

ESTUDIS D'ASSOCIACIÓ I FUNCIONALS EN GENS CANDIDATS PER A L'OSTEOPOROSI

Memòria presentada per

Mariona Bustamante Pineda

Per optar al grau de

Doctora per la Universitat de Barcelona

Tesi dirigida per la Dra. Susana Balcells Comas i pel Dr. Daniel Grinberg Vaisman
al Departament de Genètica de la Facultat de Biologia
de la Universitat de Barcelona

Dra. Susana Balcells Comas

Dr. Daniel Grinberg Vaisman

Mariona Bustamante Pineda

2007

1. PLANTEJAMENT INICIAL: LA COHORT BARCOS

L'osteoporosi és una malaltia complexa determinada tant per factors ambientals com per factors genètics, essent l'heretabilitat estimada dels fenotips ossis bastant elevada. En el present treball ens vam plantejar estudiar el component genètic que determina la pèrdua de DMO després de la menopausa en població femenina espanyola. Per això vam dissenyar un estudi d'associació de tipus secció transversal, en el qual una mostra representativa de la població s'analitza en un moment determinat. La mostra estudiada està formada per dones postmenopàusiques no emparentades procedents de l'àrea de Barcelona i totes amb avantpassats espanyols. Es van excloure de la mostra les dones que presentaven trastorns ossis, metabòlics o endocrins i aquelles que estaven en tractament farmacològic el qual pogués modificar la massa òssia. Durant el desenvolupament d'aquest treball es va anar augmentant el nombre de participants fins a 876. Algunes de les dades dels estudis d'associació englobades en aquesta tesi, també van ser analitzades en paral·lel en diverses metaanàlisis prospectives dutes a terme pel consorci GENOMOS. Dins el projecte GENOMOS, la cohort de dones postmenopàusiques espanyoles va ser anomenada BARCOS.

En els estudis d'associació s'analitza la relació existent entre unes determinades variants polimòrfiques i certes variables fenotípiques. Els fenotips ossis relacionats amb l'osteoporosi que es van analitzar en aquest treball van ser la DMO femoral i lumbar. Aproximadament un 30% de les dones de la mostra presentaven osteoporosi, percentatge equivalent al calculat en població caucàsica femenina de 50 anys d'edat. Una de les conseqüències finals de l'osteoporosi és l'aparició de fractures i per això es van registrar el nombre i el tipus de fractures que havien sofert les participants. Les fractures identificades es van desglossar en fractures traumàtiques, les quals es van descartar, i en fractures osteoporòtiques prevalents. Les fractures osteoporòtiques prevalents perifèriques es van diagnosticar a través d'un qüestionari, i les vertebrals, les quals són majoritàriament indolores, es van diagnosticar a través de mètodes radiològics. Degut al baix nombre de dones amb fractures osteoporòtiques prevalents (aproximadament un 18% de la mostra), aquest fenotip no va ser analitzat en la present tesi, però sí en les metaanàlisis prospectives del projecte GENOMOS.

També es va recollir informació sobre les dades antropomètriques i ambientals que se sap que participen en la determinació de la massa òssia. Algunes de les dades que es van incorporar en l'anàlisi estadística van ser l'edat, el pes, la talla, l'IMC o els anys transcorreguts des de la menopausa fins al moment de la densitometria (YSM). Les participants de l'estudi tenien una mitjana d'edat de 55 anys, feia 7 anys de mitjana que havien tingut la menopausa, i com en població general femenina, aproximadament un 19% eren obesas ($IMC \geq 30 \text{ kg/m}^2$).

En la *Discussió* procedirem a comparar l'efecte que han mostrat tenir els polimorfismes analitzats en aquest treball amb els descrits en la bibliografia, centrant-nos especialment en els resultats del projecte GENOMOS. Seguidament, analitzarem la possible funció dels polimorfismes del promotor del gen *COL1A1* a partir dels resultats obtinguts en aquest treball i en altres. I, finalment, discutirem els avantatges, els inconvenients i el futur dels estudis d'associació.

2. COMPARACIÓ DELS RESULTATS OBTINGUTS EN AQUEST TREBALL AMB ELS RESULTATS PUBLICATS PRÈVIAMENT

2.1. LES FREQUÈNCIES AL·LÈLIQUES I HAPLOTÍPIQUES DELS POLIMORFISMES ANALITZATS

En el present treball hem estudiat l'efecte de certs polimorfismes situats tant en gens candidats clàssics per a l'osteoporosi (*VDR*, *ESR1*, *COL1A1* i *TGFB1*) com en altres gens candidats (*IL6R* i *RUNX2*) sobre la DMO lumbar i femoral en població postmenopàusica espanyola. Els polimorfismes dels gens clàssics es van analitzar perquè fins al moment el seu efecte no s'havia estudiat en una cohort de mida mitjana formada per dones postmenopàusiques d'origen espanyol. Alhora aquestes dades van incloure's en el projecte GENOMOS. Els polimorfismes dels gens *IL6R* i *RUNX2* es van seleccionar en base a la seva possible funcionalitat (SNPs no sinònims o bé situats en el promotor) o en base a evidències funcionals prèvies.

Les freqüències al·lèliques i haplotípiques obtingudes en aquest treball van ser similars a les publicades en estudis previs realitzats en població caucàsica (Ioannidis i col., 2004; Napierala i col., 2005; Wang i col., 2005; Doecke i col., 2006; Stewart i col., 2006; Uitterlinden i col., 2006) (*Resultats 1.1.2. GENOMOS: TGFB1*) i es mostren a la Taula 7. Dels 21 polimorfismes, només quatre (-800 G/A, Arg25Pro, Thr263Ile i -1025 T/C) presenten una MAF inferior a 0,1.

Taula 7 Comparació de les MAFs i de les freqüències haplotípiques obtingudes en aquesta tesi amb les publicades prèviament

Polimorfisme	MAFs			Freqüències dels haplotips majoritaris		
	Al·lel minoritari ¹	BARCOS ²	Stewart ³	Haplotips ¹	BARCOS ²	Stewart ³
COL1A1						
-1997 G/T	T	0,13	0,14	GIS (G-8T-G)	64,7%	65,3%
-1663 indelT	"D" delT 7T	0,22	0,20	GDs (G-7T-T)	19,5%	16,8%
Sp1	"s" T	0,20	0,19	TIS (T-8T-G)	11,1%	13,3%
VDR	Al·lel minoritari ¹	BARCOS ²	Uitterlinden ³	Haplotips ¹	BARCOS ²	Uitterlinden ³
Cdx2	A	0,25	0,20	No existeix desequilibri de lligament entre els polimorfismes Cdx2, FoK1 i la resta		
FokI	"f" T	0,35	0,38	baT (G-G-T)	42,5%	41,4%
BsmI	"B" A	0,43	0,42	Bat (A-T-C)	36,7%	37,3%
Apal	"a" G	0,47	0,44	bAT (G-T-T)	10,6%	13,2%
TaqI	"t" C	0,40	0,40			
ESR1	Al·lel minoritari ¹	BARCOS ²	Ioannidis ³	Haplotips ¹	BARCOS ²	Ioannidis ³
(TA) _n	H	0,43	0,46	Lpx (L-T-A)	48,4%	47,1%
PvuII	"P" C	0,46	0,46	HPX (H-C-G)	30,5%	31,0%
XbaI	"X" G	0,38	0,35	LPX (L-C-G)	7,3%	3,4%
				HPx (H-C-A)	7,2%	9,3%
TGFB1	Al·lel minoritari	BARCOS ²	Langdahl ³	Haplotips	BARCOS ²	Langdahl ^{3,4}
-800 G/A	A	0,09	0,10	G-C-T-G-C	49,4%	53,1%
-509 C/T	T	0,33	0,29	G-T-C-G-C	29,7%	27,0%
Leu10Pro	C	0,41	0,33	A-C-T-G-C	9,2%	9,9%
Arg25Pro	C	0,07	0,07	G-C-C-C-C	7,0%	7,1%
Thr263Ile	T	0,03	0,02			

Taula 7 Comparació de les MAFs i de les freqüències haplotípiques obtingudes en aquesta tesi amb les publicades prèviament (Continuació)

RUNX2	Al·lel minoritari	BARCOS²	Napierala		
-330 G/T	T	0,15	0,15	No existeix desequilibri de lligament entre els dos polimorfismes	
-1025 T/C	C	0,07	0,12-0,06		
IL6R	Al·lel minoritari¹	BARCOS²	Wang	Haplotips	BARCOS²
-1435 C/T	T	0,14	No disponible	C-G	46,2%
-208 G/A	A	0,38	0,40	C-A	39,7%
				T-G	14,1%
Asp358Ala	Ala C	0,43	0,41	No existeix desequilibri de lligament amb el polimorfisme Asp358Ala	

¹S'noten els diferents noms que poden tenir els al·lells o haplotips

²Dades obtingudes en el present treball amb la cohort de Barcelona (BARCOS)

³La comparació ha estat realitzada amb els resultats de la cohort APOSS (Aberdeen Prospective Osteoporosis Study) que participa en el projecte GENOMOS

⁴Les freqüències anotades són les dels haplotips estimats a partir dels quatre primers polimorfismes

2.2. VDR

El primer estudi d'associació realitzat en relació a l'osteoporosi va ser dut a terme per Morrison i col·laboradors l'any 1994 amb el polimorfisme *BsmI* (rs1544410) situat a la regió 3' del gen *VDR*. Aquest polimorfisme es troba en desequilibri de lligament amb els polimorfismes *Apal* (rs17879735) i *TaqI* (rs731236) també situats a la regió 3' (revisat per Uitterlinden i col., 2004). Un altre polimorfisme interessant situat en aquest gen és *FokI* (rs2228570), el qual introdueix un inici de traducció alternatiu produint-se així un receptor tres aminoàcids menor i amb més capacitat d'unió al factor TFIIB (Gross i col., 1996; Uitterlinden i col., 2004). L'any 2001, el polimorfisme 1e - 1739 G/A (rs11568820), situat en una caixa d'unió del factor de transcripció *Cdx2*, es va trobar associat a la DMO (Arai i col., 2001). La sonda que contenia l'al·lel A d'aquest polimorfisme va mostrar una capacitat d'unió al factor *Cdx2* superior a la de la sonda que contenia l'al·lel G, i a més l'activitat transcripcional del promotor amb l'al·lel A era superior a la del promotor amb l'al·lel G. *Cdx2* és un factor de transcripció que regula l'expressió del gen *VDR* en l'intestí, el qual estimula l'absorció intestinal de calci. La posició dels cinc polimorfismes descrits, els quals es troben en tres blocs haplotípiques diferents (*Cdx2*, *FokI* i *BsmI*-*Apal*-*TaqI*), es representa en la Figura 19.

El nombre de publicacions aparegudes des de l'any 1994 en relació a l'efecte d'aquests cinc SNPs sobre els fenotips osteoporòtics és considerable, però en la majoria d'aquests estudis el nombre d'individus analitzats és petit (revisat per Liu i col., 2003; Liu i col., 2006). La creació del consorci GENOMOS va permetre realitzar una metaanàlisi amb les dades de 26.000 individus per determinar l'efecte d'aquests SNPs en relació a l'osteoporosi. Aquest estudi va evidenciar que cap dels cinc polimorfismes del gen *VDR* estava associat a la DMO, i que només l'al·lel A del polimorfisme *Cdx2* reduïa modestament el risc de patir fractures vertebrals osteoporòtiques (OR=0,87 i 95%CI: 0,78-0,97) en el conjunt d'homes i dones (Uitterlinden i col., 2006).

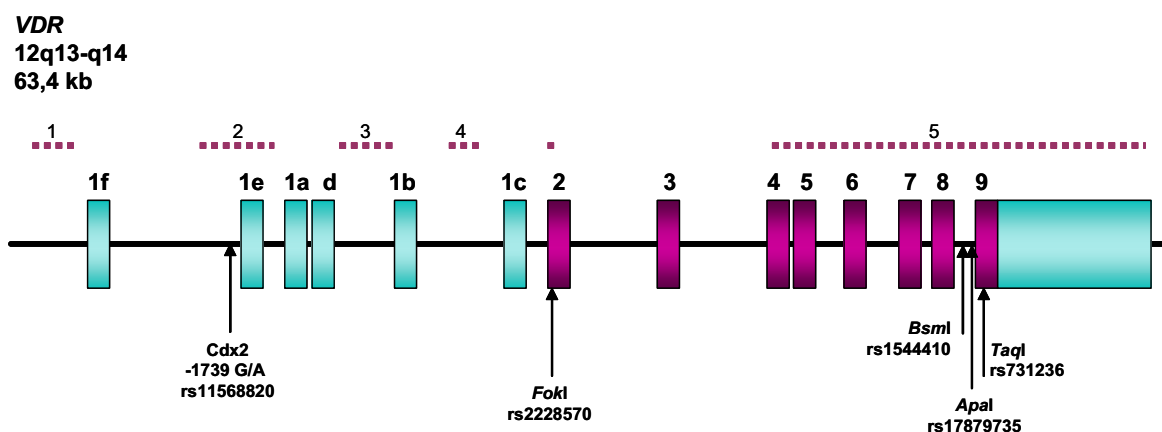


Fig. 19 Estructura gènica i polimorfismes més destacats del gen *VDR* (adaptat de Uitterlinden i col., 2004 i de Fang i col., 2005). La línia discontinua marca orientativament els diferents blocs de LD, i amb color vermell s'indica la regió codificant.

Els resultats del projecte GENOMOS coincideixen amb els resultats trobats en la cohort BARCOS en el present treball, on cap dels polimorfismes anteriors es va trobar associat a la DMO lumbar o femoral en dones postmenopàusiques espanyoles (Bustamante i col., 2007). L'únic resultat significatiu observat en relació al gen *VDR*, va ser una interacció entre el polimorfisme *Apal* (o *TaqI*) del gen *VDR* i el polimorfisme -1663 indelT (o Sp1) del gen *COL1A1* en relació a la DMO femoral. Una interacció similar havia estat descrita prèviament en relació al risc de fractura (Uitterlinden i col., 2001). Com comentarem més endavant, els resultats derivats de les interaccions d'aquest treball s'han de prendre amb precaució.

No ha estat fins l'any 2005 que s'ha realitzat un estudi exhaustiu de les variants polimòrfiques del gen *VDR* (Fang i col., 2005). Per fer-ho, els autors van definir 15 *tags* que capturaven la major part d'aquesta variabilitat en la població estudiada. Els haplotips del bloc 2 (inclou el *tag* 1e -1739 G/A o *Cdx2*) i els del bloc 3 de la regió promotora, així com també els haplotips del bloc 5 de la regió 3' del gen (que inclou els polimorfismes *BsmI*, *Apal* i *TaqI*) es van trobar associats al risc de fractura en una mostra de 6.000 individus. Dels 12 polimorfismes del promotor que es trobaven en caixes consens per a factors de transcripció, se'n van estudiar quatre: 1a -1521 G/C, 1a -1012 A/G, 1e -2090 C/T i 1e -1739 G/A (o *Cdx2*). Els estudis de retardament en gel van mostrar que els tres últims SNPs retenien factors de transcripció de manera específica d'al·lel. En concret, el polimorfisme 1e -1739 G/A (o *Cdx2*) va mostrar una retenció similar a la descrita prèviament per Arai i col·laboradors (2001). A més, els haplotips conformats pels polimorfismes 1a -1521 G/C i 1a -1012 A/G modificaven els nivells transcripcionals. D'altra banda, es va observar que els haplotips de la regió 3' del gen participen en l'estabilització del mRNA en diferents tipus cel·lulars.

Recentment s'ha publicat una metaanàlisi retrospectiva amb les dades d'uns 7.000 individus que, com l'estudi de GENOMOS, confirma la no participació dels polimorfismes *BsmI* i *TaqI*, i dels haplotips que conformen en la determinació del risc de patir fractures osteoporòtiques (Fang i col., 2006).

Finalment, cal destacar una possible interacció entre els polimorfismes del gen *VDR* i la ingesta de calci (Macdonald i col., 2006). Macdonald i col·laboradors van observar que l'efecte dels cinc polimorfismes principals del gen *VDR* sobre la DMO era nul en dones que presentaven una ingesta de calci adequada, mentre que els polimorfismes mostraven cert efecte quan la ingesta de calci era baixa. La participació de la ingesta de calci en l'efecte dels polimorfismes s'explica pel fet que la vitamina D₃ i el seu receptor *VDR* regulen l'absorció intestinal de calci, element imprescindible per a la mineralització òssia.

Vista la gran discrepància entre els resultats, la falta de treballs que analitzin conjuntament tota la variabilitat polimòrfica i la importància que sembla tenir la ingesta de calci, no podem descartar, tot i els resultats majoritàriament negatius de les metaanàlisis, que certs polimorfismes del gen *VDR* participin en la determinació de l'osteoporosi.

2.3. *ESR1*

El gen que codifica el receptor d'estrògens α, *ESR1*, ha estat associat a gran varietat de fenotips com a conseqüència de l'ampli espectre de funcions que tenen les hormones sexuals en l'organisme. En l'os, els estrògens inhibeixen l'osteoclastogènesi a través de la regulació de la producció de determinades citocines. Entre les variants polimòrfiques analitzades en el gen *ESR1* podem destacar-ne tres: el microsatèl·lit (TA)_n (rs3138774) situat en el promotor, i els SNPs *Xba*I (rs9340799) i *Pvu*II (rs2234693) situats a l'intró 1 (Fig. 20). Totes tres variants es troben en desequilibri de lligament.

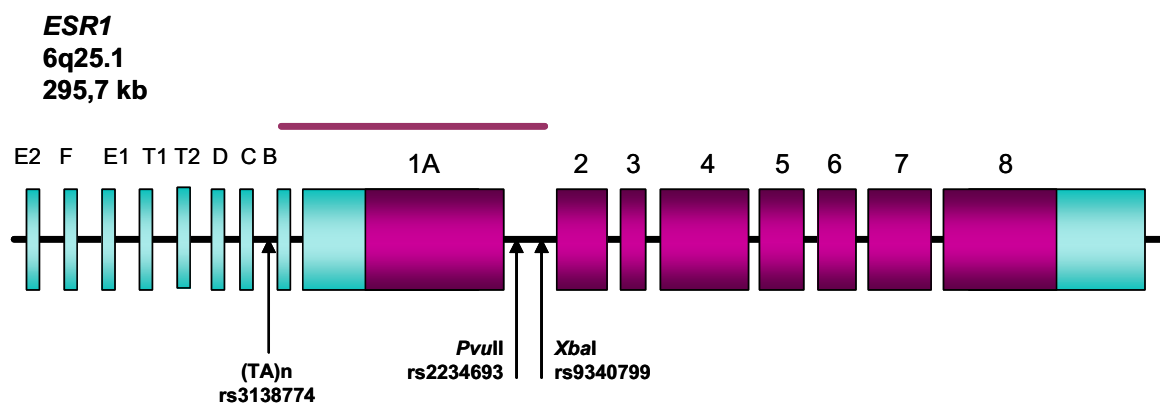


Fig. 20 Estructura gènica d'*ESR1* i localització dels polimorfismes estudiats (modificat de Gennari i col., 2002). La línia horitzontal indica que tots tres polimorfismes es troben en LD i el color vermell mostra la regió codificant.

Pel que fa a l'osteoporosi, els resultats dels estudis d'associació d'aquests polimorfismes han estat discrepants (Liu i col., 2003; Gennari i col., 2005; Liu i col., 2006). En analitzar tots tres polimorfismes en una mostra de 19.000 individus (projecte GENOMOS), es va observar que el polimorfisme *Xba*I participava en la determinació del risc de fractura (Ioannidis i col., 2004). Els individus portadors del genotip "XX" o de l'haplogenotip "HPX/HPX" (homes i dones conjuntament)

tenien una major protecció en front qualsevol tipus de fractures respecte els no portadors (aproximadament OR=0,79 i 95%CI: 0,70-0,90). El OR per les fractures vertebrals en el subgrup de dones amb aquests mateixos genotips i haplotips va ser de 0,65 (95%CI: 0,49-0,87). Cap dels tres polimorfismes va mostrar tenir un efecte estadísticament significatiu sobre la DMO lumbar o femoral. Tot i això, es va observar que els individus homozigots "XX" i "PP" tendien a presentar una major DMO. El risc de patir fractures osteoporòtiques de manera independent dels valors de DMO es pot explicar pel fet que aquests dos fenotips comparteixen només una part del component genètic (Huang i Kung, 2006).

En la cohort de dones postmenopàusiques espanyoles, vam observar que els individus amb el genotip "pp" (*Pvull*) tendien a presentar valors de DMO femoral menors, i que els portadors de l'haplotip "LPX" presentaven una mitjana de DMO femoral significativament major que la resta (Bustamante i col., 2007). De manera similar, el treball d'Albagha i col.laboradors (2005), va descriure que l'haplotip "px" es trobava associat a menor DMO. És fa difícil avaluar la correlació d'aquests resultats amb els de GENOMOS, ja que tot i que a causa del desequilibri de lligament els al·lels de risc són els mateixos, els fenotips (fractura i DMO) i els models genètics a través dels quals estan associats no són equivalents.

De la mateixa manera que l'efecte dels polimorfismes del gen *VDR* sembla dependre de la ingesta de calci, l'efecte dels SNPs del gen *ESR1* podria dependre del sexe, de l'estat de menopausa i de possibles interaccions amb polimorfismes situats en gens com el *ESR2* (Ioannidis i col., 2004; Albagha i col., 2005; Rivadeneira i col., 2006).

2.4. *TGFB1*

El TGFβ1 és un factor de creixement essencial per a l'osteoblast. Els polimorfismes analitzats en el gen *TGFB1* es poden classificar segons la seva posició en: variants del promotor [-800 G/A (rs1800468) i -509 C/T (rs1800469)], variants situades en el pèptid senyal [Leu10Pro (rs1982073) i Arg25Pro (rs1800471)] i variants situades en el pèptid latent associat (LAP), el qual protegeix i inactiva la forma madura del TGFβ1 [Thr263Ile (rs1800472)] (Fig. 21).

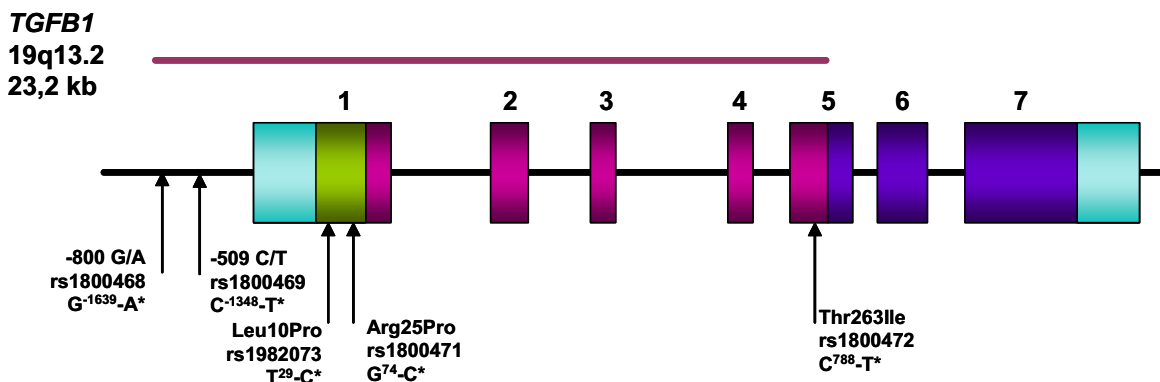


Fig. 21 Representació esquemàtica del gen *TGFB1* i d'alguns dels polimorfismes més rellevants (adaptat de Resultats 1.1.2. GENOMOS: *TGFB1*). La línia horitzontal indica la regió de LD, en verd s'indica el pèptid senyal i en lila la regió del pèptid madur. La resta de la regió codificant (vermell) dona lloc al LAP. *Nomenclatura segons GENOMOS.

Malgrat que aquests polimorfismes han estat associats a fenotips relacionats amb l'os en diversos estudis (revisat per Janssens i col., 2005), les dades obtingudes de l'anàlisi conjunt de GENOMOS van mostrar que cap dels polimorfismes ni cap dels seus haplotips influenciava la DMO o el risc de patir fractures (*Resultats 1.1.2. GENOMOS: TGFB1*). En analitzar subgrups de la mostra de GENOMOS es va observar que els polimorfismes -509 C/T i Thr263Ile influenciaven lleugerament la DMO i el risc de fractura, respectivament. Els autors, però, demanaven precaució en valorar aquests resultats. El SNP Thr263Ile presenta una MAF molt petita (0,03) i es va trobar associat només al subgrup amb fractures vertebrals incidents (n=412), i el SNP -509 C/T es va trobar associat a la DMO lumbar només en el subgrup d'homes.

En la mostra de 719 dones postmenopàusiques espanyoles es va observar que dels cinc polimorfismes estudiats només el Leu10Pro es trobava associat a la DMO (Bustamante i col., 2007). Els individus homozigots CC presentaven menor DMO lumbar que la resta. Aquest resultat coincideix amb el publicat per alguns grups (Hinke i col., 2001; Dick i col., 2003; Lau i col., 2004), però discrepa del publicat per altres, ja sigui perquè alguns van trobar l'associació inversa (Yamada, 2000; Langdahl i col., 2003; Tzakas i col., 2005) o ja sigui perquè altres no van trobar cap associació (McGuigan i col., 2006). Els haplotips estimats a partir dels cinc polimorfismes estudiats no es van trobar associats a la DMO.

Una ampliació del nombre de genotips de la mostra (aproximadament 200 genotips més) va comportar la pèrdua de l'associació significativa entre el polimorfisme Leu10Pro i la DMO (dades de la cohort BARCOS de GENOMOS (*Resultats 1.1.2. GENOMOS: TGFB1*)). En augmentar la mostra és freqüent observar oscil·lacions en els resultats, normalment a causa de la magnitud de l'efecte del polimorfisme. Com menor és la magnitud de l'efecte d'un SNP sobre el fenotip, més difícil és detectar diferències fenotípiques entre cadascun dels genotips, i petits canvis en la mostra poden fer que aquestes diferències deixin de ser significatives. Tot i que tant nosaltres com altres autors han descrit la mateixa associació pel polimorfisme Leu10Pro, cal destacar que la metaanàlisi de GENOMOS no ha estat capaç de detectar-lo.

2.5. *RUNX2*

El paper clau que juga la proteïna Runx2 en la diferenciació osteoblàstica fa que el gen *RUNX2* sigui un bon candidat pels estudis d'associació amb fenotips relacionats amb l'osteoporosi. Partint de la hipòtesi que els polimorfismes que modifiquin els nivells d'expressió d'aquest factor podrien tenir un efecte sobre els nivells de la DMO, ens vam plantejar estudiar polimorfismes situats en les regions promotores d'aquest gen. Com hem comentat en la *Introducció*, el gen *RUNX2* presenta dos promotors (P1 i P2) que disten entre ells unes 90 kb i que donen lloc a dues isoformes majoritàries en el teixit ossi (Fig. 16). Per aquest motiu vam decidir investigar l'efecte de dos polimorfismes, cada un d'ells situat en un dels dos promotors (Fig. 22). Aquests dos polimorfismes es van escollir en base a la seva funcionalitat i/o en base a estudis previs que mostraven la seva associació amb la DMO.

RUNX2
6p12.3
222,8 kb

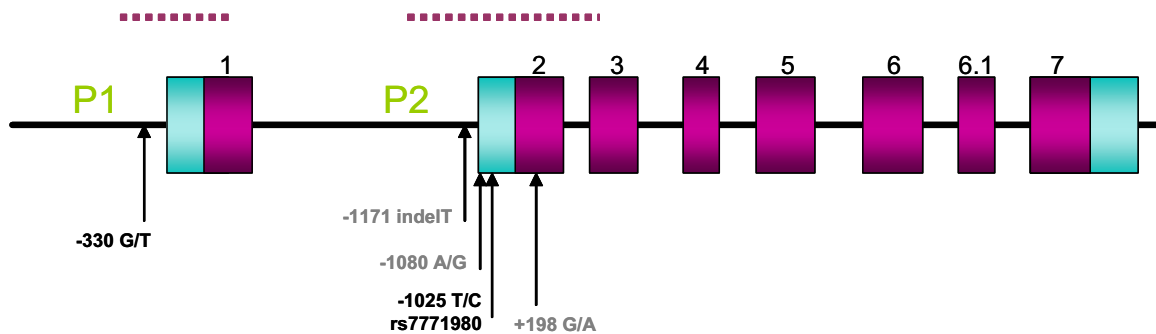


Fig. 22 Estructura gènica de *RUNX2* i localització d'alguns polimorfismes (adaptat de Doecke i col., 2006). Es mostren les regions aproximades de LD amb una línia horitzontal discontinua. El color vermell marca la regió codificant i els 2 promotors s'indiquen com P1 i P2.

El polimorfisme -1025 T/C (rs7771980) (P2) es va escollir perquè s'havia descrit que aquest SNP junt amb dos polimorfismes més del promotor 2 (P2) estaven associats a la DMO de fèmur en població australiana (Doecke i col., 2006). Els individus portadors de l'al·lel C del polimorfisme -1025 T/C presentaven una DMO femoral més elevada que els portadors de l'al·lel T. A nivell funcional, es va descriure que l'al·lel T era capaç de retenir un complex proteic que no era retingut per l'al·lel C, i que l'haplotip minoritari (que contenia l'al·lel C) mostrava major activitat transcripcional. D'acord amb aquestes dades, l'haplotip que conté l'al·lel C del polimorfisme -1025 T/C podria estar modulant a l'alça els nivells d'expressió del gen *RUNX2*, fet que afavoriria la diferenciació osteoblàstica i conseqüentment un augment de la DMO.

Els polimorfismes del promotor 2 (P2) es troben en desequilibri de lligament amb la variant sinònima +198 G/A de l'exó 2, la qual també es va trobar associada a la DMO en tres estudis: un amb població escocesa i dos amb població australiana (Vaughan i col., 2002; Vaughan i col., 2004; Doecke i col., 2006); (Fig. 22). Altres polimorfismes propers al P2 es van trobar associats a la llargada del fèmur (Ermakov i col., 2006).

En una mostra de 821 dones postmenopàusiques espanyoles vam observar que el polimorfisme -1025 T/C es trobava associat a la DMO femoral i que els individus TC presentaven una mitjana de DMO major que els individus TT (els individus amb el genotip CC, pel fet de ser només el 0,6% de la mostra, es van descartar) (*Resultats. 1.2.1. RUNX2*). Aquests resultats, obtinguts amb un mostra independent, constitueixen una rèplica dels resultats de Doecke i col·laboradors (2006). Malgrat la major mida mostral per a la DMO lumbar i una potència estadística similar o lleugerament major, no vam trobar que aquest SNP tingués cap efecte sobre la determinació de la massa òssia en columna.

D'altra banda també vam decidir genotipar el polimorfisme -330 G/T situat en el promotor 1 (P1). S'havia demostrat prèviament que els al·lells d'aquest polimorfisme comportaven una unió diferencial a factors nuclears (Napierala i col., 2005). Tot i això, a nivell transcripcional, no es van trobar diferències entre els al·lells. En el present estudi el polimorfisme -330 G/T no es va trobar

associat a la DMO, ni lumbar ni femoral. Aquest resultat manté certa correspondència amb els resultats de Doecke i col·laboradors (2006). Aquests autors, tot i que no van analitzar el polimorfisme -330 G/T, van genotipar dos polimorfismes situats en el promotor 1 (P1) (a unes 400 pb del SNP -330 G/T) i no van trobar cap associació significativa entre ells i la DMO femoral.

Els dos promotors del gen *RUNX2* es troben en dos blocs haplotípics segons les dades de HapMap (Data Rel19/phaseII Oct 2005), les de Doecke i col·laboradors (2006) i les presentades en aquesta tesi. Els polimorfismes situats en el P2 semblen influenciar la DMO, mentre que els polimorfismes situats en el P1 no mostren cap efecte sobre aquest fenotip. En un principi es pensava que la isoforma I de Runx2, procedent del P2, no era específica d'os, però recentment s'ha vist que aquesta isoforma també és necessària per a la correcta diferenciació osteoblàstica (Xiao i col., 2004; Stock i Otto, 2005).

2.6. *IL6R*

Diversos grups han detectat l'existència de lligament entre l'osteoporosi i marcadors de la regió 1q21 on es troba el gen *IL6R* (revisat per Huang i Kung, 2006). Tot i això, altres han exclòs, a través d'estudis de lligament, la presència en aquesta regió de QTLs que produeixin un efecte sobre la DMO que expliqui un 10% de la variabilitat fenotípica (Chen i col., 2006). Com s'ha comentat amb anterioritat aquest receptor presenta dues formes: una de transmembrana i una de soluble (sIL6R), la qual en comptes d'actuar com a antagonista del coreceptor gp130 actua com a agonista (Kallen, 2002). La forma soluble es produeix o bé per *splicing* alternatiu o bé pel tall proteolític de la forma de membrana (Jones i col., 2001).

En l'os, la IL6 i els seus receptors gp130 i IL6R promouen principalment l'osteoclastogènesi (Jilka i col., 1992; Tamura i col., 1993). En pacients amb diabetis de tipus II, obesitat o osteoporosi, i també en processos fisiològics com l'envelliment o la menopausa s'han detectat nivells sèrics elevats de IL6 i de sIL6R (Mohamed-Ali i col., 1997; Fried i col., 1998; Kado i col., 1999; Abrahamsen i col., 2000; Scheidt-Nave i col., 2001). Per això no és d'estranyar que polimorfismes en el gen *IL6R* s'hagin trobat associats a la diabetis i a l'obesitat (Escobar-Morreale i col., 2003; Wolford i col., 2003; Hamid i col., 2004; Wang i col., 2005; Esteve i col., 2006; Zhang i Tian, 2006), però fins ara no s'ha descrit cap associació entre aquest gen i els fenotips osteoporòtics.

Aquest és, doncs, el primer treball d'associació on s'han estudiat polimorfismes del gen *IL6R* en relació a la DMO. Les tres variants analitzades han estat el SNP -1435 C/T situat en la regió promotora, el SNP -208 G/A situat en la regió 5'UTR i un canvi d'aminoàcid (Asp358Ala) situat a l'exó 9 en la regió on té lloc el tall proteolític per donar la forma soluble (Fig. 23). Degut a que l'estudi no engloba tota la variabilitat genètica present en el gen, els polimorfismes s'han escollit en base a la seva possible funcionalitat. El polimorfisme Asp358Ala és l'únic SNP del gen que dona lloc a un canvi d'aminoàcid i que se'n coneix la MAF. Els altres dos polimorfismes es

Discussió

troben en la regió promotora i, segons evidències *in silico*, podrien unir factors de transcripció de manera específica d'al·lel.

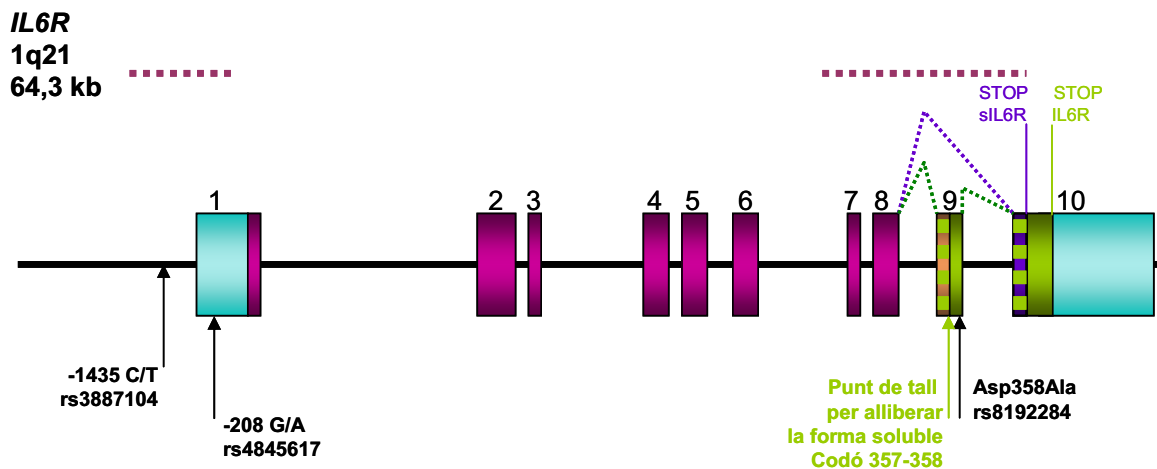


Fig. 23 Esquema del gen *IL6R* on s'indiquen els polimorfismes analitzats. Les regions de LD es mostren amb una línia horitzontal discontinua. Els exons 1-10 (exceptuant el 9 que és saltat durant el procés d'*splicing*) codifiquen la forma soluble (sIL6R). Els exons 1-10 (incloent l'exó 9) codifiquen la forma de transmembrana, la qual a través d'un tall proteolític majoritàriament entre el codó 357 i el codó 358 de l'exó 9 dona lloc a l'altre forma soluble coneguda del receptor. En vermell s'indica la regió codificant comuna en les tres isoformes. Les regions específiques s'indiquen en taronja (sIL6R provinent del tall proteolític), en lila (sIL6R provinent de l'*splicing* alternativu) i en verd (IL6R).

Després de genotipar aquests tres polimorfismes en aproximadament 450 dones, vam observar que els dos polimorfismes del promotor es trobaven associats a la DMO femoral (*Resultats 1.2.2. IL6R*). En concret es va observar que els individus homozigots CC del polimorfisme -1435 C/T o homozigots AA del polimorfisme -208 G/A tenien valors de DMO femorals superiors a la resta d'individus. Com que ambdós polimorfismes es trobaven en fort desequilibri de lligament, a nivell haplotípic es va observar que els portadors de 2 còpies de l'haplotip C-A presentaven valors de DMO femoral superiors als de la mitjana de la població i que els portadors de com a mínim una còpia de l'haplotip T-G mostraven valors de DMO menors als de la mitjana (*Resultats 1.2.2. IL6R; Fig. 1*). Aquestes dades suggereixen, doncs, que l'haplotip C-A és protector, mentre que l'haplotip T-G confereix cert risc.

Tot i l'existència de polimorfismes "redundants" amb els estudiats en el promotor (HapMap Data Rel21a/phasell Gen 2007), en el futur seria interessant realitzar estudis d'EMSA i de gen reporter per validar o descartar la seva funcionalitat, ja que els programes Genomatix i SignalScan mostren que els dos SNPs presenten caixes d'unió a factors de transcripció específics d'al·lel. En concret, l'al·lel T del polimorfisme -1435 C/T podria unir els factors de transcripció Cdx2 i GATA, mentre que l'al·lel C podria unir el factor E2F. Pel que fa al SNP -208 G/A, l'al·lel A podria unir específicament el factor AML-1a i l'al·lel G el factor NF-1.

D'altra banda vam trobar que el polimorfisme Asp358Ala es trobava associat a la DMO de columna. Es va observar que els individus homozigots per l'al·lel Ala presentaven valors més elevats de DMO lumbar. Aquest polimorfisme es troba situat, en la isoforma de transmembrana, en la regió on es produeix el tall que allibera la forma soluble del receptor (sIL6R). S'ha descrit que la

mutació puntual de l'Asp358 a Gly, Phe, Leu o Arg redueix el tall al 22-54% respecte el 100% del control, i que la doble mutació a Ala357 i Ala358 el redueix al 34% (Mullberg i col., 1994). D'acord amb aquests resultats, es pot hipotetitzar que l'al·lel Ala de la posició 358 implicaria una reducció del tall i per tant menys nombre de molècules de la forma soluble del receptor (sIL6R) (Fig. 24). La disminució de la forma soluble produiria una reducció de l'osteoclastogènesi i per tant un major valor de la DMO. Està descrit que l'efecte osteoclastogènic de la IL6 es duu a terme a través dels osteoblasts i que aquests només responen quan el receptor IL6R es troba en la seva forma soluble (Udagawa i col., 1995; Palmqvist i col., 2002; Vermes i col., 2002).

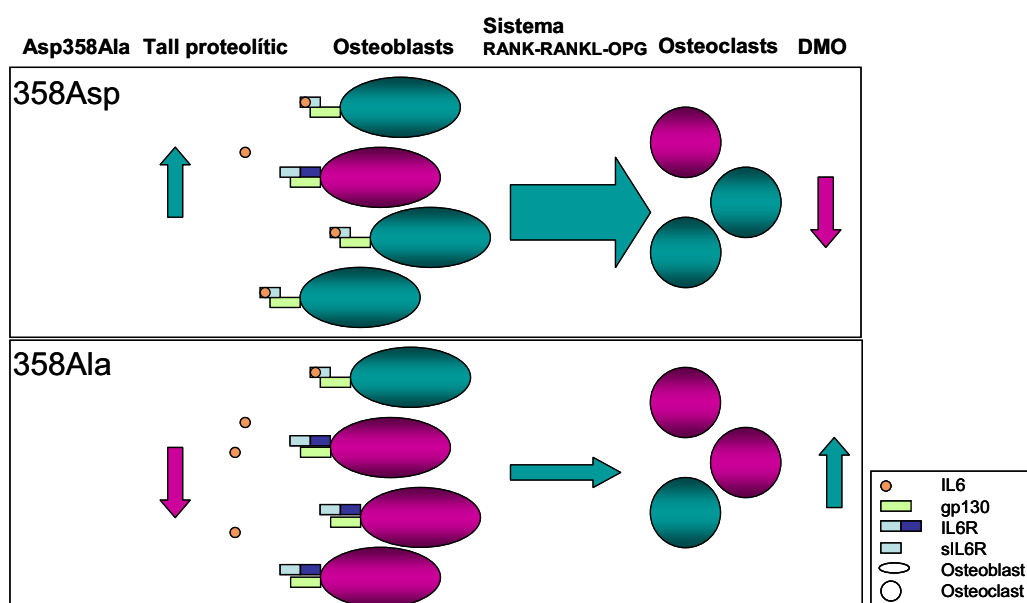


Fig. 24 Esquema del possible mecanisme d'acció del SNP Asp358Ala en relació a la DMO. Les figures verdes indiquen un increment en l'activitat, en el nombre o en el valor final de la DMO, mentre que les figures vermelles indiquen una disminució.

Malgrat aquests indicis no es pot descartar que el polimorfisme Asp358Ala estigui actuant, degut al desequilibri de lligament, com a marcador del polimorfisme causant de l'efecte. Segons les dades de HapMap (Data Rel21a/phasell Gen 2007), en el bloc haplotípic on es troba aquest SNP hi ha una sèrie de polimorfismes "redundants", però tots ells són intrònics i per ara sense evidències de ser funcionals.

En aquest treball, s'ha trobat que alguns polimorfismes del gen *IL6R* estan associats a la DMO lumbar i altres a la DMO femoral. Aquestes dues regions esquelètiques no presenten la mateixa proporció d'os trabecular i d'os cortical: la columna està formada per un 66% d'os trabecular, mentre que el fèmur està format principalment per os cortical (75%). L'os trabecular i l'os cortical no difereixen només morfològicament, sinó també fisiològicament, de manera que no és d'estranyar que l'efecte d'un polimorfisme concret no sigui equivalent en ambdues regions esquelètiques.

Finalment, es va observar que el polimorfisme -208 G/A i els haplotips del promotor també es trobaven associats a l'IMC i a l'obesitat en la cohort BARCOS. En concret, els individus homozigots GG, així com també els portadors de dues còpies de l'haplotip C-G o bé els portadors

de 0 còpies de l'haplotip C-A presentaven un major risc de patir obesitat. Prèviament s'havia descrit que el SNP Asp358Ala es trobava associat a l'IMC en població espanyola (Esteve i col., 2006), però nosaltres vam ser incapaços de replicar-ho en la cohort BARCOS. Aquest mateix SNP també es va trobar associat a l'IMC en una mostra d'indis Pima (Wolford i col., 2003), però no es va trobar associat en població danesa ni americana (Hamid i col., 2004; Wang i col., 2005). Un possible motiu de la falta de replicació entre els dos estudis realitzats en població espanyola pot ser l'estructura de les cohorts: mentre que la cohort BARCOS només conté dones postmenopàusiques, la d'Esteve i col·laboradors (2006) és una mostra mixta formada per homes i dones, un 12% dels quals presentaven la síndrome metabòlica. El LD existent entre els polimorfismes -208 G/A i Asp358Ala és baix, de manera que és poc probable que el primer SNP faci de marcador de l'altre o viceversa. També és important esmentar que tant la mostra del present estudi com la d'Esteve i col·laboradors (2006) no són molt grans, no podent-se descartar falsos positius o problemes de falta de potència estadística. Per això seria necessari estudiar l'efecte d'aquests tres polimorfismes i d'altres variants del gen *IL6R* en mostres independents de major mida mostral.

■ Relació entre el teixit ossi i el teixit adipós

Aquesta no és la primera vegada que un polimorfisme es troba associat tant a la DMO com a l'IMC. El mateix gen que codifica la citocina IL6 (liligand del *IL6R*), presenta variants polimòrfiques que s'han relacionat simultàniament a l'osteoporosi i a l'obesitat (Huang i col., 2003b). També s'ha trobat associat a la DMO i a l'obesitat un polimorfisme del gen *RANK*, gen que com *IL6* i *IL6R* participa en el procés de resorció òssia (Zhao i col., 2006). En l'estudi de Zhao i col·laboradors l'al·lel de risc per a l'obesitat també conferia més risc de patir osteoporosi. En la cohort BARCOS els individus homozigots GG pel polimorfisme -208 G/A eren els que presentaven major IMC i, alhora, els portadors de l'al·lel G eren els que tenien menor DMO femoral. És a dir, existeix una relació inversa entre la DMO femoral i l'IMC quan estratifiquem els individus en funció dels genotips del polimorfisme -208 G/A.

En general el pes o l'IMC s'han correlacionat positivament amb la DMO en ambdós gèneres i al llarg de diferents rangs d'edat (Reid, 2002). La pèrdua brusca de pes també ha estat relacionada amb disminucions de la massa òssia. Però, recentment alguns estudis han mostrat que mentre la massa magra correlaciona positivament amb la DMO, el percentatge de greix corporal (aproximadament un 16-25% del pes corporal) hi correlaciona negativament, posant en dubte, o si més no, acotant la relació descrita anteriorment (revisat per Rosen i Bouxsein, 2006; Zhao i col., 2007).

S'han formulat diverses hipòtesis per tal d'explicar la relació existent entre la DMO i l'IMC (revisat per Rosen i Bouxsein, 2006). La primera hipòtesi proposa que la pressió mecànica causada per un IMC elevat estimula la formació òssia (Robling i col., 2006). La segona hipòtesi relaciona el greix corporal amb l'augment de la secreció d'insulina i altres hormones, les quals tenen un paper anabòlic en l'os (Ronti i col., 2006). En tercer lloc, s'ha vist que l'augment dels

adipòcits comporta un augment del procés d'aromatització i per tant de la producció d'estrògens, així com també un augment de la secreció de citocines adipocitàries com la leptina i la IL6. La leptina, hormona que regula la sensació de sacietat a través del SNC, actua de manera dual sobre l'os: a través del SNC n'inhibeix la formació, però a nivell local n'estimula la formació. Per la seva banda, la IL6 estimula la resorció òssia a nivell local, i a través del SNC disminueix el greix corporal mitjançant l'augment de la despesa energètica i la disminució de la ingesta. Els ratolins als quals se'ls ha delectonat el gen *Il6* desenvolupen obesitat d'aparició tardana, però estan protegits contra la pèrdua d'os després de l'ovariectomia (Poli i col., 1994; Wallenius i col., 2002).

Un altre punt de connexió entre el teixit ossi i el teixit adipós és l'origen comú de les cèl·lules que els constitueixen (revisat per Rosen i Bouxsein, 2006); (Fig. 4). Durant l'envelliment s'ha vist que disminueix la diferenciació osteoblàstica i augmenta el nombre d'adipòcits. Les dones osteoporòtiques presenten una relació inversa entre la formació òssia i el greix situat a la medulla òssia, i les dones obeses presenten un menor índex de marcadors de formació òssia.

Finalment, a nivell genètic els estudis d'heretabilitat han mostrat que la DMO, en especial la de fèmur, i l'IMC comparteixen entre un 10-20% de la variabilitat genètica (Deng i col., 2006). Zhao i col·laboradors (2007) van descriure una correlació genètica negativa entre la massa òssia ajustada pel pes i el greix corporal, fet que suggereix que els factors genètics compartits produeixen efectes oposats sobre ambdós fenotips. A part dels gens que hem comentat anteriorment (*IL6* i *RANK*), es coneixen altres gens que influeixen tant l'osteoporosi com l'obesitat. Aquests gens són: *IGF-II*, *LEPR*, *NPY*, *VDR*, *ESR1*, *AR*, *TGFB1*, *IL6*, *TNF α* , *TNFR2*, *ApoE* i *PPAR γ* .

En resum podem dir que les connexions entre el teixit adipós i el teixit ossi són múltiples: ambdues cèl·lules provenen de les cèl·lules mesenquimàtiques, estan controlades pel SNC i comparteixen un gran nombre d'hormones i citocines reguladores (Taula 8). Si la relació entre una major taxa de greix corporal i una menor DMO es confirma i es perfilen quins són els mecanismes moleculars que hi intervenen, això tindrà importants repercussions en la clínica d'ambdues malalties (Zhao i col., 2007).

Taula 8 Propostes a favor i en contra del supòsit que "el greix corporal protegeix del risc de patir osteoporosi" (revisat per Rosen i Bouxsein, 2006)

<i>A favor</i>	<i>En contra</i>
<ul style="list-style-type: none"> ■ L'increment de l'estrès mecànic afavoreix la formació d'os cortical ■ El greix corporal incrementa la protecció enfront al risc de fractura en les caigudes ■ La leptina estimula la formació òssia de manera local ■ A major quantitat de teixit adipós, major taxa d'aromatització i per tant major quantitat d'estrògens i menor resorció òssia ■ Les adiponectines estimulen la producció d'insulina que estimula la formació òssia 	<ul style="list-style-type: none"> ■ Les citocines produïdes pels adipòcits frenen la formació òssia ■ Els àcids grassos estimulen la resorció ■ La leptina inhibeix la formació òssia via el SNC ■ La hiperglicèmia (diabetis de tipus I i II) frena la formació òssia ■ La insulinoopènia (diabetis de tipus I) frena la formació òssia ■ L'activació del PPARγ, estimula la resorció i inhibeix la formació òssia

2.7. COL1A1

2.7.1. Associació dels polimorfismes -1997 G/T, -1663 indelT i Sp1 amb la DMO i el risc de patir fractures osteoporòtiques

El gen *COL1A1* codifica la cadena $\alpha 1(I)$ que conjuntament amb la cadena $\alpha 2(I)$ constitueixen el col·lagen de tipus I, la proteïna majoritària de la matriu òssia. Un dels polimorfismes situats en el gen *COL1A1* associat a fenotips osteoporòtics és el SNP Sp1 o +1245 G/T (rs1800012). Aquest SNP, situat a l'intró 1 del gen, es va trobar associat a la DMO per primera vegada l'any 1996 i des de llavors ha estat analitzat en múltiples estudis sense resultats concloents (Grant i col., 1996; revisat per Liu i col., 2003 i per Liu i col., 2006). Malgrat la disparitat de resultats previs, la metaanàlisi del consorci GENOMOS sembla confirmar que aquest SNP participa tant en la determinació de la DMO com en el risc de patir fractures osteoporòtiques vertebrals (Ralston i col., 2006). En la metaanàlisi els individus "ss" (TT) presentaven menor DMO femoral i lumbar, així com també més risc de patir fractures osteoporòtiques. En concret el genotip "ss" implicava de mitjana una disminució de 21 mg/cm² en la DMO lumbar i una disminució de 25 mg/cm² en la DMO femoral. Les dones presentaven, per cada al·lel "s", un increment del 33% del risc de patir fractures vertebrals incidents (OR=1,33 i 95%CI: 1,00-1,77).

El polimorfisme Sp1 es troba en desequilibri de lligament amb dos polimorfismes situats en el promotor del gen, -1997 G/T (rs1107946) i -1663 indelT (rs11327935), descrits per Garcia-Giralt i col·laboradors (2002); (Fig. 25). Els autors van observar que el polimorfisme -1997 G/T i la interacció d'aquest amb el polimorfisme -1663 indelT (o bé amb el polimorfisme Sp1) estaven associats a la DMO lumbar en una mostra de 256 dones postmenopàusiques espanyoles de la cohort BARCOS. En aquesta tesi s'ha ampliat el genotipatge de tots tres polimorfismes fins aproximadament 600 mostres. En la mostra ampliada, el SNP -1997 G/T seguia trobant-se associat a la DMO lumbar, essent els individus heterozigots (GT) els que presentaven major DMO (Bustamante i col., 2007). Aquests resultats suggereixen que podria existir un avantatge pels individus heterozigots, però degut a la petita quantitat d'individus homozigots TT (n=12) aquesta hipòtesi s'ha de prendre amb precaució. De fet, sota un model dominant es van obtenir diferències estadísticament significatives entre la DMO dels homozigots GG i la DMO dels portadors de l'al·lel T (GT+TT). En la mostra ampliada, no vam detectar cap associació entre els polimorfismes Sp1 o -1663 indelT i la DMO.

A nivell haplotípic vam observar que tres dels haplotips explicaven el 95% de la variabilitat haplotípica poblacional espanyola, i que els individus portadors de dues còpies de l'haplotip GDs (G-7T-T) presentaven un valor de DMO femoral significativament menor que la resta.

Com hem dit, en la cohort BARCOS de 256 individus s'havia descrit una interacció entre el polimorfisme -1997 G/T i el polimorfisme Sp1 (o bé -1663 indelT) en relació a la DMO de columna, la qual es va mantenir després de l'ampliació de la cohort. Els individus portadors d'una o dues còpies de l'al·lel T (-1997 G/T), i alhora, portadors d'una o dues còpies de l'al·lel "s" (Sp1) presentaven valors de DMO majors que la resta. Degut al desequilibri de lligament existent entre

tots tres polimorfismes, aquests individus haurien de ser majoritàriament portadors de l'haplotip TIS (T-8T-G) i de l'haplotip GDs (G-7T-T), és a dir, de l'haplogenotip TIS/GDs i per tant triple heterozigots.

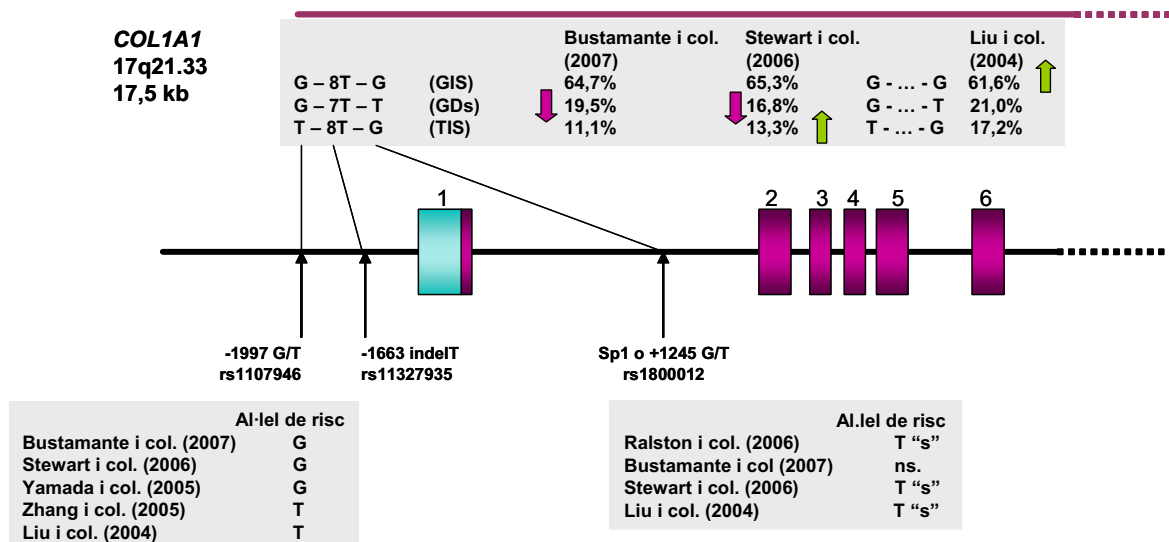


Fig. 25 Representació parcial del gen *COL1A1* i dels polimorfismes analitzats en el present treball (adaptat de Stewart i col., 2006). La línia horitzontal indica que tots tres polimorfismes es troben en una regió de LD que s'estén fins a l'intró 11. Els exons codificants estan marcats en vermell. Els resultats dels estudis d'associació dels SNPs -1997 G/T i Sp1, així com també dels haplotips majoritaris, es mostren en els requadres grisos. Pels polimorfismes associats s'indica l'al·lel de risc. Pels haplotips s'indica la freqüència present en la població així com també quins d'ells es van trobar associats a la DMO (fletxa vermella=haplotip de risc i fletxa verda=haplotip protector).

En un estudi realitzat amb unes 3.000 dones postmenopàusiques britàniques es va observar que els tres polimorfismes anteriors influenciaven la DMO, tant lumbar com femoral (Stewart i col., 2006). Els al·lells de risc van ser l'al·lel G (-1997 G/T), l'al·lel delT (-1663 indelT) i l'al·lel "s" (Sp1). Com en el present estudi, els individus homozigots per l'haplotip GDs (G-7T-T) presentaven una baixa DMO; però a diferència dels nostres resultats, els individus homozigots TIS (T-8T-T) presentaven una DMO significativament superior a la resta. Aquest fet es pot explicar per la menor potència estadística del nostre estudi, o bé per diferències ambientals o genètiques entre ambdues poblacions.

Liu i col.laboradors (2004) van trobar també que la interacció entre els polimorfismes -1997 G/T i Sp1 estava associada a la DMO en dones caucàsiques d'avançada edat. A diferència del descrit fins ara, els autors van trobar que les dones homozigotes per l'haplotip GS (G-...-G) presentaven major DMO que la resta.

El polimorfisme -1997 G/T s'ha trobat associat a la DMO en poblacions de diferent origen. Com en població caucàsica, en japonesos es va detectar l'al·lel G associat a menor DMO, mentre que en població xinesa, es va veure que l'al·lel T actuava com a al·lel de risc (Yamada i col., 2005; Zhang i col., 2005).

2.7.2. Funció dels polimorfismes -1997 G/T, -1663 indelT i Sp1

■Estudis funcionals previs a la realització d'aquesta tesi

D'acord amb els resultats dels estudis d'associació sembla, doncs, que els polimorfismes -1997 G/T, -1663 indelT i Sp1 del gen *COL1A1* podrien participar en la determinació de la DMO. Prèviament a aquesta tesi, s'havien realitzat una sèrie d'estudis amb l'objectiu de determinar si aquests SNPs tenien algun efecte en la regulació de la transcripció o en l'estabilització del mRNA.

Un d'aquests estudis va descriure que la seqüència nucleotídica on es troba el polimorfisme Sp1 unia el factor de transcripció del mateix nom de manera específica d'al·lel, sent l'al·lel "s" el que mostrava una major retenció (Grant i col., 1996; Mann i col., 2001). En els individus heterozigots, el transcrit derivat de l'al·lel "s" era tres vegades més abundant que el transcrit de l'al·lel "S". Aquests mateixos individus tenien un major ratio de proteïna $\alpha 1(I)$ respecte $\alpha 2(I)$. Es va proposar que la major quantitat de cadena $\alpha 1(I)$ podria desequilibrar la proporció entre dues cadenes $\alpha 1(I)$ i una $\alpha 2(I)$ necessària per a la formació del col·lagen de tipus I (Fig. 7), i això afectaria els nivells de DMO i la capacitat de mineralització òssia (Mann i col., 2001; Stewart i col., 2005). Cal destacar que aquests experiments, però, no van permetre discernir entre si era el mateix SNP Sp1 qui dirigia l'efecte observat o bé un polimorfisme en LD amb ell.

D'altra banda, es va observar que els polimorfismes -1997 G/T i -1663 indelT presentaven capacitat d'unió a proteïnes nuclears de manera específica d'al·lel (Garcia-Giralt i col., 2002). Les seqüències nucleotídiques de doble cadena que contenien l'al·lel més freqüent per cadascun dels polimorfismes, G (-1997 G/T) o inT (-1663 indelT), eren les que mostraven més capacitat de retenció. A nivell de cadena senzilla, la seqüència *reverse* que contenia l'al·lel C del polimorfisme -1997 G/T i la seqüència *forward* que contenia l'al·lel inT (8T) del polimorfisme -1663 indelT retenien les proteïnes nuclears més intensament que les respectives cadenes complementàries. El polimorfisme -1663 indelT es troba en una seqüència nucleotídica del gen *COL1A1* homòloga a la seqüència de rata on s'uneix la proteïna CIZ/NMP4 (*site B*) (Alvarez i col., 1998); (Fig. 26). Mitjançant un experiment de *supershift* es va veure que CIZ/NMP4 era una de les proteïnes del complex retingut per l'oligonucleòtid de cadena senzilla que contenia l'al·lel inT (8T) (Garcia-Giralt i col., 2005). També es va observar que els haplotips donaven lloc a diferències transcripcionals.

-1997 G/T

```

Humà      ACAGGACCCCTGTCGCCTATTAGGGAG--GGGCTCTC-TTGAACTGACCCACAGTAGGGGGCAGGACTTTGGTG
Gos       ACAGAATCCCTTCTCCCCTGTTGGGGAG--GGGCTCTCCTTATCTCTGACCCACA--AGGTGGGGGGAC-----
Ratolí    ACACCATCCTTGTCCACTGTTGGGGAGTGGGGATCCTTTGAACTGACC-ACAGTGAAGGCAGGACTTTGGTG
Rata      ACACCATCCCGTCCACTGTTAGGGAGTGGG-ATCCTTTGAACTGACC-ACTGTGAAGGCAGGACTTTACTG
    
```

-1663 indelT

```

Humà      ---CGGCAGCCGACGTGGCTCCCTCCCTCTCTGTTCTTTTTTTTCCCTTTGCCTTCGTTGCACAAAAACAGCT
Gos       ---TGGCTCCCTCCCCTTCTGTTCCTTGGTCTTTTTTTTTTTTTTTTCTTTGCCTTCGTTGCACAAAAACAGCT
Ratolí    ACGTGGCTCCCTCCCCTTCTGTTCCTTGGTCT-TTTTT--TTTTTTCCCTTTGCCTTCGTTGCACAAAACTAGCT
Rata      ACGTGGCTCCCTCCCCTTCTGTTCCTTGGTCT-TTTTTTTTTTTTTTCTTTGCCTTCGTTGCACAAAACTAGCT
    
```

Fig. 26 Alineament de les seqüències nucleotídiques del promotor del *COL1A1* on es troben els SNPs -1997 G/T i -1663 indelT. Els SNPs estan marcats en verd i els nucleòtids conservats en groc. Les seqüències dels oligonucleòtids que es van fer servir en els EMSAs dels treballs de Garcia-Giralt i col·laboradors (2002), i d'Alvarez i col·laboradors (1998) estan subratllades.

■ La transcripció del gen *COL1A1* en les cèl·lules Saos-2 i en les cèl·lules MG-63

Una manera d'estudiar la funció dels polimorfismes associats a un determinat fenotip és per mitjà de cèl·lules en cultiu. Els cultius cel·lulars primaris, tot i que constitueixen un sistema aproximat a la situació que es dona *in vivo*, presenten una sèrie de desavantatges com pot ser la dificultat d'extrapol·lar els resultats degut a l'heterogeneïtat genètica existent entre individus, la baixa eficiència de transfecció o la pèrdua del fenotip amb el temps (Karin i Farach-Carson, 2004). Algunes de les limitacions anteriors es poden solucionar amb l'ús de línies cel·lulars derivades de tumors o transformades. Però aquestes línies, degut al seu origen, presenten altres tipus de limitacions com una elevada i irreal taxa de proliferació i una diferenciació, en molts casos, deficient.

Existeixen diverses línies d'osteoblasts procedents d'osteosarcomes que es poden adquirir comercialment, com per exemple les MG-63 i les Saos-2. Les MG-63, que provenen d'un noi caucàsic de 14 anys, són cèl·lules allargades semblants als fibroblasts. En canvi, les Saos-2, procedents d'una noia caucàsica de 11 anys, tenen una morfologia semblant a la de les cèl·lules epitelials. Ambdós tipus cel·lulars expressen un espectre de citocines i de factors de creixement similar a l'espectre d'expressió dels osteoblasts primaris (Bilbe i col., 1996). Tot i així els elevats nivells d'osteocalcina i d'ALP de les cèl·lules Saos-2 en comparació amb els de les cèl·lules MG-63 i la capacitat de mineralització de les primeres, suggereixen que les cèl·lules Saos-2 presenten un major grau de diferenciació (Rodan i col., 1987; Pautke i col., 2004). És difícil determinar quina d'aquestes línies és més adequada i la seva selecció dependrà bàsicament de l'objectiu concret de cada experiment.

Garcia-Giralt i col·laboradors (2005) per tal de determinar quin paper tenien els polimorfismes del promotor en la regulació de la transcripció del gen *COL1A1* van estudiar en primer lloc el comportament d'aquest promotor en les cèl·lules d'osteosarcoma MG-63 mitjançant l'ús de diferents construccions fusionades a un gen reporter. Van observar que el promotor basal (SP fins a -220 pb) presentava l'expressió màxima, que era el doble de la del promotor total (LP fins a -2.483 pb) i de la del promotor total sense la regió dels polimorfismes (IR). El fet que les construccions LP i IR mostressin nivells d'expressió similars indicava que si els polimorfismes tenien un paper en la regulació del gen, aquest havia de ser discret. De fet es van trobar diferències significatives, però sempre de petita magnitud, entre les taxes de transcripció dels haplotips com comentarem més endavant en la *Discussió*. Tant en les cèl·lules MG-63 com en altres tipus cel·lulars es va descriure una regió entre les posicions -1.284 pb i -254 pb del promotor del gen *COL1A1* que actuava com a repressora (Jimenez i col., 1994; Buttner i col., 2004).

Com ja s'ha dit, les cèl·lules MG-63 i Saos-2 assoleixen un diferent grau de diferenciació. Per comprovar que el comportament transcripcional del promotor del *COL1A1* descrit en les cèl·lules MG-63 no era específic d'aquest tipus cel·lular, vam replicar els experiments en la línia cel·lular Saos-2. Els resultats van mostrar uns nivells de transcripció equivalents entre les diferents construccions del promotor del *COL1A1* en ambdós tipus cel·lulars (*Resultats 2. Estudis funcionals*).

■ El comportament del promotor del gen *COL1A1* en funció del tipus de transfecció

CIZ/NMP4, un factor de transcripció arquitectònic present en la matriu nuclear dels osteoblasts i que regula la transcripció del gen *COL1A1* en rata (Alvarez i col., 1997; Thunyakitpisal i col., 2001), va ser identificat en el complex proteic que s'uneix a la seqüència nucleotídica humana on es troba el SNP -1663 indelT. Els factors arquitectònics són proteïnes encarregades de modificar el plegament del DNA i així posar en contacte els factors de transcripció amb la resta de la maquinària transcripcional. Aquests factors arquitectònics regulen la transcripció en resposta a l'estat de la cèl·lula per mitjà de la seva connexió amb la matriu tissular (Bidwell i col., 2001; Pavalko i col., 2003). L'empaquetament del DNA és, doncs, tant important com la seqüència primària pel que fa a la regulació gènica (Buttner i col., 2004).

Fa temps que es va descriure que els DNAs de plàsmids transfectats transitòriament dins de cèl·lules de mamífer adquireixen una estructura semblant a la de la cromatina a través de la formació de nucleosomes (Reeves i col., 1985). Tot i això i degut a la importància que sembla tenir l'estructura del DNA en la regulació gènica, en el present treball es va voler comprovar si les diferents construccions del promotor del gen *COL1A1* humà mostraven nivells de transcripció equivalents tant si es trobaven en forma de vector episomal com integrats en el genoma.

En la línia cel·lular Saos-2 es va obtenir un patró de transcripció equivalent per ambdós tipus de transfecció (*Resultats 2. Estudis funcionals; Fig. 1*). Aquest fet suggereix que o bé el vector episomal transfectat transitòriament adopta un grau d'empaquetament semblant al dels cromosomes, o que l'empaquetament encara que sigui diferent no intervé en la modulació transcripcional en aquestes condicions experimentals. Per això, la resta d'experiments es van realitzar mitjançant transfeccions transitòries, metodològicament més senzilles.

■ L'efecte de la BMP2 en la regulació del promotor del gen *COL1A1* humà

Una de les funcions de la proteïna CIZ/NMP4 és actuar com a inhibidora del senyal de BMP2 (Shen i col., 2002; Morinobu i col., 2005). Per aquest motiu es va voler estudiar l'efecte tant de BMP2 com de CIZ/NMP4 sobre el promotor del gen *COL1A1*.

La proteïna BMP2 és un factor de creixement de la superfamília TGF β que participa en l'embriogènesi òssia i en la formació d'os nou en l'adult tot promovent la diferenciació de les cèl·lules mesenquimàtiques cap al llinatge osteoblàstic (Fig. 4). Diversos estudis han mostrat que la BMP2 estimula l'expressió del col·lagen (Lecanda i col., 1997; Nissinen i col., 1997; Shen i col., 2002), tot i que altres no han estat capaços de replicar-ho (Li i col., 2004). En concret Nissinen i col·laboradors (1997), van descriure que, en les Saos-2, els nivells del mRNA del *COL1A1* augmentaven després del tractament amb 100 ng/ml de BMP2 en medi sense sèrum durant 24 h.

En el present treball, després de provar diferents concentracions de BMP2 i diferents temps, vam sotmetre les cèl·lules Saos-2 transfectades amb les diferents construccions del promotor del *COL1A1* a un tractament de 2 nM de BMP2 (aproximadament 50 ng/ml) durant 24h. Després d'aplicar el test estadístic de Wilcoxon, vam observar que totes les construccions analitzades eren capaces de respondre al tractament (*Resultats 2. Estudis funcionals; Fig. 2*). El

promotor basal (SP) era inhibít, mentre que la resta de construccions (LP, IR i IRD) eren estimulades. Els nivells més elevats d'estimulació (aproximadament un 60%) s'obtenien amb les construccions LP i IRD, les quals contenen el fragment delimitat entre -2.178 pb i -1.634 pb, on es troben els SNPs.

En valorar si els haplotips participaven en l'efecte estimulador de la BMP2, vam observar que, tot i trobar diferències significatives entre ells en el percentatge d'estimulació, aquestes diferències eren de petita magnitud, no podent-se descartar que fossin causades per un artefacte experimental (*Resultats 2. Estudis funcionals; Fig. 3B*). Per aquest motiu vam decidir buscar en el fragment altres seqüències de resposta a BMP2 que poguessin estar mitjançant l'efecte.

El senyal de la BMP2 es transmet a través de la via de les proteïnes Smad (1, 5 i 8), de la via de les MAPK o bé a través de molècules mediadores (Liberatore i col., 2002; Chen i col., 2004; Miyazono i col., 2005); (Fig. 18). Utilitzant el programa informàtic Genomatix vam realitzar una búsqueda de possibles caixes d'unió de factors de transcripció que pertanyessin a alguna d'aquestes vies i vam identificar-ne una per a Dlx5. Dlx5 és un factor de transcripció que és coexpressat conjuntament amb BMP2 (Ryoo i col., 2006). Diverses evidències han mostrat que Dlx5 transdueix el senyal de BMP2 i promou els últims estadis de diferenciació osteoblàstica. D'altra banda, vam adonar-nos que en aquest fragment també hi havia una seqüència entre la posició -1.698 pb i la posició -1.692 pb (5'-GCCGaCG-3') idèntica a la seqüència del promotor del gen *Msx1* de ratolí on estava descrit que s'unia la proteïna Smad1 (5'-GCCGnCG-3') (Alvarez Martinez i col., 2002).

Malgrat la similitud d'aquestes dues seqüències amb les respectives seqüències consens, cap d'elles per separat va resultar ser responsable de l'augment de la transcripció del gen *COL1A1* en resposta a BMP2 en les cèl·lules Saos-2 (*Resultats 2. Estudis funcionals; Fig. 4*). L'afinitat d'unió de les Smads al DNA és baixa. Moltes vegades la regulació transcripcional per part de les Smads depèn de la seva capacitat d'interacció amb determinats factors de transcripció (Feng i Derynck, 2005; Miyazono i col., 2005). També se sap que la proteïna Dlx5, en resposta a la BMP2, estimula l'expressió dels factors de transcripció Runx2 i Osterix (Ryoo i col., 2006). Per tot això la via de senyalització de la BMP2 consta de múltiples ramificacions. En aquesta primera aproximació hem descartat la participació de manera separada de les caixes de resposta a BMP2 més òbvies. En el futur, seria interessant determinar, a través de la mutagenització doble de les caixes Dlx5 i Smad1, si existeix certa redundància entre elles que faci que no es vegi cap disminució de l'estimulació per part de BMP2 quan són mutagenitzades per separat. Si no fos així, per tal de facilitar la búsqueda de la seqüència de resposta a BMP2, seria útil delimitar el fragment de 544 pb que resulta estimulat mitjançant deleccions seriades.

■L'efecte de la proteïna CIZ/NMP4 en la regulació del promotor del gen *COL1A1* humà

En rata, s'ha descrit que CIZ/NMP4 s'expressa de manera ubíqua tot i que és més abundant en testicle, cor, fetge, cervell i ronyó (Nakamoto i col., 2000; Thunyakitpaisal i col., 2001), però fins al moment no es coneixia el patró d'expressió en humans.

En el present treball vam observar que CIZ/NMP4 s'expressava en tots els teixits humans analitzats (pàncrees, ronyó, múscul esquelètic, fetge, pulmó, placenta, cervell i cor), en osteoblasts humans primaris i en les línies d'osteosarcoma humanes MG-63 i Saos-2.

En totes aquestes mostres es va observar l'expressió simultània de diversos transcrits, podent-se'n identificar fins a quatre (11H, 21H, 21H-I1 i CIZ6.3) (*Resultats 2. Estudis funcionals; Fig. 5*); (*Fig. 27*). Tot i que la RT-PCR no era quantitativa, les isoformes 21H i 21H-I1 semblaven majoritàries en les cèl·lules Saos-2 i MG-63 i van ser clonades en un vector d'expressió. Les isoformes 21H i 21H-I1 es caracteritzen per la falta de l'insert 2 (I2), codificat per l'exó 6, i per la falta dels dits de zinc 4 i 5 (4+5 ZNF), codificats per l'exó 8a. Entre elles, es diferencien perquè a la isoforma 21H-I1 li manca l'insert 1 (I1), codificat per l'exó 4. En osteoblasts de rata, les principals isoformes detectades van ser la 10H, 11H, CIZ6.1 i 21H (Torrunguang i col., 2002b). La isoforma 10H no era present en els teixits humans analitzats en una quantitat suficient com per poder ser detectada a través de la RT-PCR. Finalment la isoforma 22H, que és minoritària en osteoblasts de rata, podria equivaldre, donada la seva mida, a la isoforma 21H-I1 humana.

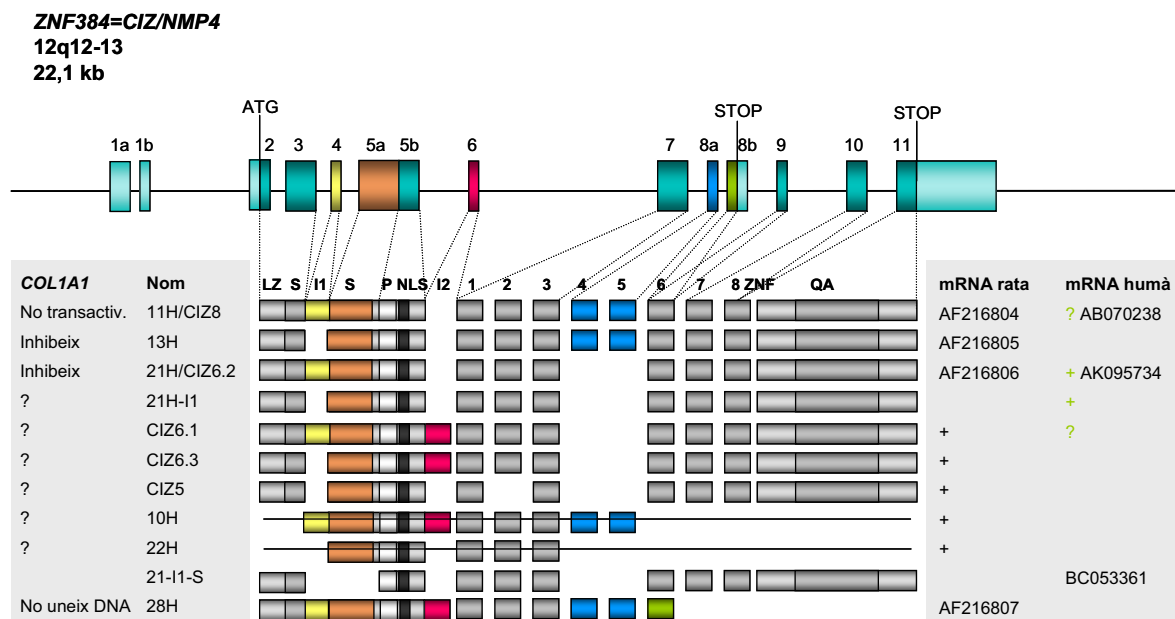


Fig. 27 Estructura del gen *CIZ/NMP4* i de les diverses isoformes descrites (Torrunguang i col., 2002b). El gen consta de 14 exons, el codó d'inici de la traducció es troba en l'exó 2 i el codó STOP en l'exó 11. Es coneix l'existència d'un codó STOP alternatiu en l'exó 8b, el qual s'utilitza en el transcrit 28H. Els exons 4, 5a, 6, 8a i 8b poden eliminar-se per *splicing* alternatiu donant lloc a les isoformes esquematitzades. A l'esquerra de cada isoforma s'indica el seu nom segons la nomenclatura dels grups que les han identificat (Nakamoto i col., 2000; Torrunguang i col., 2002b) (*Resultats 2. Estudis funcionals*), així com també la funció que s'ha descrit que tenen en la regulació del gen *COL1A1* (Thunyakitpisal i col., 2001). A la dreta es mostren les isoformes conegudes en rata i humà. Les isoformes presents en les bases de dades s'indiquen a través de les seves referències, mentre que les isoformes no presents en les bases de dades, però identificades en diferents estudis a través de seqüenciació o a través de la mida del producte de PCR s'indiquen amb els símbols "+" i "?", respectivament. La línia horitzontal negra indica que la seqüència de les isoformes no és del tot coneguda. LZ: cremallera de leucines; I1: insert 1; I2: insert 2; S: seqüència rica en serines; P: seqüència rica en prolines; NLS: senyal de localització nuclear; 1-8 ZNF: dits de Zn del tipus Krüppel; QA: zona rica en alanines i glutamines.

S'ha descrit, a través de la deleció de diferents dominis de la proteïna, que els dits de zinc 1-5 o 4-8 són necessaris i suficients per a una localització exclusivament nuclear, i que la unió de CIZ/NMP4 al DNA té lloc a través del solc menor en zones riques en ATs gràcies a, com a mínim, els dits de zinc 2, 3 i 6 (Feister i col., 2000; Torrungruang i col., 2002a). En el context de la proteïna total, la transactivació és mitjançada per la regió rica en glutamines i alanines (QA) de l'extrem C-terminal. Mentre que les isoformes 11H, 13H i 21H són capaces d'unir-se al DNA, la isoforma 28H no (Thunyakitpisal i col., 2001).

La localització de les isoformes 11H, 13H, 21H i 28H de rata és majoritàriament nuclear (Torrungruang i col., 2002b). Però mentre que les isoformes 11H i 13H presenten un patró nuclear difús (excloent el nuclèol), la isoforma 21H es troba en dos *foci* concrets del nucli, i ella o bé la isoforma CIZ6.1 semblen ser majoritàries en la fracció proteica de la matriu nuclear i en les mitocondries (Feister i col., 2000; Torrungruang i col., 2002b). Coincidint amb això, les isoformes 21H i 21H-I1 humanes transfectades de manera estable en les cèl·lules Saos-2 es van localitzar en la fracció de proteïnes nuclears i no en la de proteïnes citoplasmàtiques (*Resultats 2. Estudis funcionals; Fig. 6A*).

Els ratolins als quals se'ls ha deletat *Ciz/Nmp4* presenten una major massa òssia, una major taxa de formació d'os nou en resposta a la BMP2 i una major expressió de col·lagen (marcador de diferenciació osteoblàstica). *In vivo*, la proteïna CIZ/NMP4 inhibeix la formació òssia a través del bloqueig de l'efecte diferenciador de la BMP2 (Morinobu i col., 2005). *In vitro*, s'ha vist que CIZ/NMP4 estimula l'expressió del gen *MMP13* (Nakamoto i col., 2000; Shah i col., 2004; Fan i col., 2006), però existeix certa controvèrsia sobre l'efecte que té en la regulació del gen *COL1A1*.

En el present estudi, vam observar que la proteïna CIZ/NMP4 no inhibia la transcripció del gen *COL1A1* després del tractament amb BMP2, sinó que per contra, en absència de BMP2, tant la isoforma 21H com la 21H-I1 n'estimulaven lleugerament la transcripció (*Resultats 2. Estudis funcionals; Fig. 6B*). Encara que aquesta estimulació era estadísticament significativa, la magnitud del seu efecte era molt petit i per això el resultat s'ha d'analitzar amb precaució.

Alguns estudis *in vitro* també han mostrat un efecte estimulador de CIZ/NMP4 sobre el gen *COL1A1*. En concret, Furuya i col·laboradors (2000) van descriure que la sobreexpressió de CIZ/NMP4 estimulava l'expressió del col·lagen en les cèl·lules osteoblàstiques de ratolí MC3T3-E1. Shen i col·laboradors (2002), també van observar un lleuger augment del mRNA del *COL1A1* en cèl·lules que sobreexpressaven CIZ/NMP4 respecte les controls durant els primers dies de cultiu en absència de BMP2. Quan les cèl·lules control eren tractades amb BMP2 es duplicava aproximadament per deu l'expressió del col·lagen. En canvi, quan les cèl·lules que sobreexpressaven CIZ/NMP4 eren tractades amb BMP2 mantenien uns nivells similars als observats sense tractament. És a dir, CIZ/NMP4 inhibia l'efecte estimulador promogut per BMP2. Finalment, i contràriament als resultats d'aquest estudi, Thunyakitpisal i col·laboradors (2001) van mostrar que les isoformes 13H i 21H, però no les isoformes 11H i 28H, inhibien la transcripció de 3,5 kb del promotor del *Col1a1* en les cèl·lules d'osteosarcoma de rata UMR-106 en absència de BMP2. La inhibició observada depenia de la seqüència nucleotídica rica en timines del promotor

del *Col1a1*, anomenada *site B*, i que en humans és homòloga a la seqüència que conté el polimorfisme -1663 indelT.

Les dades contradictòries obtingudes *in vitro* podrien ser explicades per un efecte dual de la proteïna CIZ/NMP4 sobre el gen *Col1a1*, de manera similar al que s'ha descrit que té sobre el gen *Runx2*. Concretament es va observar que la transcripció de *Runx2* augmentava en presència de la Smad5 o bé de la proteïna CIZ/NMP4, però quan ambdós factors s'afegien simultàniament es produïa una disminució en la taxa de transcripció (Shen i col., 2002).

La no inhibició de la transcripció del gen *COL1A1* per part de CIZ/NMP4 també pot ser explicada per altres motius, que comentarem breument. En primer lloc hem de recordar que el comportament de les cèl·lules no és el mateix en el context de l'organisme (*in vivo*) o en cultiu (*in vitro*). Com ja hem esmentat en la *Introducció*, la regió que regula la transcripció màxima del gen *COL1A1* en l'os intacte es troba situada entre 2,3 kb i 1,7 kb, mentre que la regió important en les cèl·lules en cultiu es troba entre 3,5 kb i 2,8 kb (Krebsbach i col., 1993). Per aquests motius els resultats de la deleció de *Ciz/Nmp4* en ratolí no tenen per què correlacionar amb els resultats obtinguts *in vitro*.

En segon lloc, els estudis publicats en relació a l'efecte de la proteïna CIZ/NMP4 sobre el gen *Col1a1* han estat realitzats en cèl·lules de rata o ratolí, mentre que en el present estudi s'ha treballat amb cèl·lules d'origen humà. Tot i que la proteïna CIZ/NMP4 humana i la de rata són idèntiques en un 96% de la seva seqüència, s'han descrit algunes diferències pel que fa a la seva capacitat d'unir proteïnes (Janssen i Marynen, 2006). En aquest treball només s'ha avaluat l'efecte de les isoformes 21H (transcripcionalment activa en rata) i 21H-I1, però no es pot descartar que en humans isoformes diferents a les analitzades tinguin un efecte inhibitor sobre el gen *COL1A1*.

En tercer lloc, en aquest treball només s'ha sobreexpressat CIZ/NMP4 de manera que si la seva funció depèn de la unió a altres proteïnes per formar complexes que participin en la transcripció, la no sobreexpressió d'aquestes altres proteïnes podria anul·lar-ne l'acció en aquest sistema.

Una altra limitació de l'estudi és l'ús de cèl·lules que expressen CIZ/NMP4 de manera estable. Aquestes cèl·lules es van obtenir com a *pool* de clons, i es van transfectar transitòriament amb el promotor del col·lagen. Degut a que l'eficiència de transfecció de les cèl·lules Saos-2 és baixa i que s'utilitza un *pool* d'estables, no podem saber quin nivell d'expressió de CIZ/NMP4 presenten les cèl·lules finalment transfectades. Una possible solució consistiria en aïllar clons i seleccionar aquells amb una expressió de CIZ/NMP4 elevada. En el present treball vam optar per les transfeccions estables, tot i les limitacions indicades, perquè prèviament havíem provat de realitzar cotransfeccions transitòries de la CIZ/NMP4 i del promotor del *COL1A1* fusionat al gen repoter de la Luciferasa, però sense èxit.

Degut a la no disponibilitat d'un anticòs comercial específic contra CIZ/NMP4 vam decidir introduir l'epítot Myc a l'extrem N-terminal de la proteïna. Malgrat la mida petita d'aquest epítot (15 aminoàcids) no podem descartar que aquest pugui estar interferint en el plegament de la proteïna o en la seva funció.

Finalment, cal dir que la metodologia utilitzada en el present treball, basada en l'estudi de 2,5 kb del promotor del *COL1A1* fusionat a un gen reporter, no permet descartar l'existència de seqüències més a 5' del promotor on CIZ/NMP4 pugui actuar. De fet *in silico* hem trobat tres caixes d'unió a CIZ/NMP4 més a 5' de la regió analitzada. Per això seria interessant comparar els nivells endògens de mRNA del *COL1A1* entre les cèl·lules Saos-2 que sobreexpressen CIZ/NMP4 i les control, tractades o no amb BMP2.

En resum, hem observat que en línies generals les construccions del promotor del gen *COL1A1* humà es transcriuen de manera similar en les cèl·lules MG-63 i en les Saos-2, i en aquest darrer tipus cel·lular, tant si la transfecció és estable com transitòria. Hem observat que la BMP2 estimula la transcripció del *COL1A1* a través de la regió delimitada entre -2.178 pb i -1.634 pb de manera independent de les putatives caixes Dlx5 i Smad1. Finalment, a diferència del que esperàvem, la proteïna CIZ/NMP4 és incapaç d'inhibir la transcripció del gen *COL1A1* en les condicions de l'estudi.

■L'efecte dels haplotips sobre els nivells de transcripció del gen *COL1A1*

Degut a què l'osteoporosi és una malaltia multifactorial i que té una aparició tardana en la vida s'esperaria que en els estudis funcionals, l'efecte dels polimorfismes fos petit. Hem de puntualitzar, però, que la magnitud de l'efecte d'un SNP *in vitro* no té per què correlacionar amb la magnitud de l'efecte observat en el fenotip final (Rebeck i col., 2004). En l'organisme existeixen factors com l'organització tridimensional dels teixits, l'edat o la dieta, els quals no són replicables en els sistemes de cultius cel·lulars.

A la Taula 9 es mostren els resultats obtinguts en els estudis d'associació i funcionals relacionats amb els polimorfismes del promotor i de l'intró 1 del gen *COL1A1* (Mann i col., 2001; Garcia-Giralt i col., 2005; Stewart i col., 2006; Bustamante i col., 2007) (*Resultats 2. Estudis funcionals*). En les cèl·lules MG-63 es va observar que, en el context del promotor total (LP), l'haplotip G-7T conferia nivells de transcripció significativament superiors als dels haplotips T-8T i T-7T (Garcia-Giralt i col., 2005). En les mateixes cèl·lules però en el context del promotor que no conté la regió repressora (IRD), el comportament dels haplotips era diferent. Per contra, en les cèl·lules Saos-2 no es van observar diferències estadísticament significatives entre els nivells de transcripció dels haplotips (*Resultats 2. Estudis funcionals; Fig. 3A*). Per la seva banda l'al·lel "s" del polimorfisme Sp1, el qual degut al LD ha de definir l'haplotip G-7T-T (18%) o T-7T-T (5%), comportava major quantitat de mRNA i de proteïna que l'al·lel "S" [G-8T-G (65%) o T-8T-G (12%)].

Pel que fa a la correlació entre el fenotip i els nivells de transcripció s'observa que l'haplotip G-7T-T, associat a menor DMO tant en població espanyola com britànica, era un dels que va donar majors nivells de transcripció (Garcia-Giralt i col., 2005). Per tant semblaria que, a major transcripció del col·lagen, menor DMO, actuant doncs, segons un mecanisme similar al que ja hem comentat pel polimorfisme Sp1 que comporta un desequilibri entre les cadenes que formen el col·lagen tipus I. D'altra banda l'haplotip T-8T-G, associat a major DMO en població britànica,

Discussió

era el que semblava que responia més a l'estimulació amb BMP2 (*Resultats 2. Estudis funcionals; Fig. 3B*).

Taula 9 Resultats dels estudis d'associació i dels estudis funcionals (nivells d'expressió i de transcripció) dels SNPs i haplotips del promotor i de l'intró 1 del gen *COL1A1*

		Associació		Expressió	Transcripció				
		<i>Bustamante i col. (2007)</i>	<i>Stewart i col. (2006)</i>	<i>Mann i col. (2001) (Osteobl. primaris)</i>	<i>Garcia-Giralt i col. (2005) (MG-63)</i>		<i>Resultats 2. Estudis funcionals (Saos-2)</i>		
SNP	Al·lel	DMO	DMO	mRNA i proteïna*	Al·lel	LP	IRD	LP + BMP2	
-1997	G	-	-		G	ns.	-		
	T	+	+		T	ns.	-		
-1663	8T	ns.	+		8T	ns.	-		
	7T	ns.	-		7T	ns.	ns.		
Sp1	G "S"	ns.	+	- G-8T-G.....65% T-8T-G.....12%					
	T "s"	ns.	-	+ G-7T-T.....18% T-7T-T.....5%					
		Haplotip	DMO	DMO	mRNA i proteïna*	Haplotip	LP	IRD	LP + BMP2
		G-7T-T	-	-	suposadament +	G-7T	+	+	ns. ns.
		T-7T-T	ns.	ns.	suposadament +	T-7T	-	ns.	ns. -
		G-8T-G	ns.	ns.	suposadament -	G-8T	ns.	+	ns. -
		T-8T-G	ns.	+	suposadament -	T-8T	-	-	ns. +

ns.: valor de referència o no significatiu

+: increments significatius

-: disminucions significatives

LP: promotor total; IRD: promotor amb la regió inhibidora delecionada

*En l'experiment dut a terme amb osteoblasts primaris de diversos individus heterozigots, tot i que els autors només es fixen en l'efecte del SNP Sp1, indirectament estan analitzant l'efecte dels haplotips indicats.

En resum, si els SNPs estudiats participen en la regulació de la transcripció de *COL1A1* ho fan d'una manera discreta, depenent de factors específics de cada tipus cel·lular o bé en unes condicions que no s'ha estat capaç de replicar *in vitro*.

■ Polimorfismes en desequilibri de lligament amb els polimorfismes -1997 G/T i -1663 indelT del gen *COL1A1*

La inconsistència entre els nivells de transcripció dels haplotips en les cèl·lules MG-63 i en les cèl·lules Saos-2, així com el petit efecte observat, planteja la possibilitat que existeixin altres polimorfismes en LD, que o bé sols o bé conjuntament amb els SNPs analitzats, siguin els causants del canvi de DMO observat en els estudis d'associació.

Els tres polimorfismes analitzats es troben en un bloc haplotípic que abarca des del polimorfisme -1997 G/T (rs1107946) fins a l'intró 11 (rs2075555) (HapMap Data Rel 20/phaseII Gen 2006); (Fig. 28). A la Taula 10 es presenten els polimorfismes continguts en aquest bloc i

s'anota la seva posició, la seva MAF i els possibles factors de transcripció que s'unirien de manera específica d'al·lel. En aquest bloc haplotípic hi ha cinc polimorfismes, a més dels tres estudiats en el present treball, amb MAFs conegudes i diferents a zero. D'aquests cinc polimorfismes, quatre han estat genotipats pel projecte HapMap. Entre aquests últims, dos presenten una r^2 superior a 0,8 amb el polimorfisme -1997 G/T i un d'aquests, rs2269336, es troba situat en el promotor a la posició -1363 donant lloc a una seqüència consens per a diversos factors de transcripció de manera específica d'al·lel. D'acord amb les prediccions del programa Genomatix, l'al·lel C d'aquest SNP uniria l'heterocomplex format per VDR/RXR (valor de similitud 0,952 i 0,794 pel *core* i la matriu, respectivament). Aquesta unió quedaria anul·lada en presència de l'al·lel G del polimorfisme. L'altre polimorfisme "redundant" amb el -1997 G/T (rs2075555) es troba situat a l'intró 11.

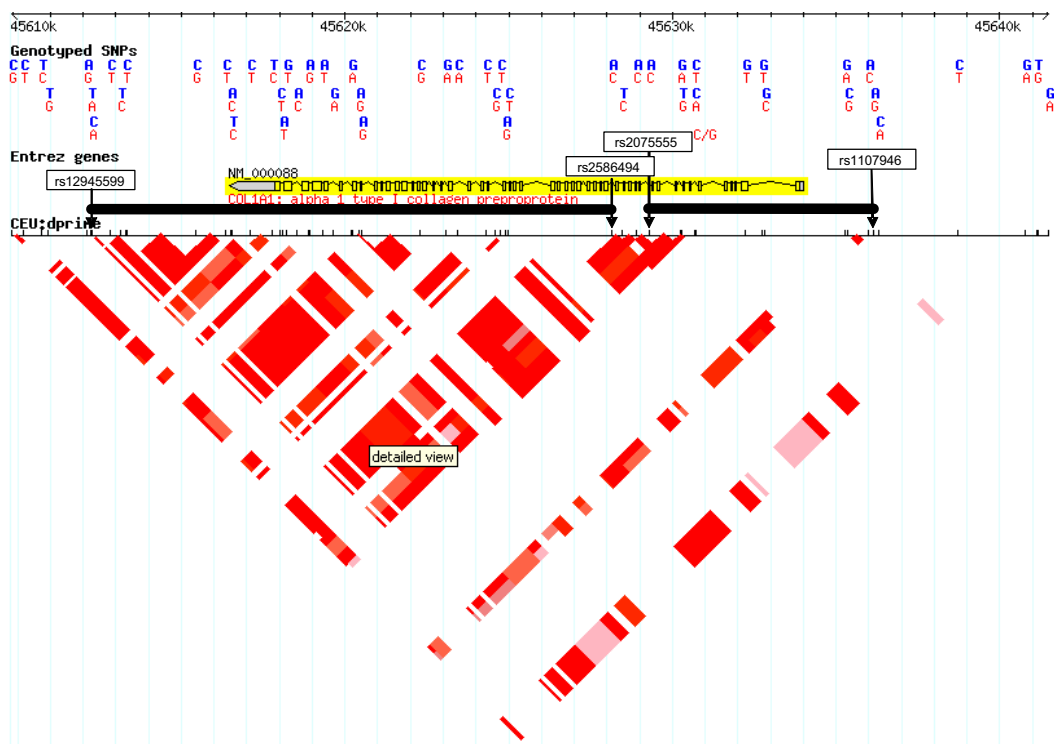


Fig. 28 Blocs haplotípics presents en el gen COL1A1 humà segons les dades del projecte HapMap (Data Rel 20/phaseII Gen 2006) i el mètode de Gabriel et al. S'observen dos blocs haplotípics, un dels quals (rs1107946-rs2075555) abarca des del promotor fins a l'intró 11 i en ell es troben els tres polimorfismes del COL1A1 analitzats en aquesta tesi.

Taula 10 Polimorfismes que es troben en el mateix bloc haplotípic que el polimorfisme -1997 G/T segons les dades de HapMap (Data Rel 20/phasell Gen 06)

	Identificació	Posició	MAF	Polimorfisme	Factors de transcripció Al·lel majoritari	Factors de transcripció Al·lel minoritari
1	<i>rs1107946</i>	<i>promotor</i>	<i>0,183</i>	<i>-1997 G/T</i>	<i>WT, CKROX</i>	<i>GC, MIF1</i>
2	<i>rs11327935</i>	<i>promotor</i>	<i>nd.</i>	<i>-1663 indelT</i>		
3	<i>rs2412298</i>	<i>promotor</i>	<i>nd.</i>	<i>-1663 indelT</i>		
4	rs8075308	promotor	nd.	T/A	MZF1, STAT, TEF	CAAT, RREB1, YY1
5	<i>rs2269336</i>	<i>promotor</i>	<i>0,164</i>	<i>-1363 C/G</i>	<i>VDR/RXR</i>	<i>MEL1, NEUROD1, YY1, ZID</i>
6	rs12603724	promotor	nd.	G/A		HMX2
7	rs2857398	promotor	nd.	A/G	CP2, PPARA	CETS1P54
8	rs2857399	promotor	nd.	T/C	CP2	NRSF, PLAG1
11	rs12601283	promotor	nd.			
12	rs2734278	exó 1	nd.	C/A	NRF1	AHRR, GSH2, PBX
13	rs2586500	intró 1	nd.	C/G	CDE	HELT
14	rs2586499	intró 1	nd.	C/G	CDE	ZF5
15	rs2696258	intró 1	nd.	C/T	ZF5, E2F, NUDR, HIC1	PAX5
16	rs9898186	intró 1	nd.	C/T	ETS1	E2F
17	rs2696257	intró 1	nd.	A/T	MTF-1	MEF3
18	rs2696256	intró 1	nd.	G/A	ATF6	PU1
19	rs2696255	intró 1	nd.	G/C	SP2	EGR1, KKLf, WHN
20	rs2696254	intró 1	nd.	G/A		HBP1
21	rs2696253	intró 1	nd.	G/T	EKLF, OLF1	P53, STAF
22	<i>rs1800012</i>	<i>intró 1</i>	<i>0,188</i>	<i>Sp1</i>		
23	rs2696252	intró 1	nd.		MINI, SP2	
24	rs1057297	exó 2	0.000		R59R	
25	rs2256835	intró 2	0.399			
26	rs2857400	exó 3	nd.	E111E		
27	rs2586498	intró 5	nd.			
28	rs2586497	intró 5	nd.			
29	rs11428831	intró 5	nd.			
30	rs5820805	intró 5	nd.			
31	rs2696251	intró 6	nd.			
32	rs2586496	intró 6	nd.			
33	rs2586495	intró 6	nd.			
34	rs2696250	intró 6	nd.			
35	rs2696249	intró 6	nd.			
36	rs8076979	intró 7	nd.			
37	rs8179178	exó 8	nd.	C197G		
38	<i>rs16948767</i>	<i>intró 8</i>	<i>0,019</i>			
39	<i>rs9913817</i>	<i>intró 8</i>	<i>0,009</i>			
40	rs16948765	intró 9	0.000			
41	rs11657761	intró 9	nd.			
42	rs11079898	intró 9	nd.			
43	rs2075554	intró 11	nd.			
44	<i>rs2075555</i>	<i>intró 11</i>	<i>0,183</i>			
45	rs2696248	intró 11	nd.			
46	rs2734280	intró 11	nd.			

El bloc de desequilibri de lligament està definit segons el mètode de Gabriel i col.

Els polimorfismes en negreta són els genotipats en aquest treball (1, 2=3, 22)

Els polimorfismes en cursiva són els genotipats pel projecte HapMap (1, 5, 38, 39, 44)

Els polimorfismes subratllats tenen una r^2 superior a 0,8 amb el polimorfisme -1997 G/T (5, 44)
nd. equival a MAF no disponible en les bases de dades

3. VALORACIÓ GENERAL DELS ESTUDIS D'ASSOCIACIÓ

3.1. AVANTATGES DELS ESTUDIS D'ASSOCIACIÓ

Tant els estudis de lligament com els d'associació són dues aproximacions vàlides per a l'anàlisi dels factors genètics que participen en la determinació de les malalties complexes. Ambdós mètodes presenten diferents avantatges i limitacions fet que els fa en molts aspectes complementaris. Tot i que no és imprescindible, primer es sol procedir a realitzar estudis de lligament per identificar *loci* lligats al fenotip, i en segon lloc es realitzen estudis d'associació amb polimorfimes situats en els gens candidats posicionals que es troben en les regions de lligament identificades prèviament (Burton i col., 2005; Dawn Teare i Barrett, 2005).

Els estudis de lligament han estat de gran utilitat en l'estudi de les malalties mendelianes però la seva eficàcia en les malalties complexes ha estat menor (Dawn Teare i Barrett, 2005). Els estudis de lligament presenten menys capacitat que els estudis d'associació per detectar factors genètics que tinguin un efecte petit sobre el fenotip.

Els estudis d'associació testen l'efecte de certs polimorfismes sobre un fenotip de manera directa o bé indirectament. Les associacions directes permeten identificar els polimorfismes concrets que causen les variacions fenotípiques. Les associacions indirectes, tot i que estan basades en el desequilibri de lligament, permeten identificar regions cromosòmiques amb més precisió que els estudis de lligament. Això és degut al fet que les regions de desequilibri de lligament són de mida menor que les regions lligades a un fenotip (Dawn Teare i Barrett, 2005). En general, dos *loci* que es trobin en desequilibri de lligament també es trobaran lligats, però dos *loci* lligats no necessàriament estan en desequilibri de lligament.

3.2. LIMITACIONS DELS ESTUDIS D'ASSOCIACIÓ

Tot i els avantatges que hem descrit pels estudis d'associació no cal oblidar que presenten moltes limitacions com queda evidenciat per la falta de replicació dels resultats. Els estudis d'associació són estudis estadístics, i com en qualsevol estudi estadístic, es poden cometre dos tipus d'errors: els errors de tipus I i els errors de tipus II (Ioannidis, 2003). En els primers acceptem com a veritable un resultat que no ho és (fals positiu) i en els segons, perdem l'ocasió d'acceptar com a veritat allò que ho és (fals negatiu). El nivell de significació (α) és la probabilitat de cometre un error de tipus I. En general el nivell de significació es fixa en 0,05, és a dir, de partida acceptem una probabilitat del 5% de descriure un fals positiu. El valor fixat per la probabilitat de cometre un error de tipus II (β), sol ser del 20%, és a dir, acceptem perdre un 20% dels resultats positius.

Entre les problemàtiques dels estudis d'associació destaquen la subestructuració de la mostra, els errors de genotipatge i de diagnosi, la potència estadística i el fet que es realitzen

proves múltiples (multiplicitat de tests) (Newton-Cheh i Hirschhorn, 2005; Palmer i Cardon, 2005; Shen i col., 2005). Tots ells afavoreixen l'aparició d'errors de tipus I o de tipus II.

3.2.1. La subestructuració de la mostra

Un dels problemes principals dels estudis d'associació és la subestructuració de la mostra tant a causa de la barreja d'ètnies com a causa de l'estratificació (Laird i Lange, 2006). La barreja d'ètnies és l'existència en la mostra d'individus procedents de diferents grups ètnics, els quals poden presentar diferent prevalença de la malaltia i/o diferents freqüències al·lèliques per un determinat polimorfisme. L'estratificació és la presència d'aparellament no a l'atzar en la mostra. Malgrat que la subestructuració pot esbiaixar els resultats, es creu que la magnitud del seu impacte no és molt gran (Shen i col., 2005).

Per disminuir la probabilitat de subestructuració de la mostra s'aconsella realitzar estudis d'associació familiars, aparellar cada cas amb un control equivalent, o bé dur a terme el genotipatge de marcadors addicionals que permetin detectar la presència de diferents ètnies i d'aparellament no a l'atzar (Cordell i Clayton, 2005; Hirschhorn, 2005; Palmer i Cardon, 2005). Aquestes mesures correctores moltes vegades són de difícil aplicació. Per exemple el genotipatge de marcadors addicionals sol ser una solució econòmicament limitant. D'altra banda, el reclutament de membres de diferents generacions d'una família per realitzar estudis d'associació familiars sol ser difícil en malalties d'aparició tardana. En el cas de l'osteoporosi, a més, els valors de DMO són de difícil comparació entre individus de diferent edat i sexe.

En aquesta tesi s'ha treballat amb una mostra constituïda per dones postmenopàusiques de l'entorn de Barcelona. A través d'una entrevista es va assegurar que totes elles tinguessin avantpassats espanyols, i a través de la cerca de cognoms repetits es va assegurar l'absència de germanes en la cohort. Aquests criteris de selecció de les participants de l'estudi permeten prevenir, en certa mesura, l'existència de subestructuració.

3.2.2. Els errors de genotipatge i de diagnosi

Els errors de genotipatge i de diagnosi poden afavorir l'extracció de falses conclusions (Hirschhorn, 2005). Per aquest motiu és important regenerar un percentatge de mostres o bé genotipar un grup de mostres controls, el genotip de les quals estigui validat simultàniament per diferents centres (Hattersley i McCarthy, 2005). En el present treball, les tècniques de genotipatge dels polimorfismes situats en els gens candidats clàssics per a l'osteoporosi van ser validades a través d'una placa control analitzada simultàniament en els diversos centres europeus que participaven en el projecte GENOMOS. Després de detectar un 1,7% d'errors en els genotips en la placa control en la cohort BARCOS, es van prendre les mesures oportunes per tal de corregir-ho. El percentatge de mostres no genotipades en la placa control va ser del 1,6%.

Com a mesura per detectar errors de genotipatge s'aconsella comprovar l'equilibri de Hardy-Weinberg (HWE) en el grup d'individus controls de l'estudi d'associació. L'equilibri de Hardy-Weinberg enuncia que sota certes condicions (no mutació, no selecció...), i després d'una generació d'aparellament a l'atzar, les freqüències genotípiques d'un *locus* es fixaran en un determinat valor d'equilibri que dependrà de les freqüències al·lèliques dels progenitors. Els errors de genotipatge no són la única variable que pot trencar l'equilibri de Hardy-Weinberg. La falta d'equilibri també pot ser causada per la deriva genètica, la selecció, l'atzar o la subestructuració de la població (Trikalinos i col., 2006). En el present treball, tots els polimorfismes analitzats es trobaven en equilibri de Hardy-Weinberg excepte la variant -800 G/A situada en el promotor del gen *TGFB1* en la mostra de dones postmenopàusiques espanyoles (Bustamante i col., 2007). Tot i això, aquest polimorfisme va assolir l'equilibri de Hardy-Weinberg quan es va augmentar en aproximadament 200 el nombre de mostres genotipades en motiu de la participació al projecte GENOMOS (*Resultats 1.1.2. GENOMOS: TGFB1*).

En els estudis d'associació és important definir un protocol de diagnosi ben establert. En el cas de l'osteoporosi, es considera que la DMO es una variable quantitativa que es pot mesurar de forma estandarditzada de manera que el seu ús és generalitzat i indiscutible. En canvi, la diagnosi de les fractures osteoporòtiques presenta més dificultat d'estandardització ja que moltes fractures lumbars, en ser indolores, poden passar desapercebudes. Pot ser per això que en els estudis de GENOMOS les associacions positives siguin majoritàriament entre els SNPs i el subgrup de fractures vertebrals incidents, les quals pel fet de ser incidents estan més ben diagnosticades.

3.2.3. La potència estadística

La potència estadística és un paràmetre clau en el disseny dels estudis d'associació. La potència estadística es defineix com la probabilitat d'observar en la mostra un efecte que existeix en la població, és a dir, és la probabilitat que no es doni un error de tipus II (fals negatiu) i val $1-\beta$. Com hem comentat anteriorment el valor de β es sol fixar en un 20% i per tant la potència estadística es sol fixar en un 80%. Aquesta depèn bàsicament de la mida mostral, de la magnitud de l'efecte que produeix el genotip sobre el fenotip, del tipus de model genètic, del desequilibri de lligament entre el marcador genotipat i la variant funcional i de les freqüències al·lèliques d'ambdós polimorfismes (marcador i funcional) (Zondervan i Cardon, 2004; Shen i col., 2005).

Una de les variables més importants per determinar la potència estadística d'un estudi d'associació és l'efecte que se suposa que té el polimorfisme sobre el fenotip. Existeixen diverses teories pel que fa l'efecte dels polimorfismes causals de les malalties complexes. La primera teoria proposa que les malalties comunes són degudes a polimorfismes comuns en la població, els quals produeixen efectes petits sobre el fenotip. La segona teoria apunta que els polimorfismes causants de les malalties comunes són polimorfismes poc comuns amb un efecte important sobre el fenotip. L'últim escenari, i el menys propici, és aquell en que SNPs poc freqüents tenen un efecte de magnitud molt petita sobre el fenotip.

Pel que fa a l'osteoporosi, les anàlisis de segregació estimen que hi ha un gran nombre de gens implicats en la malaltia, majoritàriament amb un efecte petit (Brown, 2005). En aquesta direcció van els resultats obtinguts en els estudis d'associació relacionats amb la DMO. En ells s'ha observat que les variants genètiques expliquen cada una només una petita part de la variació fenotípica observada en la mostra (<5%) (Hirschhorn i Daly, 2005; Shen i col., 2005).

Per tant, com que majoritàriament en les malalties complexes esperem que els polimorfismes tinguin un efecte molt petit sobre el fenotip, la potència estadística tendirà a ser baixa a no ser que s'augmenti considerablement la mida mostral. En la Figura 29 es pot observar l'evolució de la potència estadística en funció de la mida mostral, segons diferents magnituds de l'efecte del SNP sobre la variació fenotípica (2% i 5%) i segons la MAF del polimorfisme causant (0,2 i 0,05). S'assoleix una potència del 80% a partir de 500 mostres si la magnitud de l'efecte és relativament gran (5%) i la MAF de l'al·lel causant és de 0,2. Per tenir la mateixa potència estadística i detectar efectes del 2% amb una MAF del 0,05 cal augmentar la mostra fins aproximadament 1.200 individus. En resum, la potència estadística augmenta en augmentar la mida mostral, la magnitud de l'efecte i la MAF del polimorfisme causal. Alhora, l'efecte de la MAF sobre la potència estadística depèn del tipus de model genètic plantejat.

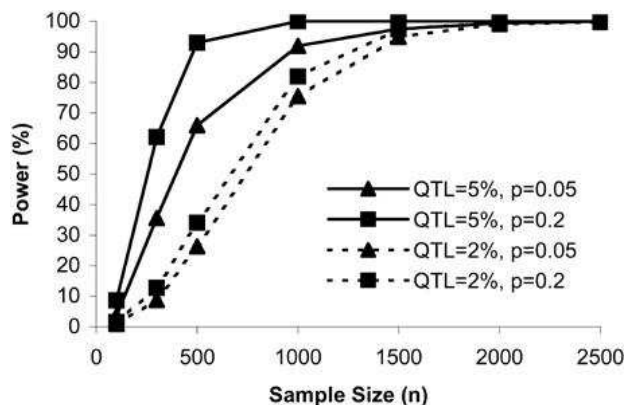


Fig. 29 Variació de la potència estadística en funció de la mida mostral per a diferents MAFs (p) i diferents magnituds de l'efecte (QTL) (extret de Shen i col., 2005).

En els estudis d'associació indirecta, com major és el desequilibri de lligament existent entre el polimorfisme causal i el marcador analitzat més elevada és la potència estadística. Però no és suficient que el valor de D' sigui proper a 1, sinó que també convé que el valor de r^2 sigui proper a 1, és a dir que les MAFs d'ambdós polimorfismes siguin similars (Shen i col., 2005).

Quan el desequilibri de lligament és baix existeixen molts haplotips en la mostra, alguns dels quals tenen freqüències poblacionals molt baixes. Per fer els càlculs estadístics es tendeix a agrupar els haplotips estranys en una sola categoria per així augmentar la potència estadística. Alguns autors, però han demostrat que en determinades malalties són els haplotips estranys els que es troben associats al fenotip i que per tant la pràctica esmentada no és útil (Liu i col., 2005). La potència estadística també disminueix en el moment en que es subdivideix la mostra en grups de menor mida per tal d'analitzar interaccions d'epístasi o interaccions entre els gens i l'ambient.

La potència estadística de la cohort BARCOS és adient per detectar efectes moderats causats pels mateixos marcadors analitzats o per polimorfismes que es trobin en fort desequilibri

de lligament amb ells (r^2 similar a 1). El present estudi no té suficient potència per detectar efectes de petita magnitud ni efectes causats per la interacció entre polimorfismes, és per això que les interaccions que s'han descrit no s'han tractat com a resultats sinó com a punts de partida per generar hipòtesis.

Una de les solucions per augmentar la potència estadística és incrementar la mostra estudiada. En la Figura 30 es representen 226 estudis d'associació relacionats amb fenotips osteoporòtics duts a terme fins l'any 2002. Els estudis s'agrupen en funció del nombre de mostres analitzades en ells (Shen i col., 2005). Aproximadament el 50% d'aquests estudis han estat realitzats amb mostres de menys de 200 individus, fet que compromet la potència estadística.

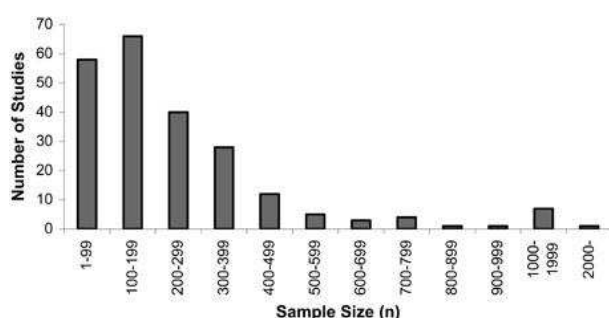


Fig. 30 Representació del nombre d'estudis d'associació amb fenotips osteoporòtics realitzats fins l'any 2002 en funció del nombre de mostres analitzades (extret de Shen i col., 2005).

Les metaanàlisis, que consisteixen en analitzar conjuntament les dades obtingudes per diversos grups d'investigació, permeten solucionar el problema de la potència estadística (Salanti i col., 2005). Hi ha dos tipus de metaanàlisis, les prospectives, on abans d'obtenir les dades de l'estudi, els centres estandarditzen els paràmetres, i les retrospectives, on s'analitzen les dades ja publicades sense prèvia coordinació entre els centres. Les metaanàlisis han mostrat que, en general, els efectes dels polimorfismes han estat sobreestimats en les publicacions que els descriuen per primera vegada (Khoury i col., 2006).

Els principals problemes metodològics de les metaanàlisis són el possible biaix de les publicacions i l'heterogeneïtat. Per solucionar el problema del biaix causat per les publicacions positives és necessari que els treballs amb resultats negatius també siguin publicats (Manly, 2005). En les metaanàlisis s'analitzen conjuntament mostres procedents de diferents poblacions, fet que pot diluir l'efecte d'un polimorfisme que només es manifesti en una població concreta on hi hagi un fons genètic o ambiental determinat. Per tal d'evitar això les metaanàlisis testen l'heterogeneïtat existent entre les mostres estudiades. Tot i els avantatges que presenten, hi ha qui dubta que aquestes realment solucionin la inconsistència dels resultats dels estudis d'associació ja que també s'han descrit resultats contradictoris provinents de diverses metaanàlisis.

La comparació dels resultats obtinguts en aquest treball (aproximadament 500 mostres) amb els resultats de les metaanàlisis del projecte GENOMOS (aproximadament 20.000 mostres) es mostra a la Taula 11. Mentre que en el projecte GENOMOS es va trobar que el polimorfisme Sp1 (*COL1A1*) estava associat a la DMO tant femoral com lumbar, nosaltres no vam detectar-ho possiblement per falta de potència estadística. En canvi, vam observar associat a la DMO lumbar

Discussió

el polimorfisme -1997 G/T (*COL1A1*), que es troba en LD amb el SNP Sp1. D'altra banda, tot i que en un primer moment vam detectar una associació entre el SNP Leu10Pro (*TGFB1*) i la DMO en la cohort BARCOS, aquesta va desaparèixer en ampliar la mostra. En el projecte GENOMOS aquest SNP no es va trobar associat ni a la DMO ni al risc de patir fractures. Finalment en ambdós estudis es va detectar una tendència, en la cohort BARCOS en alguns casos significativa, entre la DMO i els polimorfismes del gen *ESR1*.

En la cohort BARCOS l'efecte dels polimorfismes o haplotips va comportar, de mitjana, un canvi en la DMO de 0,04 g/cm². En el projecte GENOMOS els efectes detectats en relació a la DMO van ser una mica menors, aproximadament de 0,02 g/cm², i els OR descrits pel fenotip fractura oscil·laven entre 0,65 i 1,4.

Taula 11 Comparació dels resultats de la present tesi (BARCOS) amb els resultats del projecte GENOMOS

Gen	Polimorfisme	DMO lumbar		DMO femoral		Fractura
		BARCOS	GENOMOS	BARCOS	GENOMOS	GENOMOS
COL1A1						
Intró 1	Sp1	ns.	sig.	ns.	sig.	sig.
VDR						
Promotor	Cdx2	ns.	ns.	ns.	ns.	sig.
Codificant	FokI	ns.	ns.	ns.	ns.	ns.
Intró 8	BsmI	ns.	ns.	ns.	ns.	ns.
Intró 8	ApaI	ns.	ns.	ns.	ns.	ns.
Codificant	TaqI	ns.	ns.	ns.	ns.	ns.
ESR1						
Promotor	(TA) _n	ns.	ns.	ns.	ns.	ns.
Intró 1	PvuII	ns.	ns.	tend.	tend.	ns.
Intró 1	XbaI	ns.	ns.	ns.	tend.	sig.
TGFB1						
Promotor	-800 G/A	ns.	ns.	ns.	ns.	ns.
Promotor	-509 C/T	ns.	sig. (homes)	ns.	ns.	ns.
Codificant	Leu10Pro	sig. (subgrup)	ns.	ns.	ns.	ns.
Codificant	Arg25Pro	ns.	ns.	ns.	ns.	ns.
Codificant	Thr263Ile	ns.	ns.	ns.	ns.	sig. (subgrup)

ns.: no significatiu

sig.: significatiu

tend.: tendència

3.2.4. La multiplicitat de tests

En els estudis d'associació, i sobretot en els desenvolupats al llarg del genoma, es generen gran nombre de tests estadístics (multiplicitat de tests) de manera que la probabilitat de trobar per atzar un fals positiu és elevada. Totes les variables dels estudis d'associació que permetin abolir els biaixos, com són una bona diagnosi o l'absència de subestructures en la mostra, tendiran a disminuir els errors de tipus I.

En aquests moments no existeix acord sobre quina és la millor manera de corregir el valor de significació per tal de disminuir els errors de tipus I, però existeixen vàries propostes que es discuteixen breument a continuació (Newton-Cheh i Hirschhorn, 2005). La correcció de Bonferroni, que consisteix en dividir el valor de significació (0,05) pel nombre de tests independents realitzats, és una de les mesures més conegudes. Però donat que Bonferroni és una mesura excessivament conservadora, s'aconsella no aplicar-la en els estudis d'associació. En cas d'aplicar-se, cal tenir en compte que en els estudis d'associació els tests contrastats no sempre són independents ja que hi ha certa relació entre els fenotips i entre els polimorfismes si aquests es troben en desequilibri de lligament. Per tant s'ha d'estimar el nombre de tests independents segons aquests criteris.

Una altra manera d'ajustar el valor de significació és per mitjà de les permutacions (Hirschhorn i Daly, 2005). El mètode de les permutacions permet trobar un valor de significació experimental a través de permutar a l'atzar les dades i de calcular amb aquestes totes les possibles associacions. Després d'això s'obtenen uns nous valors de significació ajustats i només aquells SNPs que donin un valor igual o inferior a 0,05 seran considerats.

L'índex de falsos descobriments (FDR) és una mesura que fa referència a la versemblança que una dada observada reflexi una troballa deguda a l'atzar (Newton-Cheh i Hirschhorn, 2005). Aquest mètode permet identificar un llindar de significació a partir del qual l'índex de falsos positius queda fixat en un valor definit pels autors.

També s'ha proposat calcular, a *posteriori*, la probabilitat de descriure un fals positiu (FPRP) (Wacholder i col., 2004). Aquesta probabilitat depèn del valor de significació, de la potència estadística, de l'efecte del polimorfisme i de la probabilitat a *priori* de què existeixi associació entre el polimorfisme i la variant. La probabilitat a *priori* fa referència a tota la informació que predisposa a què una associació positiva sigui certa (la importància funcional del gen, associacions descrites prèviament, la probabilitat de funcionalitat d'un polimorfisme...).

Finalment, per confirmar una associació es pot optar per replicar els resultats d'un estudi en mostres independents però amb variables antropomètriques similars (edat, sexe, anys des de la menopausa) on s'analitzi exactament el mateix fenotip, el mateix polimorfisme i la mateixa direcció de l'efecte de l'al·lel (Moore, 2003; Manly, 2005; Newton-Cheh i Hirschhorn, 2005).

Les associacions analitzades en aquesta tesi, s'han plantejat en la majoria dels casos com a rèpliques d'estudis previs, i en els casos en que no ha estat així no s'ha estudiat un nombre molt gran de polimorfismes. Per aquests motius en aquest treball s'ha descartat corregir les associacions a través del mètode de Bonferroni o altres mètodes.

3.2.5. Complexitat, epístasi i interaccions gen-ambient

El disseny dels estudis d'associació pot comportar resultats no replicables, però altres fonts de falta de replicació poden ser els patrons de desequilibri de lligament, l'epístasi o els factors ambientals (Shen i col., 2005). En aquest sentit, variacions en el patró de desequilibri de lligament entre diferents poblacions poden fer que una associació significativa en una població no

ho sigui en una altra (Palmer i Cardon, 2005). Les diferències ambientals poden produir una conseqüència similar. Un exemple n'és la ingesta de calci en relació a l'efecte dels polimorfismes del gen *VDR* sobre la DMO (Macdonald i col., 2006).

Estadísticament, l'epístasi o interacció entre factors, és la no additivitat de dos efectes. Biològicament, l'epístasi s'entén com l'emascament de l'efecte d'un gen per l'efecte d'un altre gen. L'epístasi és una de les claus de l'arquitectura de les malalties complexes ja que els fenotips són el resultat de la interconnexió biològica de xarxes d'interaccions de biomolècules. Tot i que alguns autors defensen que a través de les noves eines estadístiques es podrà valorar l'efecte de l'epístasi en les malalties complexes, altres creuen que encara s'està lluny de poder-se fer (Moore, 2003; Hirschhorn i Daly, 2005).

3.3. NOVES PERSPECTIVES DELS ESTUDIS D'ASSOCIACIÓ

Els estudis d'associació van aparèixer per primera vegada als anys 80 (Newton-Cheh i Hirschhorn, 2005). En general aquests estudis només tenien potència estadística per detectar efectes de magnitud considerable sempre assumint que la variant analitzada era la causal. Moltes de les associacions descrites durant els anys 80 han estat difícils de replicar, i quan s'han replicat, la magnitud de l'efecte detectat ha estat bastant menor que el descrit per primera vegada. Durant els anys 90 es va apostar pels estudis de lligament com a una possible solució als resultats inconsistents dels estudis d'associació. Però els estudis de lligament relacionats amb fenotips complexes no van tenir l'èxit que s'esperava. Després d'això els estudis d'associació van tornar a guanyar terreny. Durant la darrera dècada, la millora de les tècniques de genotipatge, l'elaboració de bases de dades de polimorfismes i el coneixement dels patrons de desequilibri de lligament, ha donat una empenta als estudis d'associació. Actualment s'estan fent una sèrie de propostes encaminades a estandarditzar els estudis d'associació i racionalitzar la informació que se n'obté.

3.3.1. El projecte HapMap

Últimament l'enfocament dels estudis d'associació ha variat a causa de la informació sobre les variants genètiques aportada per bases de dades com la de HapMap. Les dades de HapMap ens permeten conèixer la disposició dels polimorfismes al llarg del genoma, els patrons de LD i el nombre de *tags* necessaris per capturar cert grau de variabilitat genètica (HapMapC., 2005). L'any 2003, el consorci de HapMap havia genotipat més d'un milió de polimorfismes en 269 individus procedents de quatre poblacions: d'Estats Units (CEU), de Nigèria (YRI), de Xina (CHB) i de Japó (JPT).

Segons el projecte HapMap, els polimorfismes es troben espaiats homogèniament excepte en el DNA mitocondrial i en el cromosoma Y. Un 46% dels polimorfismes són poc freqüents ($MAF < 0,05$) i un 9% s'han vist en un sol individu. Les poblacions CEU, CHB i JPT tenen menys

polimorfismes amb MAFs baixes que els YRI, i aquests, són la població genèticament més distant. Els individus d'una mateixa població comparteixen haplotips, però un 10% dels haplotips d'una determinada població també es troba en una altra. La mitjana d'haplotips que hi ha en un bloc és de 4 (CHB+JPT), de 4,66 (CEU) o de 5,6 (YRI). Actualment s'estan realitzant estudis que comparen les dades d'una població concreta amb les dades de les poblacions analitzades en el projecte HapMap. La majoria dels resultats apunten que els *tags* escollits a partir de les dades de HapMap són adients per usar-los en estudis d'associació en altres poblacions (Ribas i col., 2006).

El projecte HapMap ha confirmat que es necessiten pocs polimorfismes *tags* per capturar la major part de la variabilitat genètica. Considerant només polimorfismes comuns, un 20% dels polimorfismes disposa, cada un d'ells de 20 polimorfismes "redundants", un 60% dels polimorfismes en té més de 5, i el 20% dels polimorfismes restants no disposa de cap variant "redundant".

3.3.2. Els estudis d'associació al llarg d'un gen i al llarg del genoma

La informació de HapMap i les millores en les tècniques de genotipatge permeten dissenyar estudis d'associació que cobreixin gran part de la variabilitat polimòrfica d'un gen o bé estudis d'associació al llarg del genoma (GWA). L'abordatge d'aquests tipus d'estudis es pot fer seguint un model basat en el desequilibri de lligament, i per tant usant *tags*, o bé seleccionant polimorfismes potencialment funcionals. Un polimorfisme es pot considerar potencialment funcional quan és un canvi d'aminoàcid o bé quan per la seva situació podria modificar la transcripció, la traducció o l'estabilitat del mRNA. Alguns programes informàtics, a través de l'anàlisi de regions conservades entre espècies i a través de la comparació de regions que se sap que són reguladores, estimen la probabilitat de funcionalitat d'un polimorfisme.

S'ha calculat que el nombre de marcadors que cal genotipar en un estudi d'associació al llarg del genoma basat en *tags* és de 200.000 a 1.000.000, mentre que el nombre necessari per a estudis amb marcadors funcionals és de 50.000 a 100.000 (Hirschhorn i Daly, 2005; Palmer i Cardon, 2005).

3.3.3. La credibilitat

Després de moltes decepcions degudes a la falta de replicació dels estudis d'associació s'ha començat a plantejar que la informació obtinguda en aquests tipus d'estudis ha de ser considerada informació temptativa. Ioannidis (2006) apunta que només una petita part de la informació temptativa assolirà, a través de la replicació, una major taxa de credibilitat, però la major part, aproximadament un 90%, serà refutada amb el temps (Ioannidis, 2006). El mateix autor defensa la validesa dels estudis d'associació, tot i que accepta que, en aquests moments, l'epidemiologia genètica es pot considerar una tècnica en procés de maduració. Altres autors,

Discussió

però, consideren que l'epidemiologia genètica no durà enlloc i que el que cal és centrar-se més en el disseny experimental seleccionant les mostres sota criteris genètics (per exemple poblacions que hagin passat per un coll d'ampolla) (Terwilliger i Weiss, 2003).

Sota la perspectiva d'informació temptativa, es pot establir un gradient de credibilitat dels resultats obtinguts en els estudis d'associació en funció d'una sèrie de criteris (Taula 12). En primer lloc els estudis on es detecta un efecte gran són segurament més creïbles que els estudis que detecten un efecte petit. Tot i això cal remarcar que majoritàriament s'espera que l'efecte dels polimorfismes sobre el fenotip sigui de petita magnitud. En segon lloc, l'existència de rèpliques d'una associació en mostres independents comporta un *plus* de credibilitat. Altres criteris per discernir la credibilitat d'una associació són la protecció de l'estudi enfront a biaixos, la credibilitat biològica (existència d'evidències funcionals) i la rellevància clínica (Ioannidis, 2006).

Taula 12 Credibilitat dels resultats dels estudis d'associació segons la magnitud de l'efecte i el nombre de rèpliques (extret de Ioannidis, 2006)

<i>Magnitud de l'efecte (risc relatiu)</i>	<i>Rèpliques</i>	<i>Credibilitat (%)</i>
Gran (>5)	Cap	10-60
	Limitades	30-80
	Extenses	70-95
Moderat (2-5)	Cap	5-20
	Limitades	10-40
	Extenses	50-90
Petit (1,2-2)	Cap	<5
	Limitades	2-20
	Extenses	10-70
Molt petit (1-1,2)	Cap	<1
	Limitades	1-5
	Extenses	2-30

3.3.4. La coordinació entre centres

Recentment han començat a aparèixer xarxes formades per diversos centres amb l'objectiu de coordinar la investigació en el camp de l'epidemiologia genètica de les malalties complexes (Ioannidis i col., 2005). Amb la creació d'aquestes xarxes es busca estandarditzar els mètodes clínics, moleculars i estadístics. Entre aquestes xarxes es troba el consorci de GENOMOS. En paral·lel s'ha creat una xarxa de xarxes anomenada Human Genome Epidemiology Network (HuGENet), que inclou informació sobre les xarxes d'investigació de diferents malalties. Potser la col·laboració entre centres serà el futur dels estudis d'associació.