres síndromes. Els sis casos seleccionats eren FAPs clàssiques i presentaven més de 100 adenomes i en alguns casos altres manifestacions extracolòniques. La taula M.1 (vegeu l'apartat de materials mètodes) presenta dades sobre la història clínica d'aquests sis pacients. L'annex 1 mostra els arbres de les famílies on s'indica amb una fletxa l'individu que correspon al pacient d'estudi.

Les biòpsies dels adenomes es varen recollir a partir de la peça de colectomia que arribava al departament d'anatomia patològica de l'hospital seguint un protocol de treball preestablert. La peça havia d'arribar en fresc (normalment les peces que arriben a anatomia patològica des del quiròfan són immerses en formol) i es procedia a dividir de manera visual el còlon en cinc zones com les que s'indiquen en la figura 1.1. De cada zona es varen recollir un mínim de deu adenomes sempre que va ser possible i cada un d'aquests adenomes es va classificar segons la zona a la qual pertanyia i a la seva mida (<10mm, entre 10 i 20 mm i >20mm de diàmetre). De cada zona es varen recollir també mostres de mucosa macroscòpicament normal (lliure d'adenomes). En cap d'aquests sis casos hi va haver carcinoma i, per tant, no es va poder recollir. Les biòpsies es varen congelar immediatament per evitar-ne la degradació.

Per als estudis subsegüents es varen seleccionar deu adenomes de cada cas, intentant que fossin representatius de les cinc zones i tres mides diferents. La taula M.2 (vegeu l'apartat de materials i mètodes) indica les característiques físiques dels 60 adenomes estudiats.

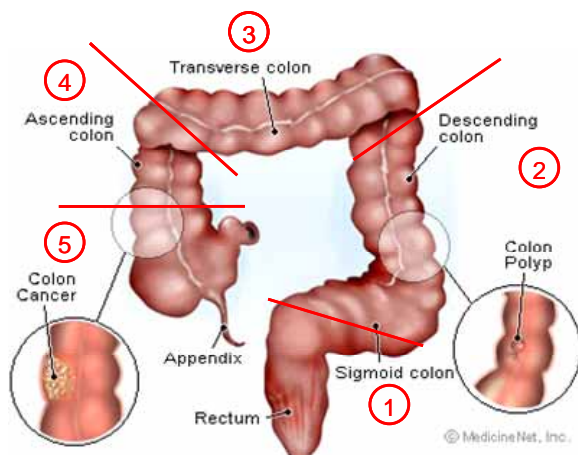


Figura 1.1. Divisió del còlon en cinc zones. 1: recte; 2: còlon descendent; 3: còlon transvers; 4: còlon ascendent; 5: cec.

Les mutacions en *APC* se seleccionen al llarg de la progressió tumoral per la seva habilitat per proporcionar un nivell òptim d'activació de la β -catenina en les cèl·lules tumorals. L'origen de l'avantatge selectiu de les diferents mutacions d'*APC* pot ser sigui el balanç entre unió i degradació de β -catenina o el balanç entre les funcions de les regions N i C terminals i pot ser que es donin per un efecte dominant negatiu per la dimerització d'*APC*.

En els adenomes analitzats s'ha trobat pèrdua al·lèlica en 26 dels 60 adenomes (43%) (figura 1.2). S'ha trobat LOH en 11/20 (55%) adenomes de pacients amb mutacions germinals properes al codó 1.300 (PCF4: c.4175C>A; p.Ser1392X i PCF9: c.3183_3187del AAACA; p.K1061KfsX2) i en 15/40 (37,5%) en la resta de pacients (p-valor 0.155, test exacte de Fisher) (taula 1.1).

Mutació germinal	LOH+	LOH-
Propera codó 1.300	11/20	9/20
Llunyana codó 1.300	15/40	25/40

Fisher's Exact test p-valor 0,155

Taula 1.1. Relació entre la localització de la mutació germinal i la presència d'LOH (pèrdua al·lèlica).

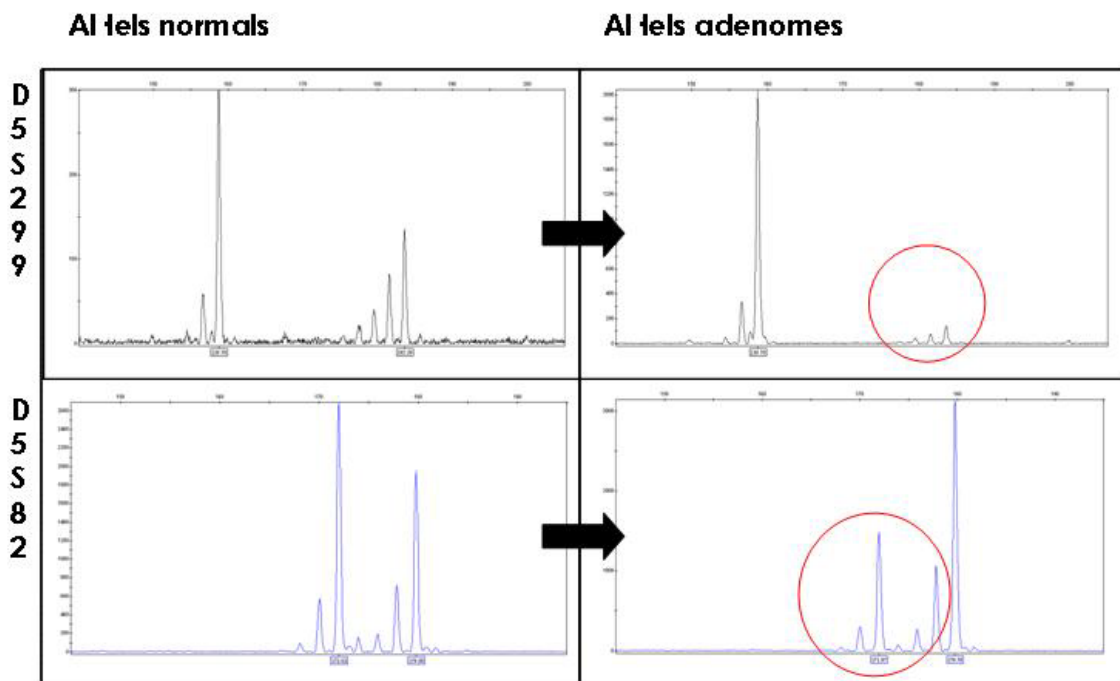


Figura 1.2. Exemples de pèrdues al·lèliques en adenomes de FAP. A) Pèrdua total de l'al·lel al locus *D5S299*. B) Pèrdua parcial (deguda a la policlonalitat o existència de teixit normal contaminant) de l'al·lel al locus *D5S82*.

És probable que amb la mida mostral no hi hagi potència estadística per trobar l'associació entre mutacions germinals properes al codó 1.300 i presència d'LOH com a *second hit* proposada per Crabtree *et al.* (Crabtree *et al.*, 2003) (vegeu introducció). Curiosament els dos casos que presenten mutacions germinals prop del codó 1.300 tenen comportaments oposats. El cas PCF4 no presenta cap adenoma amb mutacions puntuals en l'MCR i el cas PCF9 és el pacient que en presenta més (n=2). Els estudis que presenten associacions d'aquest tipus fan anàlisis amb més de 200 adenomes. Cal destacar que en el present estudi amb 60 adenomes s'ha detectat un 43% d'LOH, mentre que en l'estudi de Crabtree *et al.* amb 233 adenomes han detectat només un 22% d'LOH i això suposa una diferència cabdal que afecta de manera important l'anàlisi d'associació.

Pel que fa a les mutacions puntuals a l'MCR, s'han identificat 4/60 mutacions somàtiques truncants (6,66%), totes són mutacions *nonsense*. Una d'elles és una transició de C>T i les altres 3 són transversions (taula 1.2). Cal destacar que d'aquestes quatre mutacions detectades totes són molt properes al codó 1.300 i això reforça la idea que són aquestes mutacions les que són més freqüentment seleccionades per abolir de manera més significativa la funció supressora tumoral d'APC.

Pacient Mutació germinal	Pèrdua al·lèlica	Mutació a l'MCR				Adenomes amb alteracions somàtiques
		n	ID adenoma	Mutació truncant	Alteració molecular	
PCF1 c.1958+1G>A; r.spl	5/10	0/10				5/10
PCF3 c.1958+3G>A; r. spl c.1959G>A; ("p.=")	3/10	1/10	AD2	Q1338X	C>T	4/10
PCF4 c.4175C>A; p.Ser1392X	4/10	0/10				4/10
PCF6 c. 4612_4613delGA; p.Glu1538IlefsX5	3/10	0/10				3/10
PCF8 c.1262_1263delGGinsAA; p.Trp421X	4/10	1/10	AD7	E1397X	G>T	4/10
PCF9 c. 3183_3187del AAACA; p.K1061KfsX2	7/10	2/10	AD1 AD4	S1356X E1309X	C>G G>T	7/10

Taula 1.2. Alteracions somàtiques del gen APC detectades en els 60 adenomes de PAF.

2. Determinació del nivell d'expressió i localització cel·lular de la β -catenina

S'ha demostrat que la inactivació bial·lèlica d'*APC* en les cèl·lules tumorals provoca l'increment dels nivells de β -catenina (CTNNB1). La tinció nuclear intensa de β -catenina és un tret immunohistològic ben conegut en cèl·lules tumorals (Blaker *et al.*, 2003). Aquest tret no només es restringeix a tumors amb mutacions que tenen *APC* inactivat o mutacions activadores de β -catenina. Per tant, a part de l'efecte directe de les mutacions d'*APC* sobre la localització de la β -catenina, han d'existir altres mecanismes que condueixin al mateix fenomen.

Es va realitzar una immunohistoquímica per detectar els nivells i la localització de la β -catenina en els 60 adenomes de la sèrie de FAP i les seves mucoses normals corresponents. Mentre que en cap de les mucoses normals s'ha trobat β -catenina nuclear. En un 83% dels adenomes (50/60) s'ha trobat tinció clara de β -catenina nuclear, fet que indica una activació de la via de Wnt a través de la funció transcripcional de la β -catenina (figura 1.3 B).

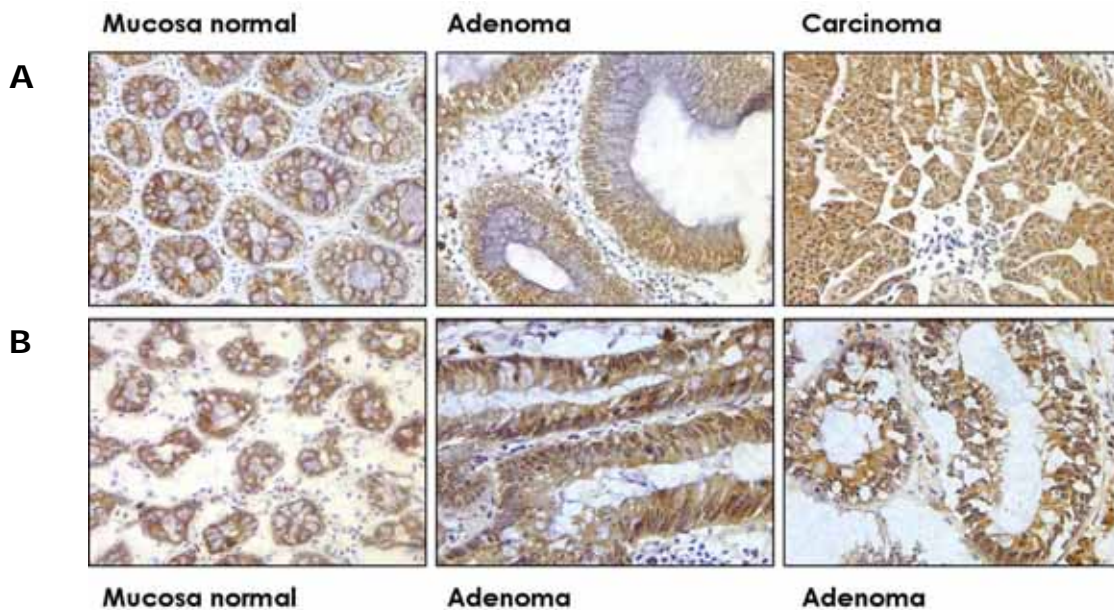
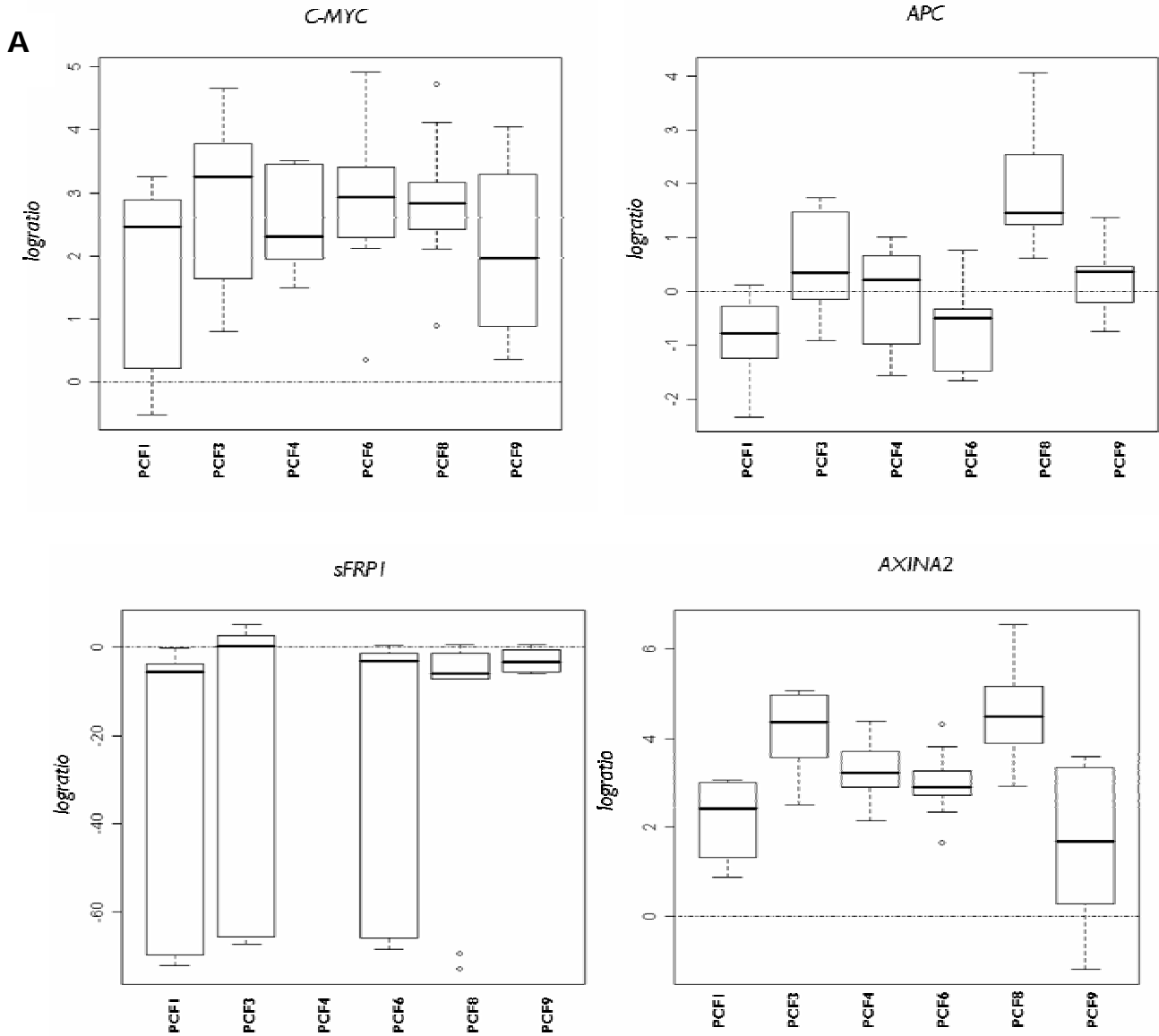


Figura 1.3. Immunohistoquímica de la β -catenina. A) Sèrie adenoma-carcinoma de càncer colorectal esporàdic. B) Mucosa i dos adenomes representatius de la tinció de la β -catenina en mostres de FAP.

A més, es varen fer immunohistoquímiques de mostres seriades de mucosa normal, adenoma i carcinoma (n mucosa normal= 4; n adenomes=4; n carcinomes= 10) de casos de càncer colorectal esporàdic per tal de determinar quin és el comportament immunohistoquímic de la β -catenina durant la progressió tumoral. S'ha pogut constatar un augment dels nivells de β -catenina i de la seva localització nuclear al llarg de l'evolució tumoral. Pel que fa a l'activació de la via de Wnt, donada per la localització nuclear de la β -catenina, s'ha detectat un 100% de tinció nuclear en els adenomes esporàdics (4/4) i un 80% en el cas dels carcinomes (8/10) (figura 1.3 A).

3. Proteïnes de la via de Wnt i els seus nivells d'expressió en el càncer colorectal

La via de Wnt controla processos del desenvolupament embrionari i processos de homeòstasi cel·lular en diferents òrgans i està conservada al llarg de l'evolució. El fet de controlar processos tan fins, implica que els nivells de les proteïnes que hi estan implicades hagi de ser acurat i controlat. Un fenomen que es dona en el càncer en general és la desregulació dels nivells d'expressió de determinats gens que seria el causant del seu fenotip. D'aquesta manera, és fàcil pensar que alguns dels gens que formen part de la via de Wnt poden tenir els seus nivells alterats per afavorir la capacitat transcripcional de la via que sabem que condueix a un augment de la proliferació cel·lular. Amb l'objectiu de caracteritzar els nivells d'expressió de determinats gens de la via en mostres de càncer colorectal per comparar-les amb el teixit normal, es varen fer experiments de PCRq. Els gens que es varen analitzar són: *APC*, *C-MYC*, *AXINA2* i *SFRP1* (figura 1.4).



B

	<i>C-MYC</i>		<i>APC</i>		<i>sFRP1</i>		<i>AXINA2</i>	
	<i>foldchange</i>	<i>p-valor</i>	<i>foldchange</i>	<i>p-valor</i>	<i>foldchange</i>	<i>p-valor</i>	<i>foldchange</i>	<i>p-valor</i>
PCF1	3,43	0,019*	-1,78	0,004*	-69754,56	0,002*	5,48	0,002*
PCF3	7,61	0,002*	1,42	0,160	-472514,80	0,064	24,44	0,002*
PCF4	5,57	0,002*	-1,02	1	ND	ND	9,27	0,002*
PCF6	7,04	0,002*	-1,63	0,027*	-174663147,19	0,002*	7,94	0,002*
PCF8	7,32	0,002*	3,72	0,002*	-180295,02	0,006*	25,05	0,002*
PCF9	3,96	0,002*	1,20	0,375	-8,34	0,084	3,08	0,013*

* p-valor significatiu (<0,05)

ND: no disponible

continuació figura 1.4

C

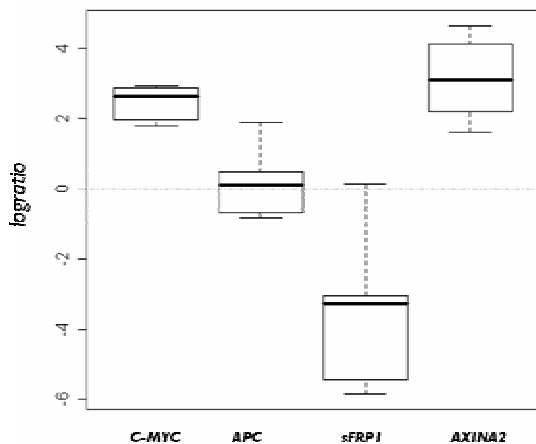
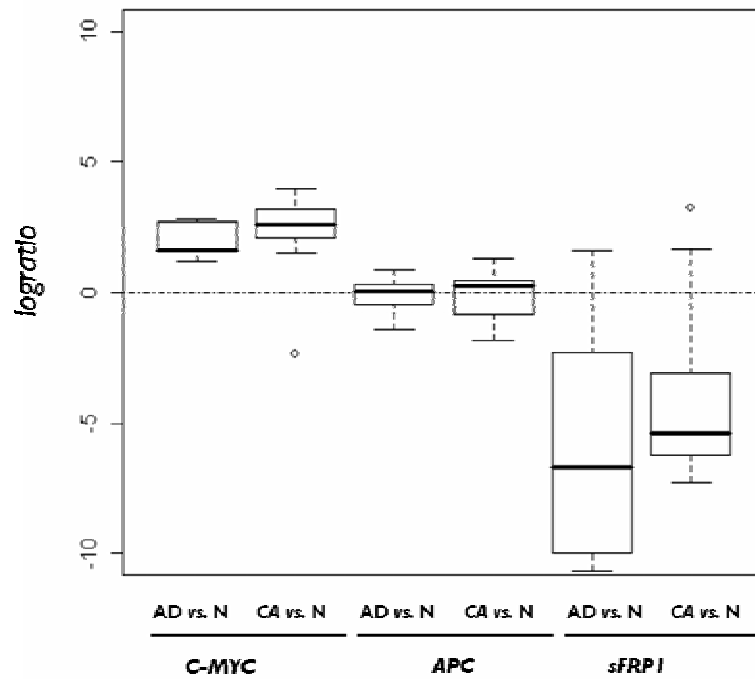


Figura 1.4. Anàlisi d'expressió dels gens *sFRP*, *APC*, *C-MYC* i *AXINA2*. A) *Boxplots* per a representar els resultats dels experiments de PCRq on es va analitzar l'expressió de *C-MYC*, *APC*, *sFRP1* i *AXINA2*. El nivell d'expressió s'ha expressat en *log2ratio* (adenoma/mucosa normal) on el valor de 0 representa que no hi ha canvi, valors positius representen sobreexpressió i valors negatius infraexpressió en l'adenoma respecte al teixit normal. El rang de valors d'*sFRP1* és ampli, ja que en molts adenomes es deixava d'expressar totalment i per no tenir un valor de *log2ratio* infinit es va posar un valor màxim d'infraexpressió que correspon a una *log2ratio* d'aproximadament -70, la resta de valors es corresponen a adenomes que no perden totalment l'expressió d'*sFRP1*. B) Es mostra una taula amb la mitjana dels *foldchange* (valor de canvi) pels 10 adenomes de cada cas i els p-valors obtinguts amb el Wilcoxon-test per avaluar la significació estadística. C) *Boxplot* resum amb la mitjana de les dades de tots els pacients per cada un dels gens.

El mateix procediment es va aplicar a deu mostres aparellades de càncer colorectal esporàdic (mucosa normal-adenoma-carcinoma), amb l'excepció del gen *AXINA2*, i es va obtenir la mitjana dels valors de *logratio* dels adenomes i carcinomes per cada gen (figura 1.5).



	<i>C-MYC</i>		<i>APC</i>		<i>sFRP1</i>	
	<i>foldchange</i>	p-valor	<i>foldchange</i>	p-valor	<i>foldchange</i>	p-valor
AD vs. N	3,87	0,009*	-1,12	0,846	-39,56	0,006*
CA vs. N	4,81	0,001*	-1,01	0,846	-14,15	0,014*

* p-valor significatiu (<0,05)

Figura 1.5. Anàlisi d'expressió dels gens *C-MYC*, *APC* i *sFRP*. *Boxplot* per a representar els resultats dels experiments de PCRq on es va analitzar l'expressió de *C-MYC*, *APC* i *sFRP* amb mostres de CCR esporàdic. El nivell d'expressió s'ha expressat en *log2ratio* (adenoma (AD) o carcinoma (CA)/ mucosa normal (N)). Es mostra una taula amb la mitjana dels *foldchange* (valor de canvi) dels deu adenomes o carcinomes i els p-valors obtinguts amb el Wilcoxon-test per avaluar la significació estadística.

Aquests resultats mostren un augment en el nivell d'expressió dels gens *C-MYC* i *AXINA2* en mostres tumorals, tant de FAP com de CCR esporàdic. Els canvis en el gen *APC* no són consistents i varien entre diferents individus, de manera que no es varen obtenir valors de canvi significatius ni es va poder establir una tendència. En els adenomes de FAP es va trobar pèrdua al·lèlica en alguns casos, de manera que sembla que no hi ha una associació directa entre l'LOH i canvis a nivell d'mRNA del gen.

Pel que fa al gen *SFRP1*, mostra una clara infraexpressió i fins i tot una manca d'expressió tant en mostres de FAP com de CCR esporàdic. *SFRP1* és un membre de la família de proteïnes sFRP que conté un domini ric en cisteïna homòleg al lloc d'unió de Wnt amb el receptor Frizzled. sFRP actua com a modulador de la senyalització a través de Wnt, ja que impedeix la unió de Wnt amb el seu receptor. Les mostres tumorals analitzades presenten nivells inferiors d'expressió respecte a la seva mucosa normal corresponent. Caldwell *et al.* varen descriure el mateix fenomen en dotze adenomes i en carcinomes i varen indicar un paper supressor tumoral per a aquest gen (Caldwell *et al.*, 2004, Caldwell *et al.*, 2006).

C-MYC és un potent activador del programa de proliferació de la cèl·lula. Actualment es considera el gen diana per excel·lència de la via de Wnt. Com caldria esperar en un context on *APC* es troba inactivat, els nivells de *C-MYC* són sobreexpressats en la majoria d'adenomes i carcinomes analitzats.

L'axina2 (també coneguda com a conductina) forma part d'un complex multiproteic juntament amb *APC*, β -catenina i *GSK3 β* que contribueix a la degradació de β -catenina. Els resultats presentats indiquen sobreexpressió d'aquest gen en adenomes de FAP, circumstància que també ha estat descrita per a la proteïna per Lustig i col·laboradors (Lustig *et al.*, 2002). A més, *AXINA2* és un gen diana de β -catenina i en teixits normals això pot representar un bon mecanisme de *feedback* negatiu quan la via és activa.

Els resultats mostren una desregulació global dels components de la via de Wnt. En estadis primerencs de la progressió tumoral, les mutacions d'*APC* no només promouen l'acumulació de β -catenina, sinó que també es donen alteracions a nivell transcripcional que afecten gens importants en la regulació d'aquesta via de senyalització.

4. Identificació de mutacions en el gen *KRAS*

Les mutacions en gens de la família de *RAS* s'han descrit com a fortament implicades en els processos de progressió neoplàstica. En els humans, s'han identificat mutacions en els gens *HRAS1*, *KRAS2* i *NRAS* en diferents tumors entre els quals es troben la leucèmia mieloide, el carcinoma del còlon i el carcinoma de pàncrees, entre d'altres. En els carcinomes de còlon, s'ha observat que les mutacions en el gen de *RAS* es donen en adenomes propers als carcinomes. Això suggereix que les mutacions en aquest gen no són només un esdeveniment primerenc en el desenvolupament del càncer, sinó que contribueixen directament al fenotip maligne (Burmer & Loeb, 1989). El 40% dels adenomes i carcinomes contenen mutacions en aquest gen i la mutació més freqüent amb molta diferència és una transició de G a A en la segona posició dels codons 12 o 13, que provoca un canvi de glicina a àcid aspàrtic.

Amb l'objectiu de detectar les mutacions de *KRAS* presents en els 60 adenomes de FAP analitzats i descriure així el seu grau d'evolució, es va seqüenciar el fragment del gen que conté els codons 12 i 13 que és on es localitzen més freqüentment les mutacions.

Es varen identificar mutacions en el 10% dels adenomes (6/60) totes localitzades en el codó 12 a la primera i segona posició (taula 1.3). En dos dels adenomes es va trobar la mutació més freqüent que és un canvi de G a A en la segona posició del codó 12.

Es va confirmar la baixa prevalència de mutacions en el gen que prèviament havien senyalat Vogelstein i col·laboradors en 40 adenomes inferiors a 1 cm de set pacients de FAP (Vogelstein *et al.*, 1988).

Pacient	Adenomes amb mutació a <i>KRAS</i>	Identificador adenoma	Mutació	Tipus de mutació
PCF1	1/10	AD2	G12A	G>C
PCF3	3/10	AD1	G12C	G>T
		AD3	G12S	G>A
		AD7	G12D	G>A
PCF4	0/10			
PCF6	0/10			
PCF8	1/10	AD3	G12D	G>A
PCF9	1/10	AD7	G12A	G>C

Taula 1.3. Mutacions en el gen *KRAS* en adenomes de FAP.

5. Canvis en el genoma detectats per *array comparative genomic hybridization*

Es varen realitzar experiments d'*array comparative genomic hybridization* amb mostres d'adenomes de pacients de PAF. La plataforma elegida per a fer les hibridacions va ser el *microarray* d'oligonucleòtids de fabricació pròpia a la Universitat d'Àmsterdam (vegeu materials i mètodes per a més detalls sobre la plataforma).

Es va observar que divuit adenomes (62%) dels 29 analitzats tenen almenys un guany i que setze adenomes (55%) tenen almenys una pèrdua. Només dos adenomes (6,8%) no tenen cap canvi de nombre de còpies.

Es va determinar la magnitud d'aberracions de nombre de còpia en cada adenoma calculant el percentatge de canvis en les 129 regions analitzades (representatives de tot el genoma). Els tumors varen ser classificats com a grau baix (< 5% de canvis), grau intermedi (5-10 % de canvis) i grau alt (> 10% de canvis) en funció del nombre de canvis per cada adenoma respecte al total de canvis (figura 1.8).

La majoria dels adenomes corresponen al grup de grau baix (69%), mentre que un 21% pertanyen al grup de grau intermedi i un 10% al grup de grau alt. Es varen detectar un total de 180 canvis de nombre de còpia, 75 dels quals eren pèrdues (42%) i 105 eren guanys (58%). Per cada regió cromosòmica, es va calcular el nombre de tumors pels quals hi havia pèrdua o guany i se'n va extreure el percentatge respecte al total de 29 adenomes analitzats (figura 1.6). Es va establir un llindar del 10% per trobar les regions més freqüentment alterades. Les regions que més freqüentment es guanyen són la regió 5q31 del cromosoma 5, tot el cromosoma 7 amb la regió 7q22.1 més recurrentment alterada, tot el cromosoma 13 (guanyat en un 10% dels casos) i la regió 20p11.22 del cromosoma 20. En la regió 5q31 es troba el *locus* de la protocadherina beta i gamma (*PCDHB1*). En el cromosoma 7, que es guanya sencer, hi trobem la regió 7q22.1 que és més recurrent la qual conté diversos gens entre els quals hi ha *CUTL1*, *SH2B2*, *ALKBH4*, *POLR2J* i *HSPC047*. El cromosoma 13 conté gens coneguts com la trombospondina (*THSD1*) i la ferritina (*FTHL7*). Curiosament, la regió 20p11.22 conté el pseudogen *suppression of tumorigenicity 13 (colon carcinoma) (ST13P)*, a part del gen *PAX1*.

Les regions més freqüentment perdudes són la regió 5q15-5q31.1 del cromosoma 5, la regió 14q32.33 del cromosoma 14 i la regió 8q24.3 del cromosoma 8 (figura 1.7). En la regió 5q15-5q31.1, el *locus* que es perd més recurrentment és on es localitza el gen

APC que es perd en els adenomes per inactivar totalment la proteïna. En tots els adenomes on es detecta pèrdua d'aquest *locus*, s'ha detectat pèrdua prèviament amb els estudis d'LOH, exceptuant un dels adenomes on es detecta pèrdua per aCGH i no per LOH. En la regió 14q32.33 s'hi troba la cadena pesant de la immunoglobulina (*IGH*). La regió 8q24.3 conté els gens *block of proliferation 1 (BOP1)* i *heat shock transcription factor 1 (HSF1)*, entre d'altres.

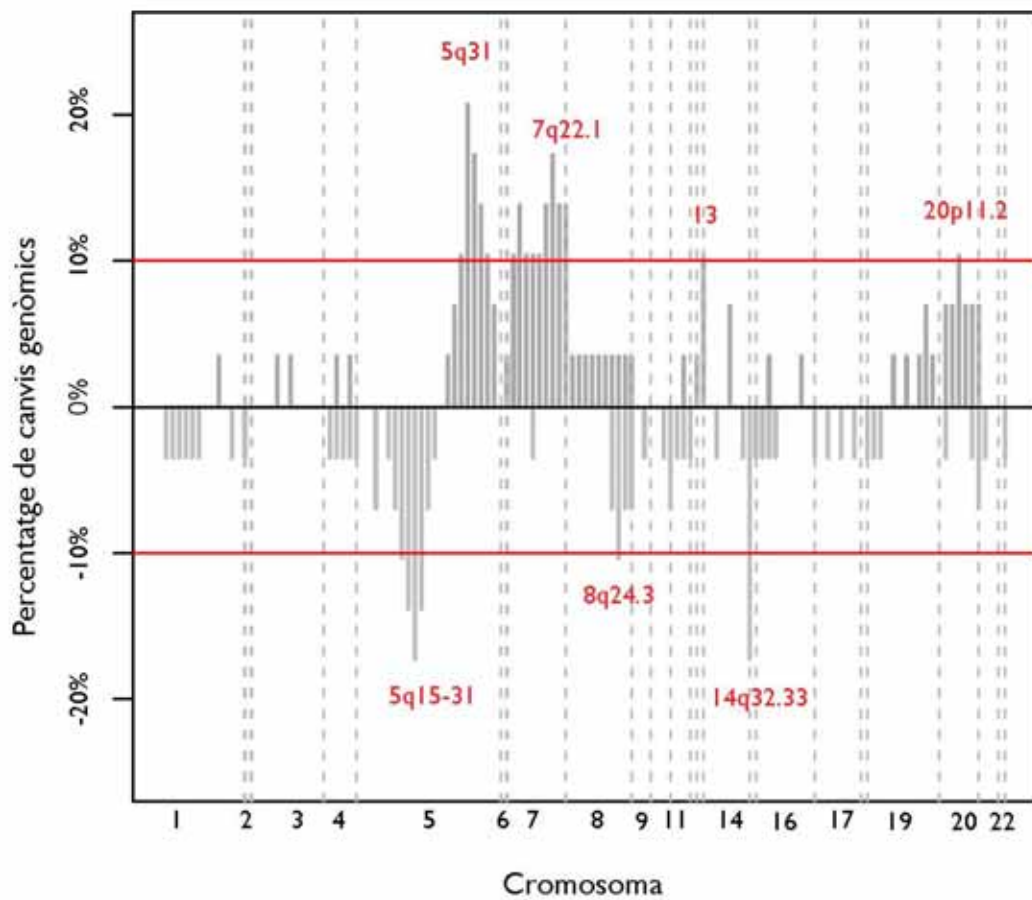


Figura 1.6. Canvis globals de guanys i pèrdues en els 29 adenomes analitzats. El percentatge de canvis (els guanys per sobre del 0 i les pèrdues per sota) es representa amb barres en cada una de les posicions representades al *microarray*. Es va establir un llindar del 10% per identificar les regions cromosòmiques amb canvis genòmics comuns.

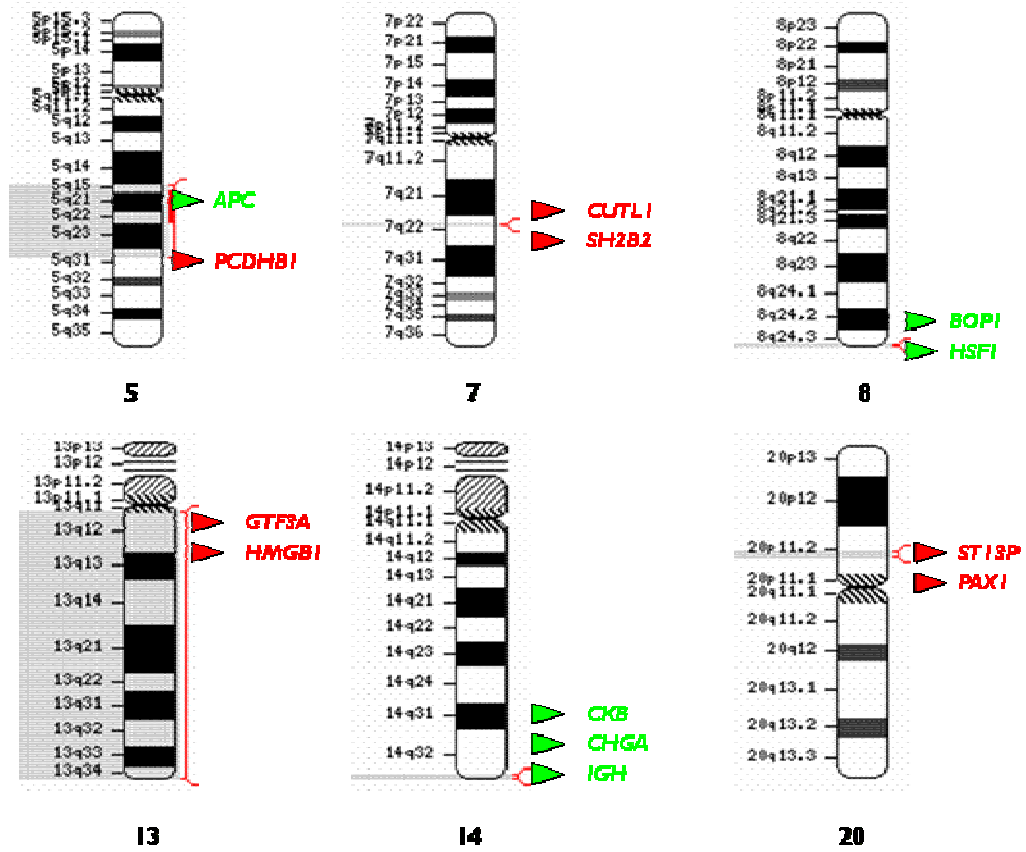


Figura 1.7. Cromosomes amb canvis genòmics. Es representen aquells adenomes amb canvis en més d'un 10% dels adenomes analitzats. Les fletxes verdes indiquen pèrdues i les vermelles indiquen guanys. Al costat de les fletxes s'indica el nom del gen candidat de cada regió.

6. Representació global de les alteracions dels adenomes de FAP

S'ha fet una gràfica que engloba tots els canvis observats per cada un dels adenomes de FAP analitzats (figura 1.8). Tots els adenomes presenten algun tipus d'alteració, ja sigui en el DNA (mutacions, LOH o canvis de nombre de còpia) o bé en l'RNA (canvis d'expressió).

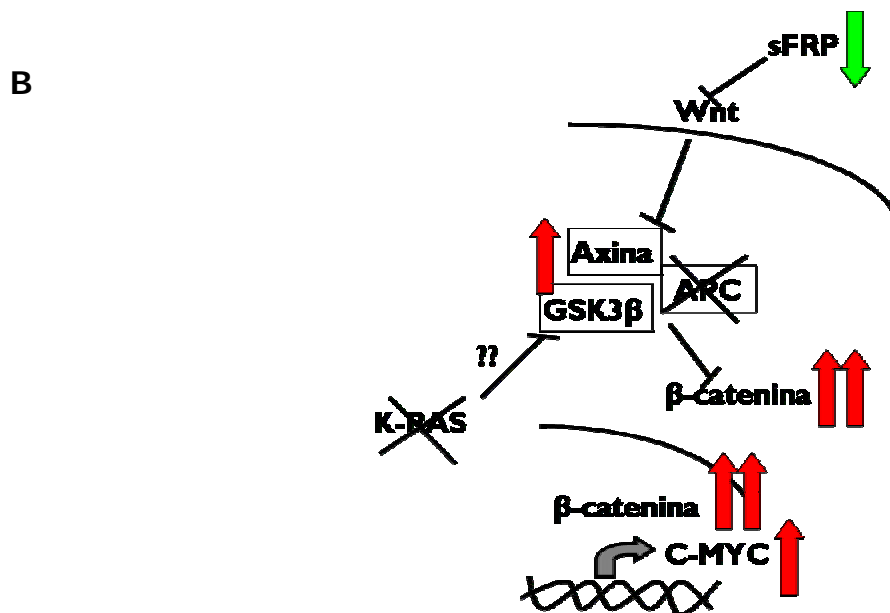
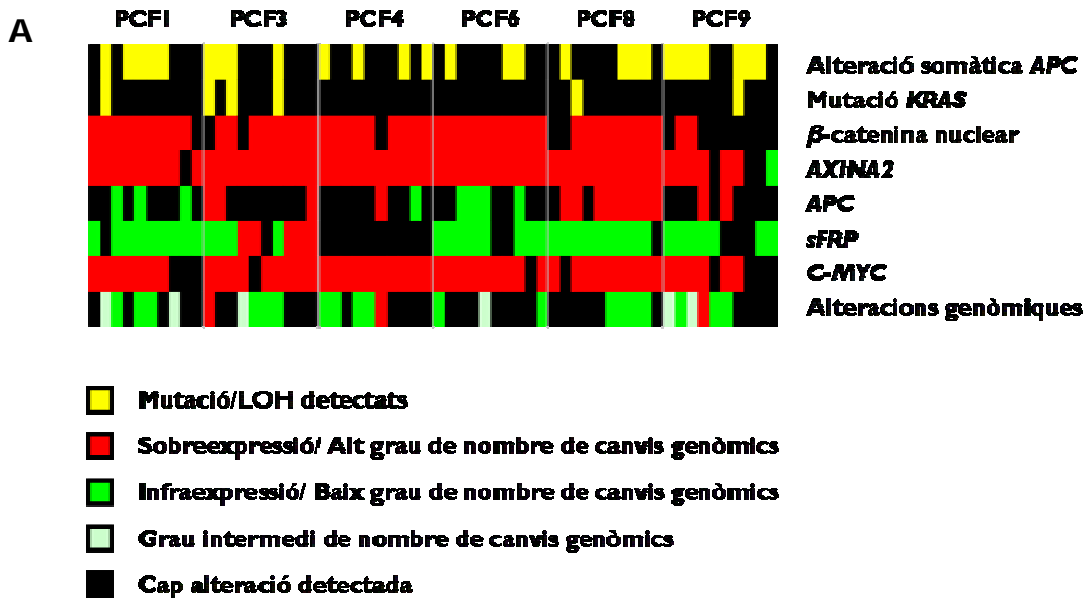


Figura 1.8. Representació global de les alteracions moleculars estudiades en la sèrie d'adenomes de FAP provinents de 6 pacients (PCF1, PCF3, PCF4, PCF6, PCF8, PCF9). (A) Per cada pacient hi ha 10 adenomes analitzats. El color negre indica que no hi ha canvi, excepte per l'última fila (Alteracions genòmiques) que indica que l'adenoma no s'ha analitzat. (B) Proposta de model de la via de Wnt alterada en CCR.

Capítol 2

APC i inestabilitat cromosòmica en
el càncer colorectal

S'ha proposat que el gen *APC* té un paper en la inestabilitat cromosòmica del càncer colorectal. Aquesta funció va ser suggerida després d'observar cèl·lules mare embrionàries de ratolí portadores de mutacions en el gen *APC* que condueixen a la formació d'una proteïna truncada en l'extrem carboxi-terminal (Fodde *et al.*, 2001, Kaplan *et al.*, 2001). *APC* interacciona amb la proteïna d'unió als microtúbuls EB1 que es requereix per a la correcta funció en la dinàmica d'aquestes estructures del citoesquelet (Green *et al.*, 2005). A part d'aquestes observacions sobre la contribució d'*APC* a la inestabilitat en càncer colorectal, la CIN és encara en bona part desconeguda a nivell molecular.

1. Inactivació del gen *APC* i expressió dels gens *BUB1B*, *MAD2L1* i *STMN1*

Tal i com es descriu a l'article adjunt a l'annex 2 (Abal *et al.*, 2007), es va estudiar el paper de la inactivació del gen *APC* en les alteracions durant la progressió de la mitosi en tumors del còlon en models animals i mostres humanes. Els models animals d'elecció foren un model de FAP, el ratolí *APC*^{1638N} i, en contraposició, el model que desenvolupa tumors colorectals en absència de mutacions d'*APC*, el ratolí pVillin-*KRAS*^{V12G}.

Amb l'estudi de la progressió del cicle cel·lular de cèl·lules provinents d'adenomes humans i de models animals per la tècnica de video-microscòpia es va observar que aproximadament un 20% de cèl·lules que estaven en mitosi eren capaces de separar les cromàtides germanes, però no eren capaces de finalitzar la segregació cel·lular i, com a conseqüència, apareixien cèl·lules polinucleades. Es varen trobar diferències estadísticament significatives quan es comparaven aquestes dades amb les de cèl·lules on *APC* no està mutat (ratolins pVillin-*KRAS*^{V12G} i cèl·lules epitelials intestinals normals) (p-valor 0,0006).

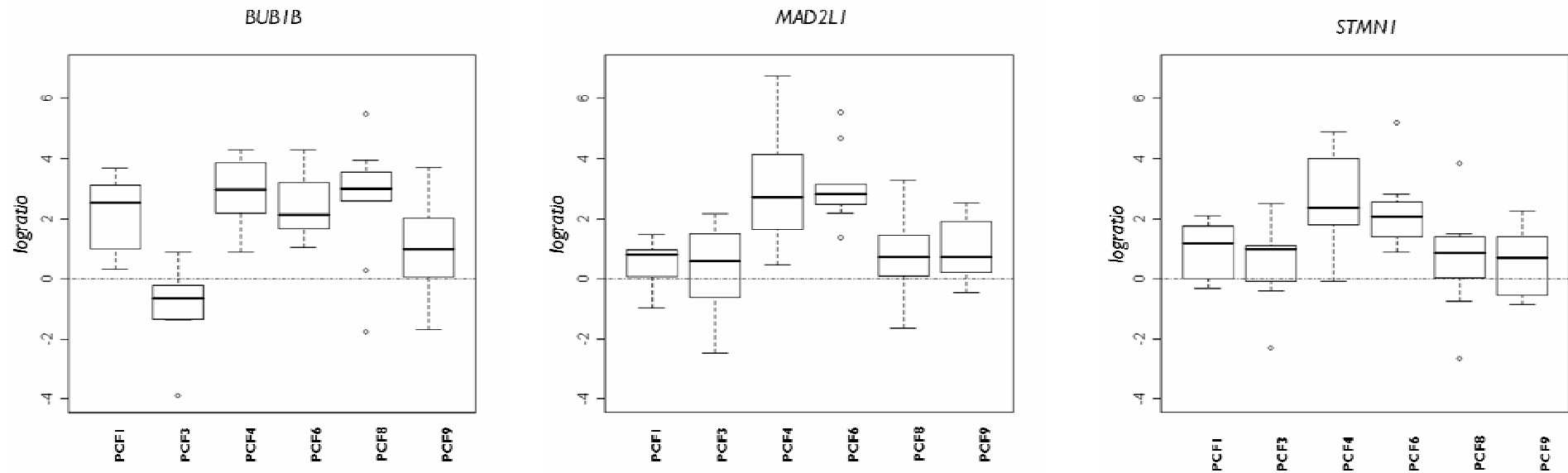
Per aprofundir en la base molecular de les anomalies observades durant la mitosi es varen fer experiments d'anàlisi massiva del transcriptoma amb *microarrays*. Les lesions provinents de ratolins pVillin-*KRAS*^{V12G} (són diploides, no presenten inestabilitat genètica ni defectes en la mitosi) es varen comparar amb lesions provinents de ratolins *APC*^{+1639N}/ pVillin-*KRAS*^{V12G}. L'anàlisi del transcriptoma va permetre identificar un conjunt de gens relacionats amb mitosi que es trobaven diferencialment expressats entre els dos tipus de ratolins (taula 2.1).

Gen	Descripció del gen	Foldchange
<i>STMN1</i>	<i>stathmin 1/oncoprotein 18</i>	4,7702
<i>CCNB2</i>	<i>cyclin B2</i>	3,6262
<i>RANBP1</i>	<i>RAN binding protein 1</i>	3,4847
<i>BUB1B</i>	<i>budding uninhibited by benzimidazoles 1 homologue b</i>	3,4311
<i>CCNB1</i>	<i>cyclin B1</i>	3,2234
<i>NUSAP1</i>	<i>nucleolar and spindle associated protein 1</i>	2,9636
<i>WEE1</i>	<i>WEE1 homologue</i>	2,704
<i>SMC4L1</i>	<i>SMC4 structural maintenance of chromosomes 4-like 1</i>	2,6086
<i>RAN</i>	<i>RAN, member RAS oncogene family</i>	2,5616
<i>CNAP1</i>	<i>chromosome condensation-related SMC-associated protein 1</i>	2,554
<i>MAD2L1</i>	<i>MAD2 mitotic arrest deficient-like 1</i>	2,487
<i>CETN3</i>	<i>centrin, EF-hand protein, 3</i>	2,1549
<i>ANAPC4</i>	<i>anaphase promoting complex subunit 4</i>	2,0603

Taula 2.1. Gens diferencialment expressats en un model animal de càncer colorectal. Augment d'expressió (*foldchange*) detectat en ratolins *APC^{+/-1639N}/pVillin-KRAS^{V12G}* quan són comparats amb ratolins *pVillin-KRAS^{V12G}*.

Entre aquests gens diferencialment expressats es va posar especial atenció al gen amb un canvi més alt: *STMN1* i als gens implicats en el control del *checkpoint* de mitosi: *BUB1B* i *MAD2L1*. *STMN1* és un gen que pertany a la família de gens d'*stathmin*. Codifica per una fosfoproteïna citoplasmàtica ubiqua que està implicada en la regulació del sistema de microtúbuls, prevenint l'acoblament, tot desestabilitzant-los. *BUB1B* codifica per una cinasa localitzada al cinetocor que té un paper en la inhibició d'*anaphase promoting complex/cyclosome* (APC/C) i retarda l'inici de l'anafase per assegurar una correcta segregació cromosòmica. *MAD2L1* és també un component del *checkpoint* del fus mitòtic que impedeix l'entrada en anafase fins que tots els cromosomes es troben degudament alineats en la placa mitòtica (vegeu la figura IN.10 de la introducció).

Per confirmar i validar els resultats trobats al model animal, es va analitzar el nivell d'expressió d'aquests tres gens en la sèrie d'adenomes de FAP prèviament caracteritzada (capítol 1). Per PCRq es varen obtenir nivells d'expressió relatius de cada adenoma amb la seva mucosa corresponent (vegeu els materials i mètodes). Els valors mitjans per cada pacient varen mostrar que els tres gens tenien uns nivells d'expressió més elevats en els adenomes que en la mucosa normal, encara que es podia observar una certa variabilitat intra-cas i inter-cas (t-test *BUB1B* p-valor 0,03; *MAD2L1* p-valor 0,03; *STMN1* p-valor 0,02)(figura 2.1).

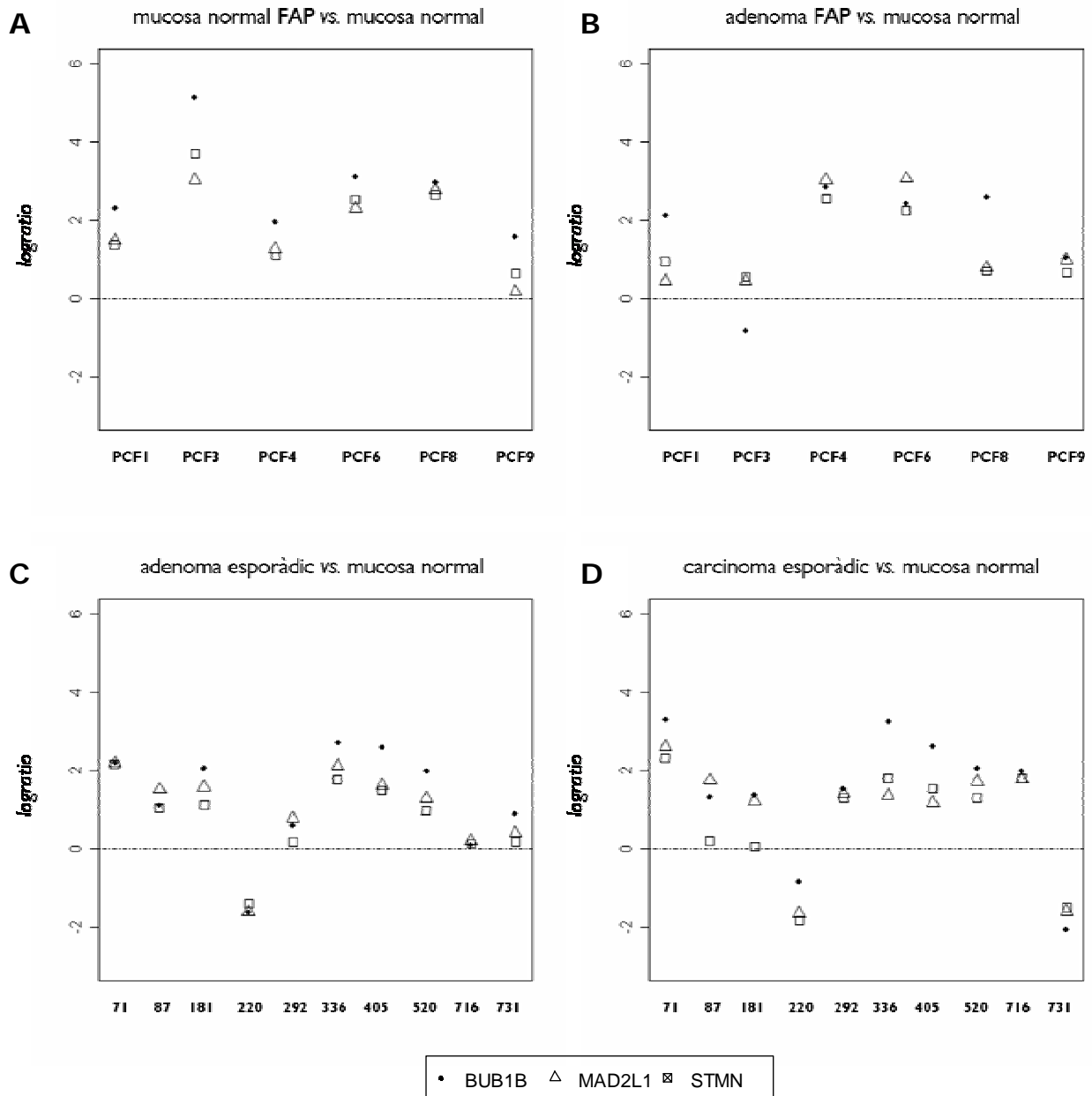


	<i>BUB1B</i>		<i>MAD2L1</i>		<i>STMN1</i>	
	<i>foldchange</i>	<i>p</i> -valor	<i>foldchange</i>	<i>p</i> -valor	<i>foldchange</i>	<i>p</i> -valor
PCF1	4,31	0,002*	1,36	0,168	1,92	0,019*
PCF3	-1,79	0,064	1,36	0,432	1,46	0,232
PCF4	7,12	0,002*	8,13	0,002*	5,82	0,004*
PCF6	5,37	0,002*	8,37	0,002*	4,74	0,002*
PCF8	5,96	0,006*	1,72	0,084	1,63	0,105
PCF9	2,03	0,084	1,98	0,014*	1,59	0,160

* *p*-valor significatiu (<0,05)

Figura 2.1. Anàlisi d'expressió dels gens *BUB1B*, *MAD2L1* i *STMN1*. El nivell d'expressió s'ha expressat en *log2ratio* (adenoma/ mucosa normal). La majoria dels valors estan per sobre de 0, fet que indicaria la sobreexpressió d'aquests tres gens en la majoria dels pacients. La taula indica la mitjana dels valors de *foldchange* (valor de canvi) dels 10 adenomes de cada cas i els *p*-valors obtinguts després d'aplicar el Wilcoxon-test.

Un fet destacable és que els nivells d'expressió dels adenomes d'un mateix pacient per als tres gens són molt similars, fet que suggereix una possible coexpressió (figura 2.2 B). Un altre fet interessant és que quan es varen comparar les mucoses normals d'aquests pacients de FAP amb un *pool* de mucoses provinents de pacients de càncer colorectal esporàdic, aquests tres gens també mostraven un nivell d'expressió més alt en les mucoses normals de FAP (t-test *BUB1B* p-valor 0.002; *MAD2L1* p-valor 0.008; *STMN1* p-valor 0.008) (figura 2.2 A). Aquests pacients de FAP presenten una mutació d'*APC* germinal coneguda i els resultats observats poden suggerir que la inactivació d'un sol dels al·lells pot ser suficient per promoure nivells més alts de transcripció de determinats gens.



continuació figura 2.2

Figura 2.2. Nivells d'expressió relatius de *BUB1B*, *MAD2L1* i *STMN1* en mucosa normal, adenomes de FAP, adenomes esporàdics i carcinomes esporàdics. A) *log2ratio* per a cada mucosa normal de FAP/*pool* de mucoses normals; B) Valors mitjans de *log2ratio* dels 10 adenomes de cada cas/mucosa normal corresponent; C) *log2ratio* per a cada adenoma esporàdic/mucosa normal corresponent; D) *Log2ratio* per cada carcinoma esporàdic/mucosa normal corresponent.

Per estudiar el comportament d'aquests tres gens en la progressió del càncer colorectal, es va expandir l'anàlisi a un sèrie de deu adenomes i deu carcinomes esporàdics aparellats (figura 2.2 C i D). Tant adenomes com carcinomes varen mostrar nivells més alts d'mRNA comparats amb la seva mucosa normal corresponent en la majoria dels casos analitzats (8/10). Totes les comparacions varen ser estadísticament significatives, exceptuant *MAD2L1* i *STMN1* en els carcinomes (adenomes: t-test; *BUB1B* p-valor 0,01 *MAD2L1* p-valor 0,02; *STMN1* p-valor 0,04) (carcinomes: t-test *BUB1B* pvalor 0,02; *MAD2L1* p-valor 0,06; *STMN1* p-valor 0,19). Cal destacar que en dos dels pacients (amb l'identificador 220 i 731) es va observar una infraexpressió d'aquests gens. Com s'havia observat per als adenomes de FAP, els tres gens segueixen un mateix patró d'expressió en totes les mostres adenoma-carcinoma aparellades, fet que suggereix una potencial coexpressió.

Per estudiar els canvis d'expressió a nivell de proteïna i veure si es correlacionen amb els canvis a nivell d'mRNA, es va fer immunohistoquímica de *BUB1B* i *MAD2L1* en una sèrie de quatre adenomes de FAP i quatre mucoses aparellades i en una sèrie de mostres de càncer colorectal esporàdic amb deu adenomes, deu carcinomes i deu mucoses aparellades (figures 2.3 i 2.4). La sobreexpressió de *BUB1B* i *MAD2L1* es va detectar de manera clara en tots els adenomes i carcinomes analitzats.

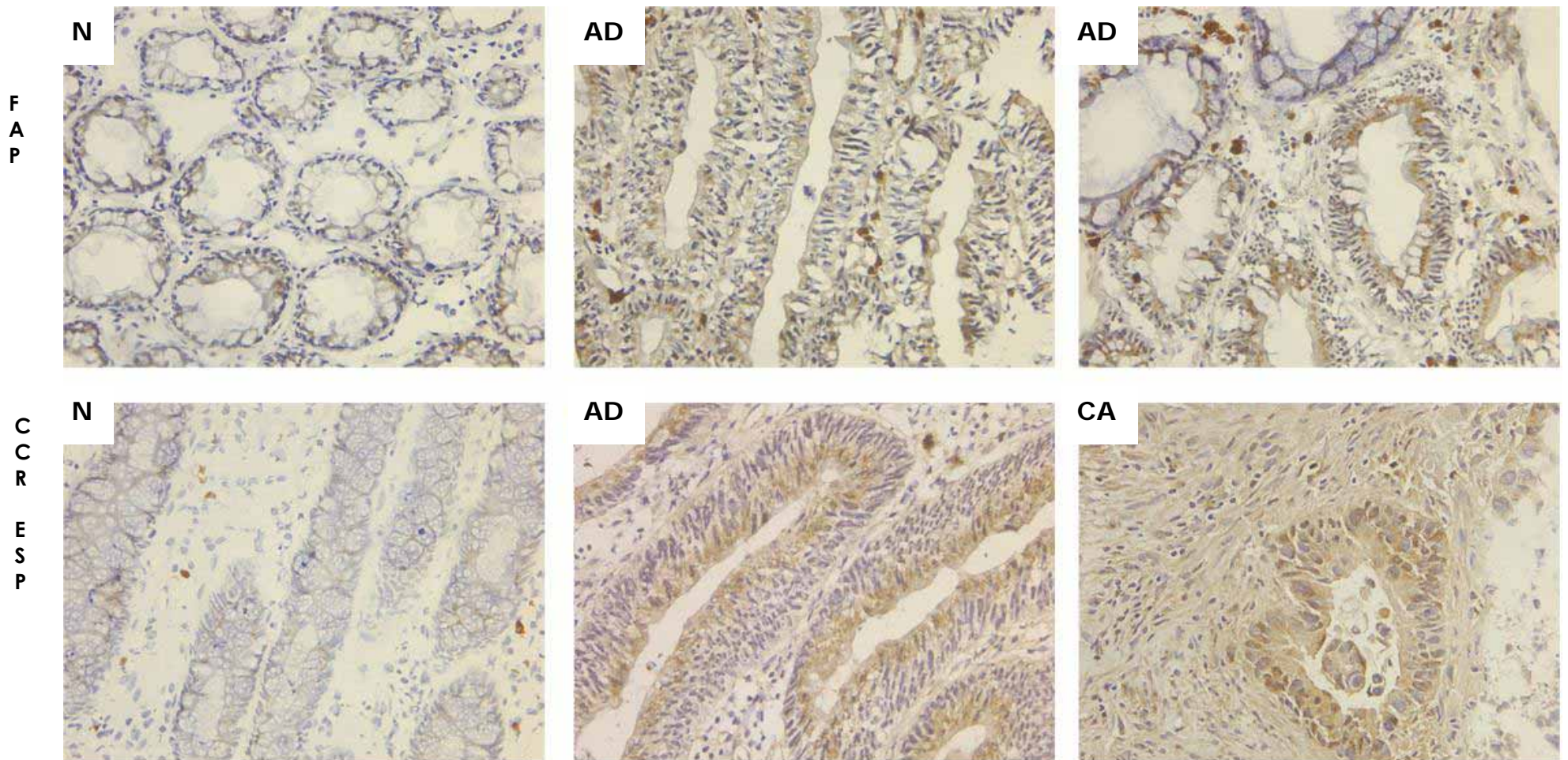


Figura 2.3. Immunohistoquímica de BUB1B de mostres representatives del conjunt analitzat de FAP i CCR esporàdic (CCR ESP). En tots els adenomes i carcinomes la sobreexpressió era evident, mentre que en les mostres de mucosa normal la tinció era molt més feble (augment 40X). N: mucosa normal; AD: adenoma; CA: carcinoma.

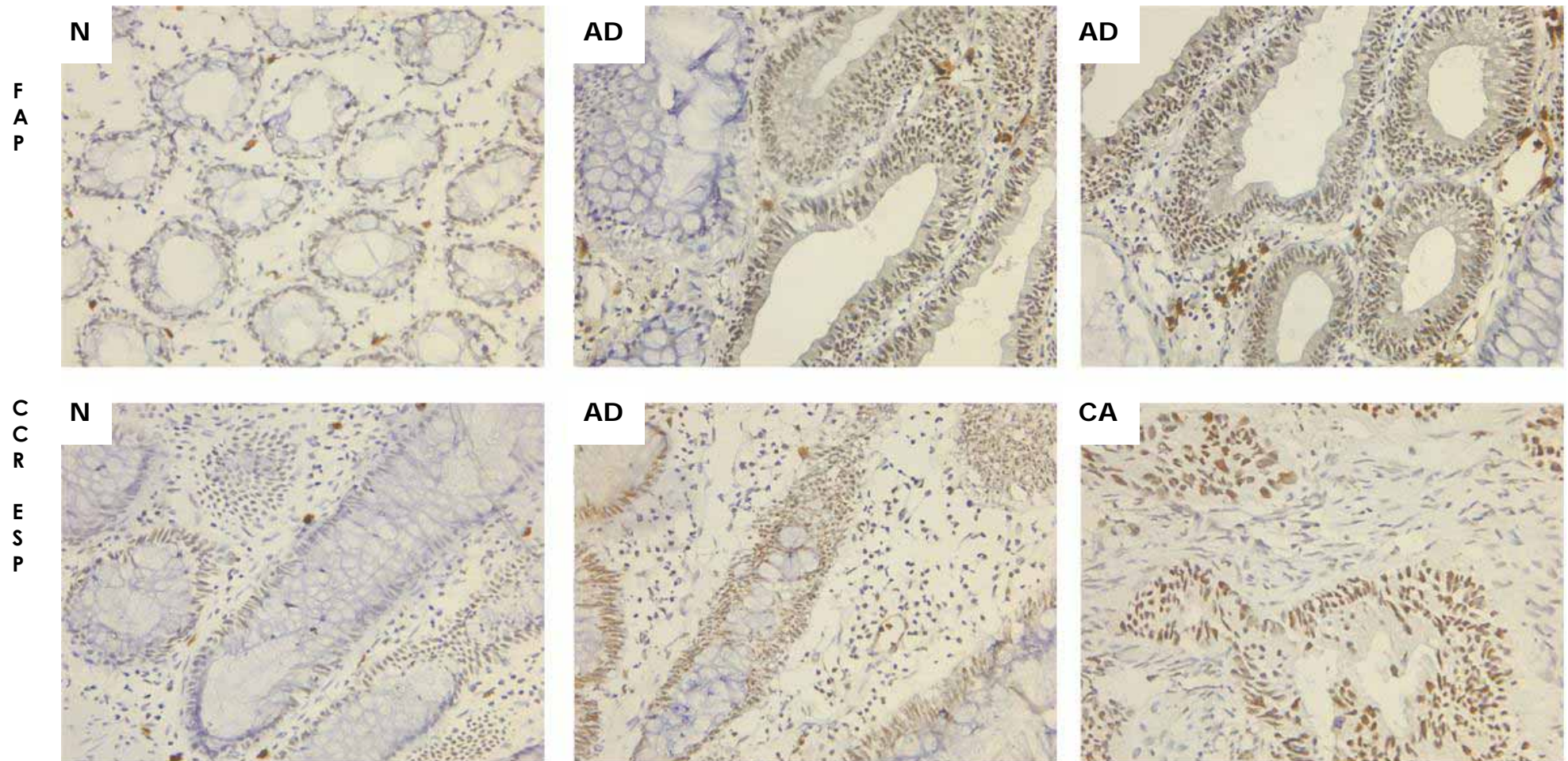


Figura 2.4. Immunohistoquímica de MAD2L1 de mostres representatives del conjunt analitzat de FAP i CCR esporàdic (CCR ESP). MAD2L1 mostrava una localització nuclear en adenomes i carcinomes. Es va observar una major tinció en teixits tumorals que en la mucosa normal. Les regions de teixit d'estroma no mostraven tinció, mentre que es veia una tinció inespecífica dels macròfags en tots els teixits. N: mucosa normal; AD: adenoma; CA: carcinoma.

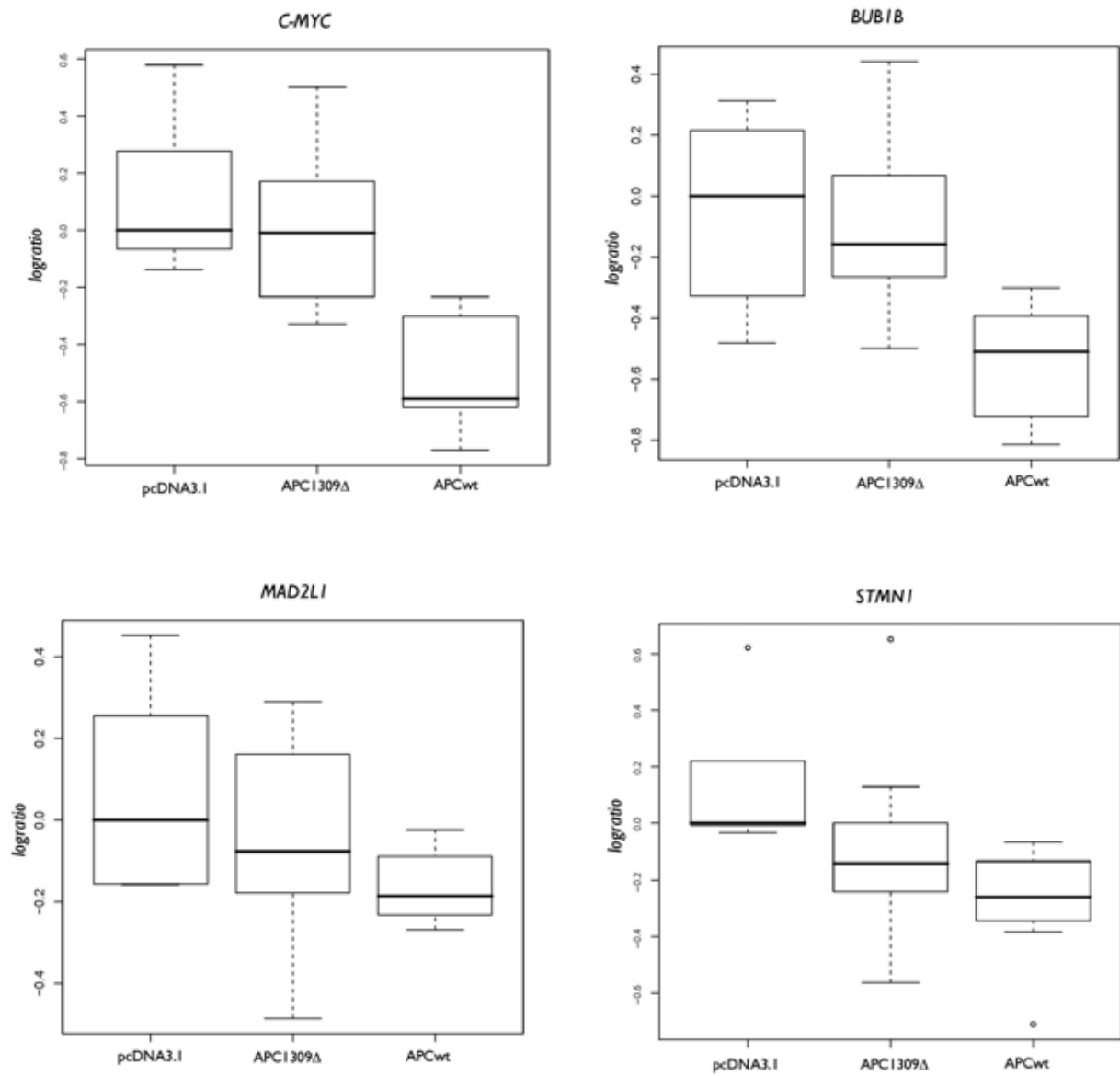
2. Canvis en la transcripció de *BUB1B*, *MAD2L1* i *STMN1* per la modulació de l'expressió d'APC

Amb l'objectiu d'investigar la relació entre la inactivació d'APC i l'increment d'expressió d'aquestes proteïnes lligades a la mitosi, es va avaluar si els seus nivells variaven quan es transfectaven proteïnes d'APC wt en la línia de CCR SW480 que té el gen totalment inactivat. La transfecció d'APC funcional se sap que restaura la capacitat cel·lular de minvar els nivells de β -catenina i, d'aquesta manera, reduir la senyalització de β -catenina/TCF4 que es troba constitutivament activada en aquestes cèl·lules (Morin *et al.*, 1997, Munemitsu *et al.*, 1995). La davallada en els nivells d'expressió de *C-MYC*, la diana de la via per excel·lència, es va fer servir com a control de l'eficiència de la introducció d'APC wt. La transfecció d'APC wt en cèl·lules SW480 va provocar la reducció de l'expressió de *BUB1B*, *MAD2L1* i *STMN1*, amb diferències estadísticament significatives per a *BUB1B* i *STMN1* (figura 2.5). A més, es va transfectar en aquesta mateixa línia el plasmidi que conté APC mutat. Els canvis observats amb el plasmidi d'APC mutat (APC1309 Δ) no són significatius com caldria esperar, ja que les cèl·lules SW480 ja tenen la proteïna inactivada (figura 2.5).

A més, es va portar a terme un experiment de transfecció d'siRNA en cèl·lules 293 les quals expressen APC wt. Es va obtenir una reducció del 50-65% de l'expressió endògena d'APC. Aquesta reducció en els nivells d'APC wt va tenir com a conseqüència un increment en l'expressió de *C-MYC* d'1,5 vegades, fet que demostrava una reducció en la senyalització a través de la β -catenina. De la mateixa manera, es va poder constatar un increment modest, però consistent, dels nivells d'expressió d'*MAD2L1* (1,22 vegades) i *STMN1* (1,17 vegades). Per *BUB1B* es varen detectar, però, resultats heterogenis que no varen permetre treure conclusions sobre el canvi d'expressió (figura 2.6). Per *MAD2L1* encara que el p-valor no sigui significatiu, la tendència de canvi és la mateixa que per *STMN1* i *C-MYC*.

Aquests dos experiments de transfecció, de formes wt i truncades d'APC i d'RNA d'interferència contra APC, representen dues maneres diferents de modular els nivells d'APC en dos tipus cel·lulars oposats. Per una banda les cèl·lules de càncer de còlon SW480 tenen la proteïna inactivada i per l'altra, les cèl·lules normals 293 tenen APC wt. Els canvis que s'observen amb els dos sistemes demostren canvis de *BUB1B*, *MAD2L1* i *STMN1* en el mateix sentit. D'aquesta manera podem concloure que el canvi

en el nivell d'APC funcional condiona el nivell d'expressió d'aquests gens, encara que el mecanisme a través del qual actua aquest canvi d'expressió és desconegut.



	APC1309Δ vs. pcDNA		APCwt vs. pcDNA	
	foldchange	p-valor	foldchange	p-valor
<i>C-MYC</i>	1,15	0,2	1,63	0,0002*
<i>BUB1B</i>	1,02	0,87	1,39	0,0125*
<i>MAD2L1</i>	1,1	0,29	1,19	0,069
<i>STMN1</i>	1,21	0,14	1,39	0,018*

* p-valor significatiu (<0,05)

continuació figura 2.5

Figura 2.5. Nivells relatius d'expressió de *C-MYC*, *BUB1B*, *MAD2L1* i *STMN1* en la línia SW480 transfectada. Els *boxplots* representen el canvi de la proteïna respecte a la situació basal (pcDNA 3.1) i la taula mostra els valors de *foldchange* (valor de canvi) i els p-valors pels 4 gens analitzats comparant APC 1309Δ i WT APC respecte a pcDNA 3.1.

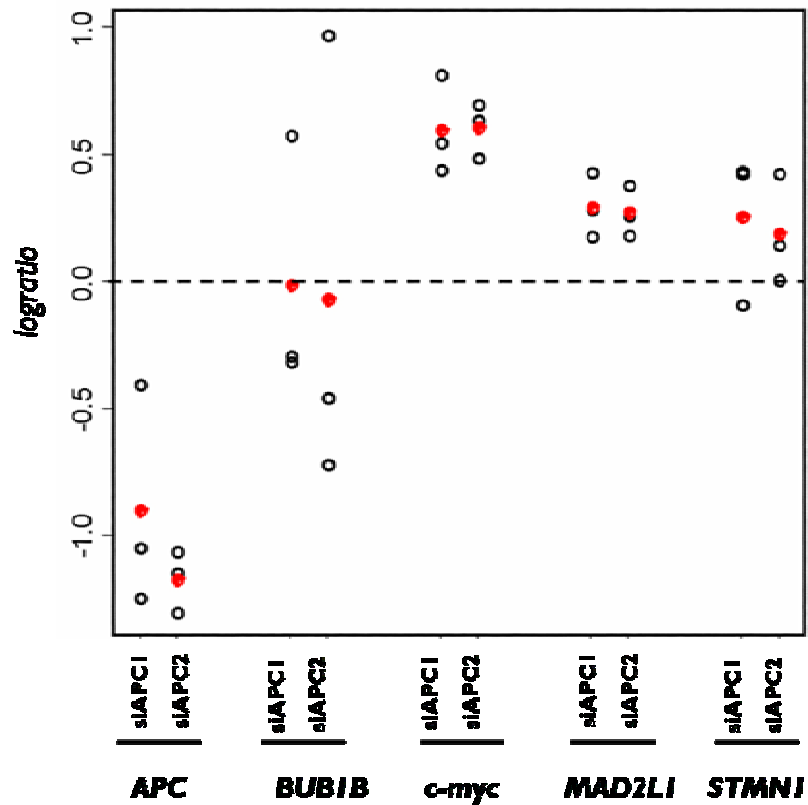


Figura 2.6. Nivells relatius d'expressió de *C-MYC*, *BUB1B*, *MAD2L1* i *STMN1* en la línia 293 transfectada amb dos RNA d'interferència contra APC (siRNA1 i siRNA2). Els punts buits representen els valors de *logratio* per a les tres rèpliques de l'experiment. Els punts vermells indiquen el valor mitjà de les tres rèpliques de cada experiment independent.

3. Estudi de coregulació de gens de la via de Wnt amb els gens *BUB1B*,

MAD2L1 i *STMN1*

En els apartats anteriors d'aquest capítol, s'ha descrit que existeix una relació entre un entorn cel·lular on el gen *APC* es troba inactivat amb un augment en el nivell d'expressió d'aquests gens (*BUB1B*, *MAD2L1* i *STMN1*). Quan el gen *APC* s'inactiva els gens diana de la via de Wnt passen a ser transcripcionalment activats pels factors de transcripció de la família de TCF. Això va suggerir que aquests gens de mitosi i els gens de la via de Wnt podrien estar sota el mateix control transcripcional. Per estudiar aquesta relació *in silico* es varen fer servir les dades de transcriptoma de pacients de càncer colorectal que provenen d'un *data set* (conjunt de dades) d'una sèrie de 36 carcinomes i mucoses normals aparellades (vegeu el capítol 3). Dels gens disponibles en el *microarray*, es varen seleccionar els que formen part d'aquestes dues vies (taula 2.2).

Gen	Descripció	Categoria
<i>SFRP1</i>	<i>secreted frizzled-related protein 1</i>	Wnt
<i>APC2</i>	<i>adenomatosis polyposis coli 2</i>	Wnt
<i>WSP2</i>	<i>WNT1 inducible signaling pathway protein 2</i>	Wnt
<i>FRAT1</i>	<i>Proto-oncogene FRAT1</i>	Wnt
<i>MET</i>	<i>Hepatocyte growth factor receptor</i>	Wnt
<i>WNT7B</i>	<i>Wnt-7b protein</i>	Wnt
<i>THC2208085</i>	<i>Axin 2</i>	Wnt
<i>FSD1</i>	<i>fibronectin type 3 and SPRY domain containing 1</i>	Wnt
<i>LRP4</i>	<i>Low-density lipoprotein receptor-related protein 4</i>	Wnt
<i>VEGFB</i>	<i>Vascular endothelial growth factor B</i>	Wnt
<i>MMP3</i>	<i>Stromelysin-1</i>	Wnt
<i>VEGFC</i>	<i>Vascular endothelial growth factor C</i>	Wnt
<i>EPHB2</i>	<i>Ephrin type-B receptor 2</i>	Wnt
<i>LRP8</i>	<i>Low-density lipoprotein receptor-related protein 8</i>	Wnt
<i>POSTN</i>	<i>Periostin</i>	Wnt
<i>LRP10</i>	<i>Low-density lipoprotein receptor-related protein 10</i>	Wnt
<i>DLK1</i>	<i>Delta-like protein</i>	Wnt
<i>FSTL3</i>	<i>Follistatin-related protein 3</i>	Wnt
<i>FN1</i>	<i>Fibronectin</i>	Wnt
<i>FSTL1</i>	<i>Follistatin-related protein 1</i>	Wnt
<i>LEF1</i>	<i>Lymphoid enhancer binding factor 1</i>	Wnt
<i>PPARD</i>	<i>Peroxisome proliferator activated receptor delta</i>	Wnt
<i>SFRP4</i>	<i>secreted frizzled-related protein 4</i>	Wnt

<i>EGFR</i>	<i>Epidermal growth factor receptor</i>	Wnt
<i>MYC</i>	<i>Myc proto-oncogene protein</i>	Wnt
<i>CD44</i>	<i>CD44 antigen</i>	Wnt
<i>GSK3B</i>	<i>Glycogen synthase kinase-3 beta</i>	Wnt
<i>GCG</i>	<i>Glucagon</i>	Wnt
<i>SOX9</i>	<i>Transcription factor SOX-9</i>	Wnt
<i>LRP2</i>	<i>Low-density lipoprotein receptor-related protein 2</i>	Wnt
<i>CTNNB1</i>	<i>Beta-catenin</i>	Wnt
<i>TCF1</i>	<i>Hepatocyte nuclear factor 1-alpha</i>	Wnt
<i>BTRC</i>	<i>F-box/WD-repeat protein 1A</i>	Wnt
<i>LRP1B</i>	<i>Low-density lipoprotein receptor-related protein 1B</i>	Wnt
<i>DVL2</i>	<i>Segment polarity protein dishevelled homolog DVL-2</i>	Wnt
<i>CDX1</i>	<i>Homeobox protein CDX-1</i>	Wnt
<i>PMF1</i>	<i>Osteocalcin</i>	Wnt
<i>APC</i>	<i>Adenomatous polyposis coli protein</i>	Wnt
<i>LRP16</i>	<i>Protein LRP16</i>	Wnt
<i>FGF18</i>	<i>Fibroblast growth factor-18</i>	Wnt
<i>EFNB2</i>	<i>Ephrin-B2</i>	Wnt
<i>LRP6</i>	<i>Low-density lipoprotein receptor-related protein 6</i>	Wnt
<i>CDH1</i>	<i>Epithelial-cadherin</i>	Wnt
<i>MITF</i>	<i>Microphthalmia-associated transcription factor</i>	Wnt
<i>VEGF</i>	<i>Vascular endothelial growth factor A</i>	Wnt
<i>THC2042689</i>	<i>c-myc binding protein</i>	Wnt
<i>BX102467</i>	<i>c-myc binding protein</i>	Wnt
<i>THC2043443</i>	<i>c-myc binding protein</i>	Wnt
<i>PFDN5</i>	<i>Prefoldin subunit 5</i>	Wnt
<i>THC2201082</i>	<i>Prefoldin subunit 5</i>	Wnt
<i>AV691423</i>	<i>c-myc binding protein</i>	Wnt
<i>BF830132</i>	<i>c-myc binding protein</i>	Wnt
<i>BUB1</i>	<i>Mitotic checkpoint serine/threonine-protein kinase BUB1</i>	Mitosi
<i>STMN2</i>	<i>Stathmin 2</i>	Mitosi
<i>BUB1B</i>	<i>Mitotic checkpoint serine/threonine-protein kinase BUB1 beta</i>	Mitosi
<i>STMN3</i>	<i>Stathmin 3</i>	Mitosi
<i>STMN1</i>	<i>Stathmin</i>	Mitosi
<i>THC2017722</i>	<i>Budding uninhibited by benzimidazoles 3 homolog</i>	Mitosi
<i>MAD1L1</i>	<i>Mitotic spindle assembly checkpoint protein MAD1</i>	Mitosi

Taula 2.2. Gens del *microarray* seleccionats que pertanyen a les categories Wnt i Mitosi.

Amb la mateixa anàlisi estadística que l'aplicada per a la totalitat dels gens descrita en el capítol 3 (vegeu el capítol 3 per a l'explicació del disseny experimental de les hibridacions), s'ha comprovat que els gens de la via de Wnt i els gens candidats (els

quals participen en la regulació de la mitosi) tenen un canvi molt petit en el seu nivell d'expressió per les comparacions del teixit normal amb els teixits tumorals (figura 2.7). Atès que el gen *MAD2L1* no està representat en el *microarray*, es va fer l'estudi de correlació dels gens *BUB1B* i *STMN1* amb els gens de la via de Wnt.

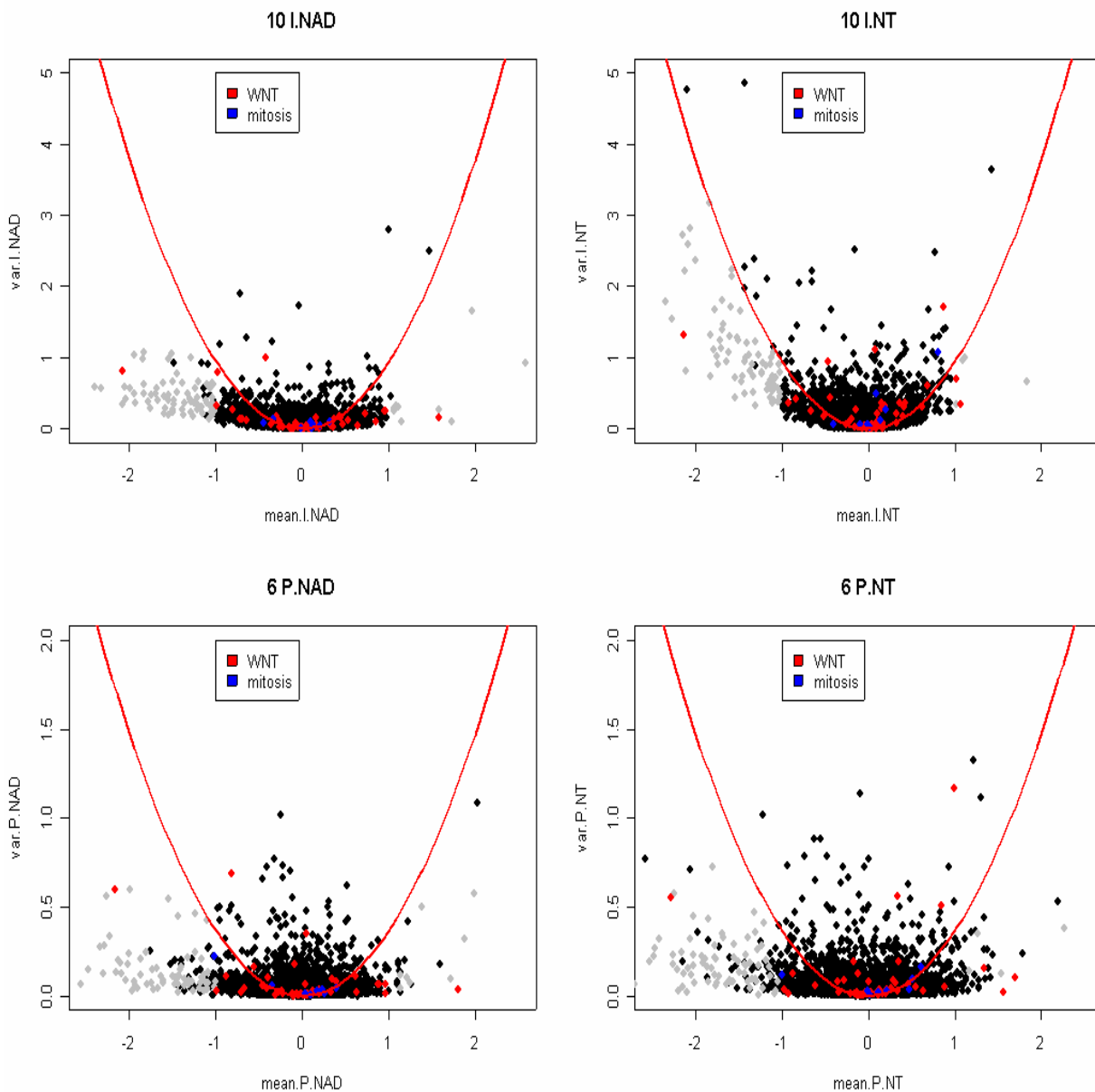


Figura 2.7. Representació de tots els gens del *microarray*. S'han destacat en vermell els gens de la via de Wnt i en blau els gens de mitosi per a les comparacions individuals (fila superior) i les comparacions de *pools* (fila inferior) i adenoma *vs.* normal (columna de l'esquerra) i carcinoma *vs.* normal (columna de la dreta). Els gens que queden per sota de la línia vermella són els gens significatius amb el t-test.

S'ha obtingut un clúster per *hierarchical clusternig* on sembla raonable tallar l'arbre en cinc clústers (figura 2.9) de manera que els gens candidats (*MAD2L1*, *BUB1B* i *STMN1*) queden en els clúster 1: *BUB1B* (sonda 5) i *MAD2L1* (sondes 1 i 2) i clúster 2 : *BUB1B* (sondes 1 a 4) i *STMN1*(sondes 1 i 2) (figura 2.9). Cal destacar que en aquests dos clústers és també on s'han distribuït totes les sondes que representen el gen *C-MYC*, que és el gen diana de la via de Wnt per excel·lència. També en el clúster 1 s'hi troben alguns dels gens diana de Wnt més ben estudiats: *CICLINA D1*, *CD44* i *C-JUN*.

Per tal de realitzar una altra anàlisi de coregulació més àmplia amb aquestes mateixes dades, es va calcular el coeficient de correlació de les sondes candidates amb: 1) totes les sondes de l'*array*; 2) només les sondes de la via de Wnt. Es varen obtenir dues llistes de coeficients de correlació per cada una de les comparacions. Es va aplicar el test de Kolmogorov-Smirnov per tal d'avaluar la diferència en les distribucions de les dues llistes de coeficients de correlació. Set de les nou sondes varen donar p-valors significatius, fet que indica que la correlació de les sondes candidates no segueix la mateixa distribució per la correlació amb totes les sondes que per la correlació amb les sondes de la via de Wnt. El resultat doncs, dóna una evidència que els nivells d'expressió dels gens candidats poden estar correlacionats amb els dels gens de la via de Wnt.

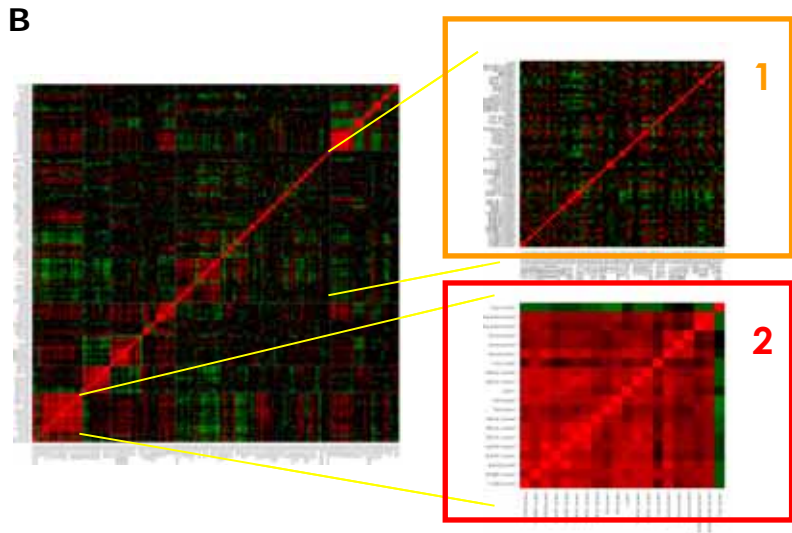
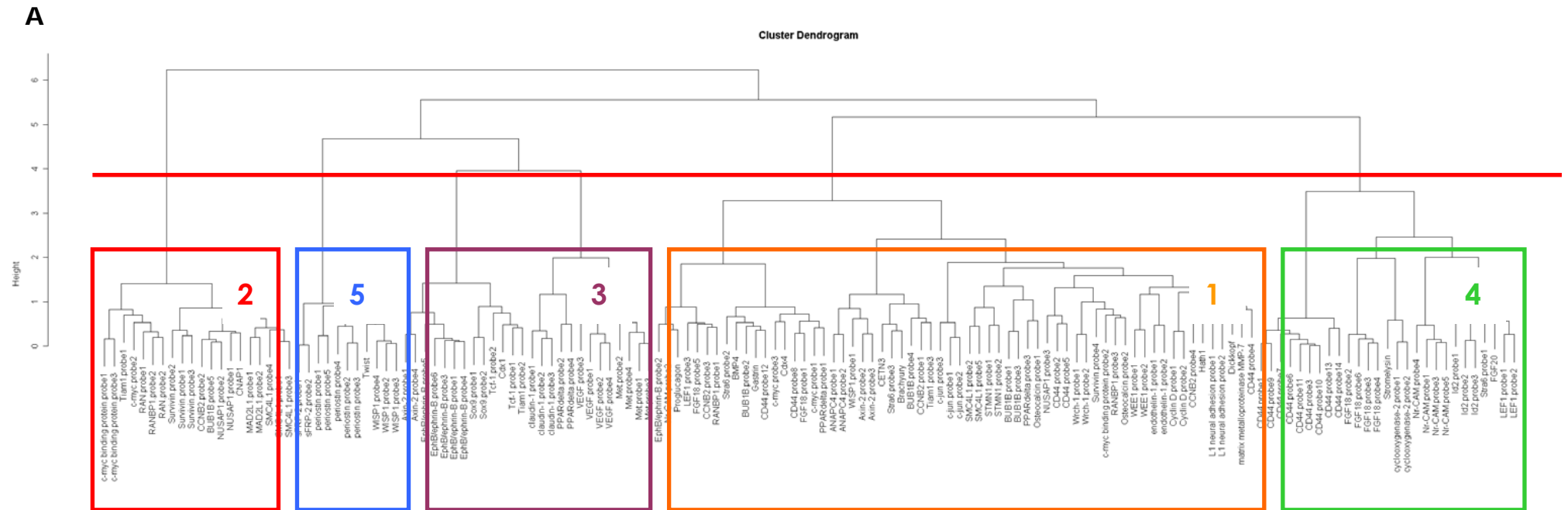


Figura 2.9. Anàlisi de coregulació en la sèrie de dades de Bittner. A) Cinc clústers de gens agrupats segons el nivell de correlació del seus nivells d'expressió. Les sondes que representen els gens *BUB1B*, *MAD2L1* i *STMN1* queden repartides entre els clústers 1 i 2 on també es troben alguns dels gens diàna de la via de Wnt més representatius. B) Detall dels clústers 1 i 2 de la matriu de correlacions dels 5 clústers.

Capítol 3

Perfils d'expressió del càncer
colorectal

Els estudis dels perfils d'expressió utilitzant els *microarrays* tenen el potencial de millorar el diagnòstic i el tractament del càncer. Fins ara s'han identificat molts gens que es troben desregulats en el càncer colorectal quan es comparen amb el teixit normal. L'àmplia aplicació d'aquesta tecnologia fa que sigui una eina amb molta prospecció en la recerca translacional.

1. Detecció de gens diferencialment expressats entre els teixits de mucosa normal, adenoma i carcinoma de càncer colorectal esporàdic. Anàlisi global.

S'han analitzat els perfils d'expressió de 108 biòpsies aparellades de mucosa normal, adenoma i carcinoma de 36 pacients de càncer colorectal esporàdic mitjançant *microarrays* d'oligonucleòtids que contenen 4.528 gens. Les mostres es varen distribuir en sis grups d'acord amb el seu estadi i la localització del carcinoma (vegeu la figura M.6 dels materials i mètodes). Es varen agrupar els RNA (*pools*) de cada un dels teixits (normal, adenoma o carcinoma) per grup i es varen fer les hibridacions dels *pools* de cada grup: adenoma *vs.* mucosa normal i carcinoma *vs.* mucosa normal. A part de les hibridacions de les mostres agrupades, es varen seleccionar deu pacients representatius dels sis grups i el seu RNA es va hibridar de manera individual. En total es disposava de 32 hibridacions.

Les anàlisis es varen fer en paral·lel per als dos tipus de dades que hi ha: provinents d'hibridacions de *pools* i provinents d'hibridacions individuals. Les dades que s'obtenen en els dos casos estan altament correlacionades, fet que indica que els canvis que s'observen no són dependents d'individu, sinó que són alteracions recurrents en la progressió del CCR. Per tant, l'anàlisi s'ha fet amb totes les dades de manera conjunta (hibridacions de *pools* juntament amb hibridacions individuals).

Per identificar els gens diferencialment expressats (GDE) entre els tres tipus de teixits, s'ha aplicat un t-test i s'han seleccionat el gens d'acord amb el següent criteri: una *logratio* major d'1 o menor de -1 i un p-valor inferior a 0,01. Aquest criteri es va decidir després de realitzar una correcció per contrastos múltiples basada en el FDR (*false discovery rate*), la qual indica la taxa de falsos positius entre els gens seleccionats com a significatius. Aquest mètode assumeix que hi ha una uniformitat en la distribució dels p-valors sota la hipòtesi nul·la que no hi ha cap gen diferencialment expressat.

Utilitzant la distribució empírica dels p-valors (histograma) el mètode busca el percentatge de gens vertaderament no diferencialment expressats per aproximar el FDR, és a dir, el percentatge de falsos positius esperats en el conjunt de gens declarats com a significatius. Es va obtenir un resultat que inicialment va semblar paradoxal, ja que el llindar del p-valor que s'obtenia per a un FDR del 5% o de l'1% era superior a 0,05 i 0,01 respectivament. Això vol dir que els gens que se seleccionen com a GDE amb un FDR del 5 o de l'1% tenen un p-valor superior a 0,05 i en aquest cas se seleccionen molt gens i gairebé tots són vertaders gens diferencialment expressats. Aquest fet s'ha descrit alguna vegada i es dona quan hi ha molts gens significatius i, en conseqüència, l'histograma de p-valors mostra una concentració de valors al voltant de 0. Es pot observar que hi ha 3.414 sondes que representen GDE amb un FDR del 5% (figura 3.1). Això equival a agafar aquests gens amb un p-valor del t-test inferior a 0,22 (molt major que 0,05). Per tant, es va optar per seleccionar com a significatives les sondes amb p-valors inferiors a 0,05 que garanteixen de sobra un FDR inferior a 5% i es va emprar com a segon filtre el llindar de la *logratio*. Això permet dir que els gens que se seleccionen amb aquest doble criteri, a la llum dels resultats de el FDR, són gens altament fiables. Amb aquest criteri s'han identificat 78 GDE entre mucosa normal i adenoma i 76 GDE entre mucosa normal i carcinoma (figura 3.2). La llista d'aquests gens no es pot mostrar, ja que en aquests moments està protegit perquè es troba en el procés d'avaluació de patents.

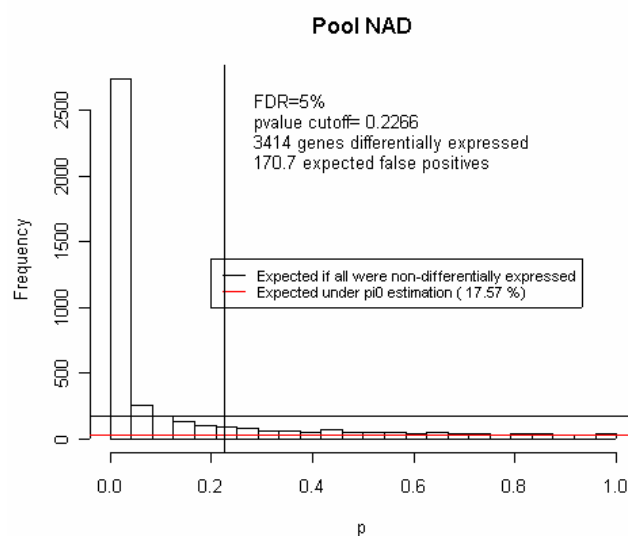


Figura 3.1. Exemple d'histograma que representa els p-valors dels gens diferencialment expressats entre adenoma i teixit normal. La majoria de valors es concentren al 0. Si s'aplica un FDR del 5%, només es detecten 170 falsos positius entre els 3.414 GDE.

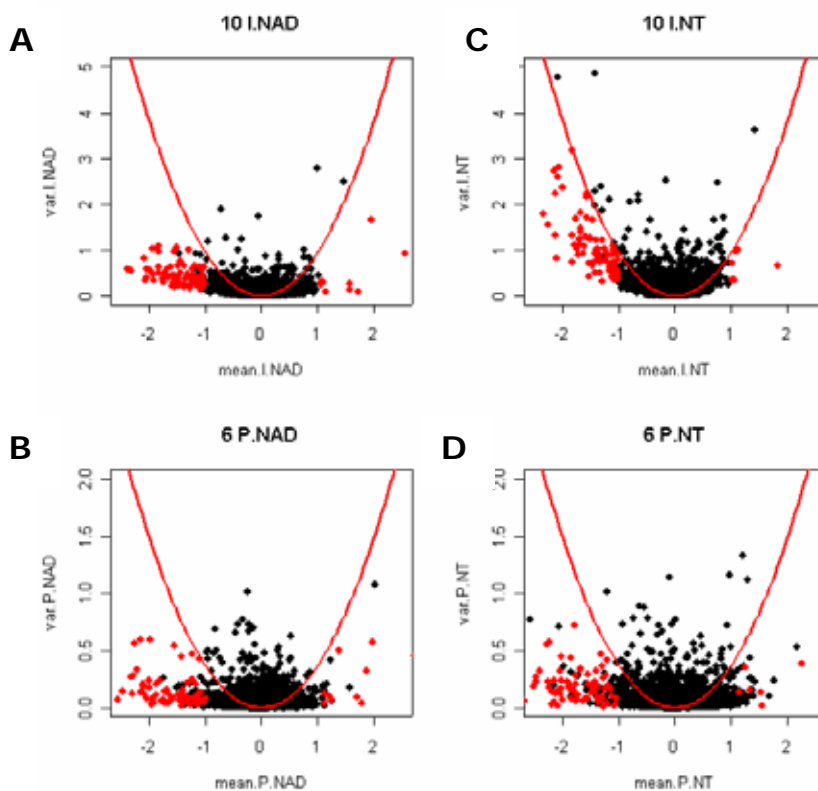


Figura 3.2. Representació de tots els gens del *microarray* segons la seva *logratio*. La línia vermella indica els gens significatius després d'aplicar el t-test i un FDR del 5%. Com que s'ha aplicat el criteri de seleccionar els gens que tenen una *logratio* major d'1 i menor de -1, els punts marcats en vermell representen els gens diferencialment expressats. A la primera columna hi ha les dades per a la comparació adenoma *vs.* mucosa normal (NAD) (A: dades d'hibridacions individuals; B: dades de *pools*) i a la segona columna la comparació carcinoma *vs.* mucosa normal (NT) (C: dades d'hibridacions individuals; D: dades de *pools*).

Convé destacar que s'ha constatat una gran superposició (*overlap*) de GDE dels dos tipus d'hibridacions i, per aquesta raó, es varen voler estudiar les diferències entre adenomes i carcinomes amb més detall. Atès que no s'havien fet cohibridacions d'aquests dos tipus de teixits, les mostres d'adenoma i carcinoma es varen comparar de manera indirecta. Només s'han detectat 22 gens diferencialment expressats (taula 3.1) entre aquests dos teixits. En ser la comparació indirecta, la variabilitat és major i es perd poder estadístic per trobar gens diferencialment expressats. Hi ha, per tant, una alta correlació entre els gens desregulats en els estadis d'adenoma i de carcinoma (figura 3.3).

Gen	Descripció
<i>IGFBP2</i>	<i>Insulin-like growth factor binding protein 2</i>
X01	Pendent revisió de patents
X02	Pendent revisió de patents
<i>STMN3</i>	<i>Stathmin 3</i>
<i>POSTN</i>	<i>Periostin</i>
X03	Pendent revisió de patents
<i>FN1</i>	<i>Fibronectin</i>
X04	Pendent revisió de patents
<i>FCGBP</i>	<i>Fc fragment of IgG binding protein</i>
X05	Pendent revisió de patents
<i>COL1A2</i>	<i>Collagen alpha 2(I) chain</i>
<i>AREG</i>	<i>Amphiregulin</i>
X06	Pendent revisió de patents
X07	Pendent revisió de patents
X08	Pendent revisió de patents
<i>BCMP11</i>	<i>breast cancer membrane protein 11</i>
X09	Pendent revisió de patents
<i>SPARC</i>	<i>Secreted protein acidic and rich in cysteine</i>
<i>UBD</i>	<i>Gamma-aminobutyric acid type B receptor, subunit 1</i>
X10	Pendent revisió de patents
<i>ITLN1</i>	<i>Intelectin 1</i>
X11	Pendent revisió de patents

Taula 3.1. Gens diferencialment expressats entre adenoma i carcinoma en mostres de càncer colorectal esporàdic. En vermell els gens sobreexpressats en carcinomes respecte a adenomes i en verd els gens infraexpressats en carcinomes. Els gens codificats com a X01-X11 no mostren el seu nom real, ja que estan pendents de redacció/avaluació d'aplicacions de patent.

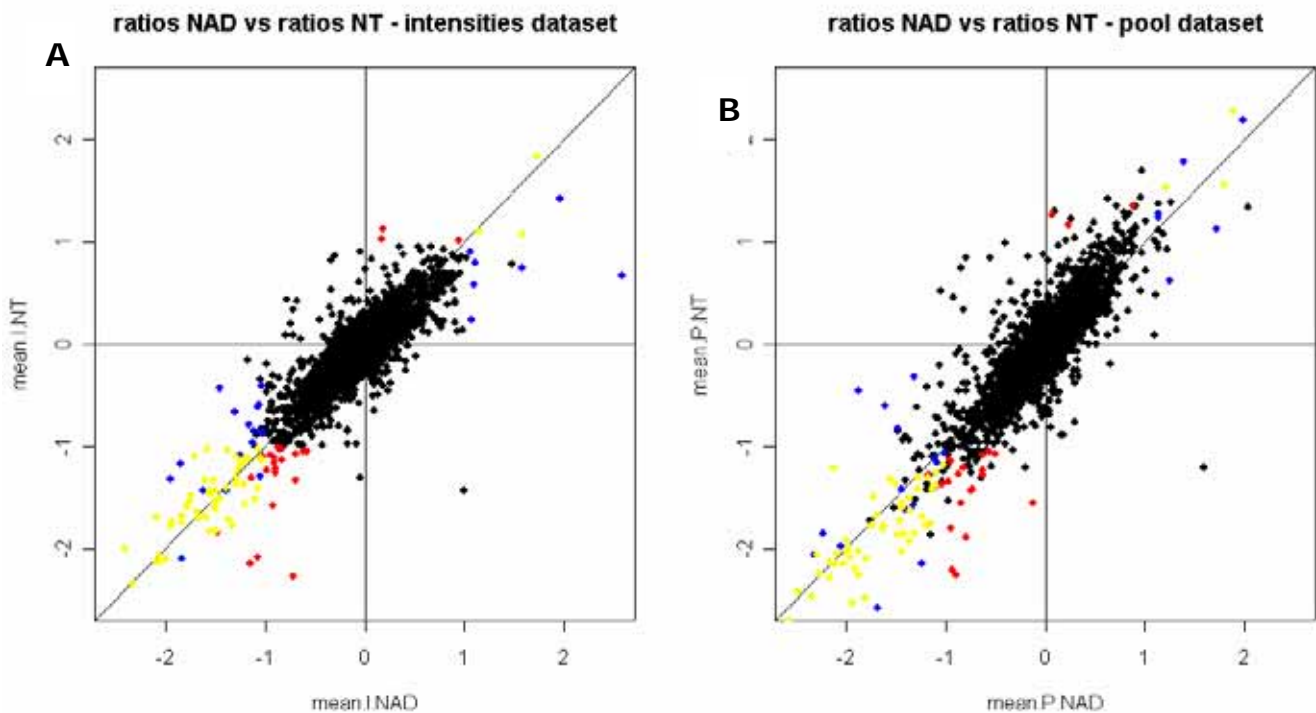


Figura 3.3. Comparació de dades dels teixits d'adenoma i carcinoma per a l'anàlisi de dades individuals (A) i per a l'anàlisi de dades de *pools* (B). Gens considerats significatius per la comparació adenoma *vs.* mucosa normal (blaus); els significatius per la comparació carcinoma *vs.* mucosa normal (vermells); i aquells gens significatius per les dues comparacions (grocs).

Per resumir els resultats obtinguts de la comparació dels tres teixits, s'ha construït un diagrama de Venn on apareix el nombre de gens diferencialment expressats per cada una de les comparacions (figura 3.4).

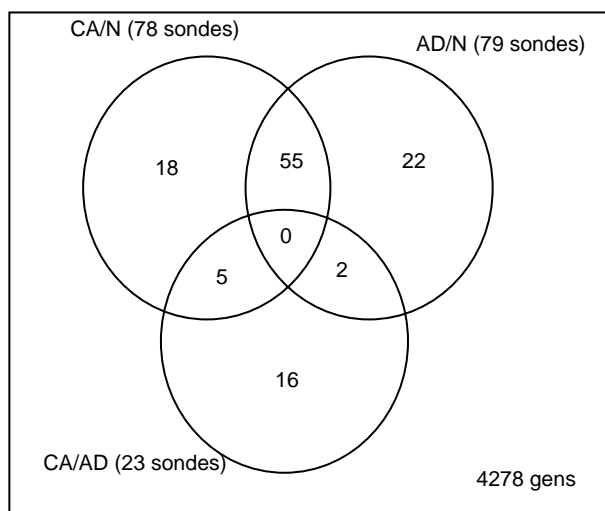


Figura 3.4. Diagrama de Venn per a les tres comparacions: adenoma *vs.* mucosa normal (AD/N), carcinoma *vs.* mucosa normal (CA/N) i carcinoma *vs.* adenoma (CA/AD). Aquí es representen les sondes diferencialment expressades per cada una de les comparacions. Un gen pot estar representat per més d'una sonda.

Com a següent pas es va analitzar si hi havia diferents patrons de comportament dels gens al llarg de la progressió adenoma-carcinoma. Així es varen definir cinc grups de gens (figura 3.5):

- 1) Es manté el mateix nivell d'expressió entre la mucosa normal i l'adenoma i disminueix la seva expressió en el carcinoma (n=7).
- 2) Augmenta el nivell d'expressió entre la mucosa normal i l'adenoma i torna a disminuir en el carcinoma (n=1).
- 3) Disminueix el nivell d'expressió entre la mucosa normal i l'adenoma i torna a augmentar en el carcinoma (n=13).
- 4) Disminueix el nivell d'expressió a l'adenoma i es manté en el carcinoma. Un exemple d'aquest grup de gens és *TNFR* (n=90).
- 5) Augmenta el nivell d'expressió a l'adenoma i es manté en el carcinoma. Exemples d'aquest grup de gens serien *AXINA2* i *C-MYC* (n=12). És interessant destacar que s'obté aquesta mateixa tendència amb els experiments de PCRq sobre mostres de FAP i de CCR esporàdic exposats al capítol 1.

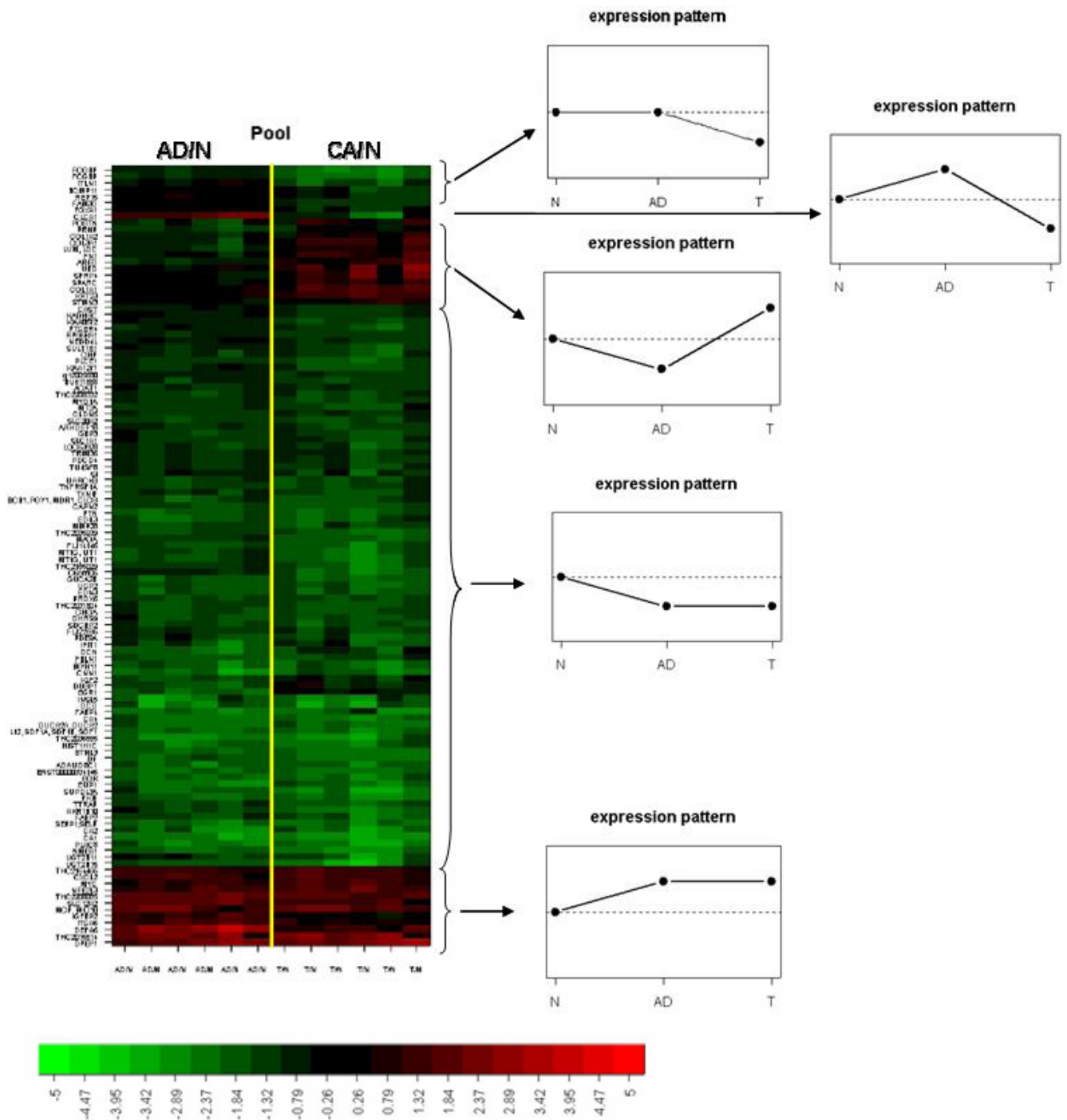


Figura 3.5. *Heatmap* per l'anàlisi de dades de *pools* on es poden veure els gens diferencialment expressats per les comparacions adenoma vs. mucosa normal (AD/N) i carcinoma vs. mucosa normal (CA/N). Els colors depenen del nivell d'expressió. Així gens verds són els que s'infraexpressen i els vermells són els que es sobreexpressen.

2. Anàlisi de perfils d'expressió diferencials segons la localització i l'estadi

Un cop identificats els GDE i els patrons d'expressió seqüencials adenoma-carcinoma es va aplicar una altra anàlisi estadística per tal d'investigar si els perfils d'expressió correlacionaven amb variables clíniques com l'estadi i la localització (figura M.6 de materials i mètodes). Es va aplicar el test de la Xi-quadrat per estudiar la distribució de les variables i el coeficient V de Cramer per explorar si les variables estan relacionades. Els resultats mostren que les variables es troben homogèniament distribuïdes ($p\text{-valor} \gg 0,05$) i que no hi ha relació entre elles ($V \ll 0,3$), cosa que permet fer les anàlisis de grups. Per tant, es varen fer comparacions dels subgrups de mostres amb totes les combinacions de les variables de dues en dues: localització-estadi; tipus de tumor (normal vs. adenoma o normal vs. carcinoma)- localització; estadi- tipus de tumor.

Comparació dels diferents estadis amb la localització del carcinoma

En la primera comparació (taula 3.2), localització- estadi, es diferencia molt bé entre estadis I-II (primerenc) i III-IV (avançat) quan el carcinoma es troba localitzat al recte i al còlon esquerre amb 157 i 198 gens diferencialment expressats (GDE) respectivament amb un baix grau de superposició (*overlap*). Cal destacar que aquesta diferència no s'observa en els carcinomes del còlon dret i això pot indicar que no necessiten acumular més alteracions mentre progressen.

Comparació dels estadis estratificant per localització

I-II	III-IV	GDE	overlap		
recte	recte	157	x	5	
còlon esquerre	còlon esquerre	198	x		x 1
còlon dret	còlon dret	5			x

Comparació localització-estadi primerenc (I-II)

1-II		GDE	overlap			
recte	còlon esquerre	544	x	229	x	x
recte	còlon dret	458	x		x 77	x 3
còlon esquerre	còlon dret	404		x 66	x	x

Comparació localització-estadi avançat (III-IV)

III-IV		GDE	overlap		
recte	còlon esquerre	0			
recte	còlon dret	110	x		
còlon esquerre	còlon dret	560	x	71	

Taula 3.2. Gens diferencialment expressats per a la comparació localització- estadi.
GDE: gens diferencialment expressats.

Pel que fa a la localització, quan es posa atenció als estadis primerencs, es constata que hi ha diferències molt grans entre les diferents localitzacions: recte- còlon esquerre GDE= 544, recte- còlon dret GDE= 458 i còlon esquerre- còlon dret GDE= 404 amb un alt grau de superposició entre els gens que canvien a còlon esquerre i dret respecte a recte mentre que aquest grau de superposició és inferior entre els gens que canvien de còlon dret a esquerre o recte. Curiosament amb els estadis avançats no hi ha diferències en GDE entre recte i còlon esquerre mentre que continuen fent-se diferents del còlon dret. Veiem doncs que, pel que fa als grups de mostres seleccionats, són grups ben diferenciats i que existeixen diferències entre l'estadi i la localització.

L'anàlisi dels termes de *Gene Ontology* (GO) va identificar els gens implicats en diferents processos cel·lulars, entre els quals hi ha l'homeòstasi dels ions de coure com els més enriquits en la comparació dels estadis primerencs i avançats en tumors rectals (FDR ajustada p-valor= 0,00004) (figura 3.6). Es varen identificar alguns gens en aquesta categoria que podrien indicar una interacció amb factors ambientals que contribueixen a la susceptibilitat al càncer colorectal. Un altre terme identificat va ser el dels gens implicats en el pas del control d'inici de la mitosi en el cicle cel·lular. Dins aquest grup es varen identificar dos gens de control del cicle cel·lular, fet que reforça la idea que les alteracions en el *checkpoint* del cicle cel·lular són clau en la tumorigènesi colorectal.

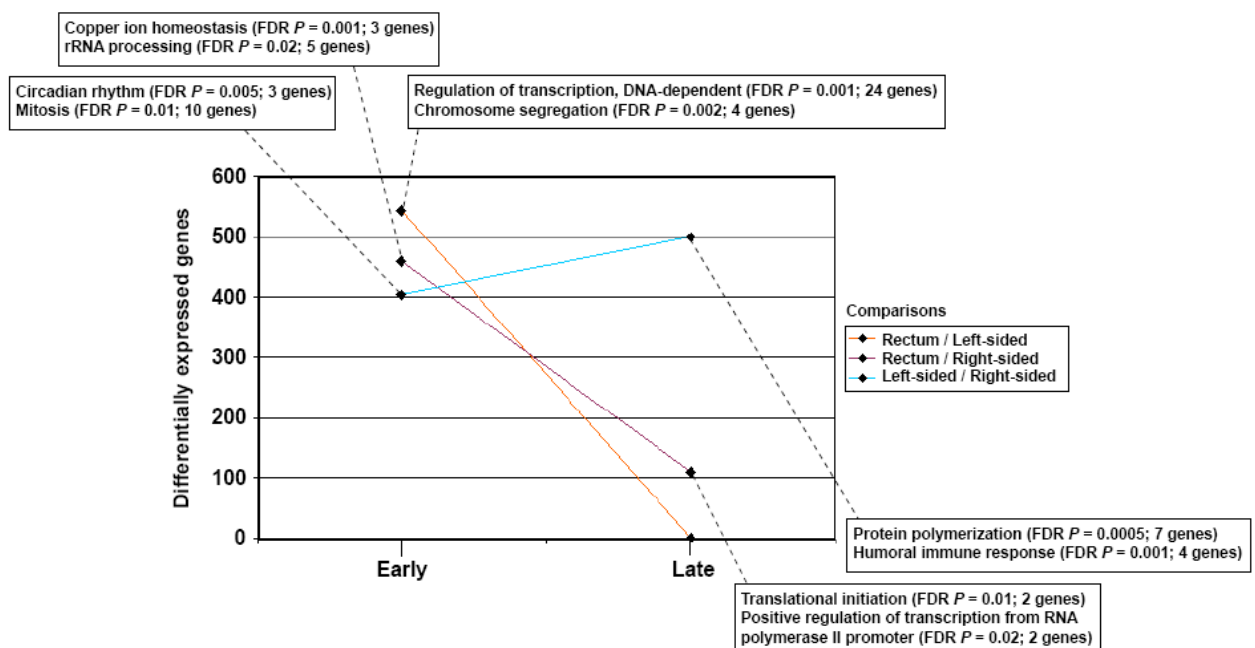


Figura 3.6. Anàlisi dels GO terms en la comparació dels estadis primerencs i tardans per a les diferents localitzacions del carcinoma.

Els canvis més petits en expressió observats entre els estadis primerencs i més avançats es donen al còlon dret on només cinc gens estan diferencialment expressats (taula 3.2). Quan es varen comparar les localitzacions es va observar que els tumors de còlon dret es comporten com una entitat totalment diferent. L'anàlisi dels gens diferencialment expressats mostra un enriquiment pel al grup de gens de la resposta immune (FDR ajustada p-valor 0,001). Aquesta observació és coherent amb el fet que aquests tumors acostumen a mostrar infiltració per cèl·lules inflamatòries.

Comparació dels adenomes i carcinomes amb la localització del carcinoma

A continuació es va voler explorar si hi havia diferències entre adenoma i carcinoma segons la seva localització (taula 3.3). Tal com s'ha esmentat abans, no hi ha grans diferències en el nombre de GDE entre tumors benignes i malignes (vegeu l'apartat 1 d'aquest capítol) que, de fet, es limiten al recte on hi ha 51 GDE entre adenomes i carcinomes. Respecte a les diferències per localització, ja des del principi, el recte es diferencia fortament del tumors de còlon dret (GDE= 430) i en carcinomes hi ha força diferència entre còlon dret i esquerre (GDE= 377). Tot plegat suggereix que les diferències entre adenomes i carcinomes són més evidents quan estan situats al recte.

Comparació per tipus de tumor estratificant per localització

N/AD		N/CA		GDE
recte	nad	recte	nca	51
còlon esquerre	nad	còlon esquerre	nca	2
còlon dret	nad	còlon dret	nca	0

Comparació dels adenomes estratificant per localització

N/AD		GDE	overlap	
recte	còlon esquerre	12	x	x
recte	còlon dret	430	x	1
còlon esquerre	còlon dret	1		x

Comparació dels carcinomes estratificant per localització

N/CA		GDE	overlap	
recte	còlon esquerre	1		x
recte	còlon dret	10	x	1
còlon esquerre	còlon dret	377	x	x

Taula 3.3. Gens diferencialment expressats segons el tipus de lesió: adenoma o carcinoma. GDE: gens diferencialment expressats. N/AD: hibridació normal *vs.* adenoma. N/CA: hibridació normal *vs.* carcinoma.

Per tal d'aprofundir una mica més en quins eren els gens implicats en la transició adenoma-carcinoma es va estudiar quin paper podria tenir l'activació constitutiva de la via WNT/ β -catenina, ja que aquesta via és clau en carcinogènesi colorectal (Clevers, 2006). Per això es va analitzar si la regió compresa entre -5.000 i +200 bp (regió on potencialment es localitza el promotor del gen) de cada un dels 53 gens diferencialment expressats entre adenomes i carcinomes (taula 3.3) per si contenien llocs d'unió de factors de transcripció de la via. Dels gens analitzats, n'hi ha 18 (33%) que contenen llocs d'unió potencials de factors de transcripció de la via, fet que suggereix que la senyalització activa de WNT/ β -catenina és important en aquesta fase de la progressió del càncer colorectal.

3. Validació dels resultats dels *microarrays* amb la tècnica de la PCRq

Amb l'objectiu de validar les dades obtingudes amb les anàlisis dels *microarrays*, es varen fer servir els resultats dels valors d'expressió dels gens *BUB1B*, *STMN1*, *sFRP1*, *APC* i *C-MYC* determinats per PCRq (vegeu els capítols 1 i 2), ja que les mostres de càncer colorectal esporàdic utilitzades per PCRq també varen ser hibridades sobre els *microarrays*.

Es va considerar interessant aplicar un t-test per avaluar la significació dels canvis observats amb les dues tecnologies. Les dades dels *microarrays* contenen valors per a dades individuals i per a dades de *pools*. S'ha vist que les dades dels *pools* són menys variables, i això dona un poder estadístic major per detectar diferències (taula 3.4).

S'ha aplicat el test de correlació de Pearson per veure quina és la correlació de les dades dels *microarrays* amb les de la PCRq (figura 3.7). El valor de correlació és moderadament elevat ($r= 0,31$) amb un p-valor de 0,005. El rang de valors per a les dues tècniques són diferents, ja que en els *microarrays* els valors són molt petits per l'efecte de compressió de les *logratios*. En els *microarrays* es poden perdre alguns canvis, ja que és una tècnica a gran escala, però per a canvis subtils no és tan sensible com una PCRq. Per determinar la correlació del sentit del canvi (sobreexpressió o infraexpressió) amb les dues tècniques, es va aplicar un test de correlació d'Spearman. El valor de correlació és llavors molt elevat ($r= 0,78$), la qual cosa indica que els valors d'expressió canvien en el mateix sentit per a les dues tècniques i això permet fer una validació de les dades dels *microarrays*.

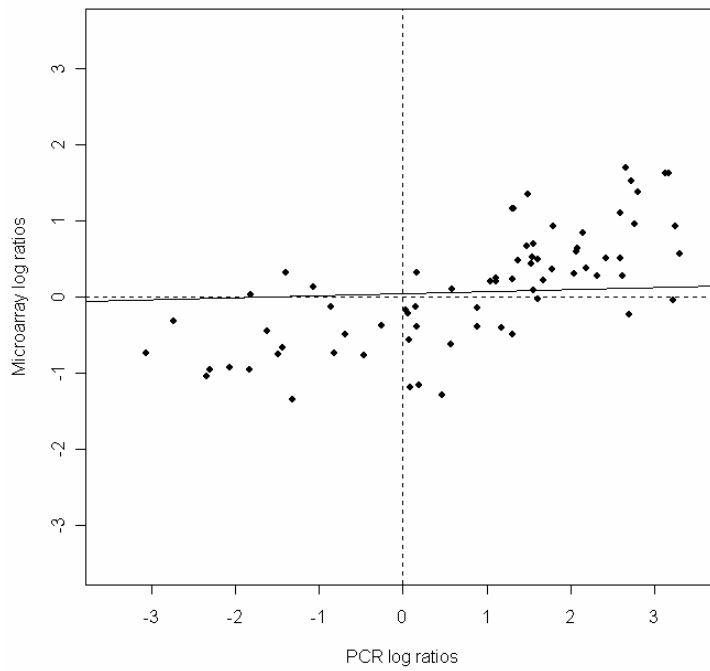


Figura 3.7. Comparació de les *logratios* obtingudes per PCRq i *microarrays* per als gens *BUB1B*, *STMN1*, *sFRP1*, *APC* i *C-MYC*. En l'eix de les Xs es representen les *logratios* obtingudes amb la PCRq i en l'eix de les Ys les obtingudes amb els *microarrays*.

	<i>sFRP1</i>		<i>BUB1B</i>		<i>STMN1</i>		<i>C-MYC</i>		<i>APC</i>	
	N vs. CA	N vs. AD	N vs. CA	N vs. AD	N vs. CA	N vs. AD	N vs. CA	N vs. AD	N vs. CA	N vs. AD
PCRq										
Mitjana de la <i>logratio</i>	-3.824	-12.355	1.444	1.253	0.700	0.757	2.267	1.952	-0.015	-0.169
Desviació estàndard	3.564	22.685	1.708	1.335	1.423	1.026	1.766	0.651	0.992	0.753
<i>t-score</i>	-3.392	-1.722	2.672	2.967	1.555	2.334	4.060	8.997	-0.047	-0.709
p-valor	0.008*	0.119	0.026*	0.016*	0.154	0.044*	0.003*	0.000*	0.963	0.496
microarray individual										
Mitjana de la <i>logratio</i>	-0.6604	-0.6947	0.2052	0.1014	0.0939	0.3189	10.137	0.9435	-0.5145	-0.7195
Desviació estàndard	0.5024	0.3736	0.5104	0.2958	0.7049	0.3228	0.8380	0.4871	0.4030	0.3565
<i>t-score</i>	-41.570	-58.796	12.714	10.839	0.4212	31.245	38.254	61.255	-40.372	-63.822
p-valor	0.0025*	0.0002*	0.2355	0.3066	0.6835	0.0122*	0.0041*	0.0002*	0.0029*	0.0001*
microarray pool										
Mitjana de la <i>logratio</i>	-0.8790	-0.8818	0.5087	0.2272	0.4760	0.3909	13.398	0.8807	-0.6160	-0.6736
Desviació estàndard	0.3560	0.3273	0.3095	0.1925	0.1901	0.1995	0.3981	0.2528	0.2402	0.1966
<i>t-score</i>	-60.480	-65.996	40.263	28.922	61.330	47.981	82.435	85.346	-62.814	-83.939
p-valor	0.0018*	0.0012*	0.0101*	0.0341*	0.0017*	0.0049*	0.0004*	0.0004*	0.0015*	0.0004*

Taula 3.4. Paràmetres estadístics de l'anàlisi d'expressió dels gens *BUB1B*, *STMN1*, *sFRP1*, *APC* i *C-MYC* obtingudes amb les tècniques de PCRq i *microarrays* d'expressió. *p-valor significatiu (< 0.05)