

Roca-Ayats N¹, Cozar Morillo M¹, Gerousi M¹, Czwan E², Urreiziti R¹, Martínez-Gil N¹, García-Giralt N³, Mellibovsky L³, Nogués X³, Díez-Pérez A³, Balcells S¹, Grinberg D¹

¹ Departamento de Genética, Microbiología y Estadística - Facultad de Biología - Universidad de Barcelona - CIBERER (Centro de Investigación en Red de Enfermedades Raras) - IBUB (Instituto de Biomedicina de la Universidad de Barcelona) - Barcelona (España)

² Laboratorios Roche Diagnostics Deutschland GmbH - Mannheim (Alemania)

³ Unidad de Investigación de Patología Ósea y Articular (URFOA) - IMIM (Instituto Hospital del Mar de Investigaciones Médicas) - Parque de Salud Mar - RETICEF (Red Temática de Investigación Cooperativa en Envejecimiento y Fragilidad) - Barcelona (España)

Identificación de variantes genéticas asociadas con la densidad mineral ósea (DMO) en el gen *FLJ42280*

DOI: <http://dx.doi.org/10.4321/S1889-836X2017000100005>

Correspondencia: Daniel Grinberg - Departamento de Genética, Microbiología y Estadística - Facultad de Biología - Avda. Diagonal, 643 - 08028 Barcelona (España)
Correo electrónico: dgrinberg@ub.edu

Fecha de recepción: 15/10/2016

Fecha de aceptación: 15/12/2016

Trabajo premiado con una beca de Investigación FEIOMM 2014.

Resumen

FLJ42280 es un posible gen de susceptibilidad a la osteoporosis. Distintos estudios de GWAs han identificado 4 SNPs no-codificantes en este gen que se asocian a la densidad mineral ósea (DMO) y el riesgo de fractura.

Para descubrir la causa de la asociación entre estos SNPs y la osteoporosis, se realizó una búsqueda de variantes genéticas mediante resecuenciación de 28 kb que contienen el gen, en una selección truncada de mujeres con DMO muy baja (n=50) o muy alta (n=50) de la cohorte BARCOS (Barcelona Cohorte Osteoporosis, cohorte de mujeres postmenopáusicas de Barcelona). Las variantes encontradas se filtraron y se analizó su frecuencia en cada grupo. Se analizó el solapamiento de las variantes con elementos funcionales del proyecto ENCODE y también se calculó el desequilibrio de ligamiento entre los SNPs de la región. Finalmente, se hizo un análisis de eQTL de los 4 SNPs no-codificantes respecto a los niveles de expresión de genes cercanos a *FLJ42280* en linfoblastos.

Se seleccionaron 110 variantes. Las diferencias de sus frecuencias entre los dos grupos estuvieron por debajo del poder estadístico del diseño experimental. Sin embargo, 3 variantes solaparon con posibles *enhancers* y una solapó con un *enhancer* activo en osteoblastos (rs4613908). Se observó un fuerte desequilibrio de ligamiento entre los 4 SNPs no-codificantes y el SNP rs4613908, que pertenecen a un bloque que abarca el gen casi por completo. Ninguno de los SNPs no-codificantes mostró asociación con los niveles de expresión de genes cercanos a *FLJ42280*.

En conclusión, el SNP rs4613908 podría estar implicado funcionalmente en la determinación de la DMO. Serán necesarios experimentos concretos para confirmarlo.

Palabras clave: *FLJ42280*, densidad mineral ósea, variantes genéticas, eQTLs, enhancers.

Identification of genetic variants associated with bone mineral density (BMD) in the *FLJ42280* gene

Summary

FLJ42280 is a possible gene for susceptibility to osteoporosis. Different studies of GWAs have identified 4 non-coding SNPs in this gene associated with bone mineral density (BMD) and fracture risk.

In order to ascertain the cause of the association between these SNPs and osteoporosis, we searched for genetic variants by resequencing the 28-kb gene, in a truncated selection of women with very low (n=50) or very high BMD (N=50) of the BARCOS cohort (Barcelona Cohort Osteoporosis, cohort of postmenopausal women in Barcelona). The variants found were filtered and their frequency analyzed in each group.

The overlap of the variants with functional elements of the ENCODE project was calculated. Finally, an eQTL analysis of the 4 SNPs-coding was performed on the expression levels of *FLJ42280* neighbor genes in lymphoblasts.

In all, 110 variants were selected. The differences in their frequencies between the two groups were below the statistical power of the experimental design. However, three variants overlapped with possible enhancers and one overlapped with an active enhancer in osteoblasts (rs4613908). A strong linkage disequilibrium was observed between the 4 non-coding SNPs and the SNP rs4613908, which belong to a block spanning the gene almost completely. None of the non-coding SNPs showed association with the expression levels of *FLJ42280* neighbor genes.

In conclusion, the SNP rs4613908 could be involved functionally in determining BMD. Tangible experiments will be required to confirm this.

Key words: *FLJ42280*, bone mineral density, genetic variants, eQTLs, enhancers.

Introducción

La osteoporosis es una enfermedad compleja caracterizada por baja masa ósea y deterioro de la microarquitectura del tejido óseo que conduce a un mayor riesgo de fractura. Por ejemplo, en EE.UU. se producen 1,5 millones de nuevos casos de fracturas cada año, lo que representa una enorme carga económica para los sistemas de atención de la salud. La osteoporosis se define clínicamente a través de la medición de la densidad mineral ósea (DMO), que sigue siendo el mejor predictor de fractura^{1,2}. Estudios de heredabilidad utilizando gemelos o familias han demostrado que el 50-85% de la variación en la densidad mineral ósea está determinada genéticamente³. Las fracturas osteoporóticas también muestran heredabilidad independiente de la densidad mineral ósea⁴.

Los estudios de asociación en genoma completo (GWAs) han ampliado muchísimo la comprensión de la arquitectura genética de enfermedades comunes y complejas⁵. Esta aproximación genómica está proporcionando información clave sobre los mecanismos de la enfermedad, con perspectivas para el diseño de estrategias más eficaces de evaluación del riesgo de enfermedad y para el desarrollo de nuevas intervenciones terapéuticas⁶. Sin embargo, las variantes genéticas que se identifican en los GWAs se encuentran con frecuencia en las zonas no-codificantes del genoma, cuya posible función es menos conocida y en muchos casos estas señales pueden estar en desequilibrio de ligamiento con variantes causales que no han sido genotipadas. El meta-análisis de GWAs para DMO y fractura osteoporótica de Estrada *et al.*⁷ identificó hasta 56 loci genómicos asociados con la DMO, 14 de los cuales también estaban asociados con las fracturas osteoporóticas. Uno

de los SNPs cuya asociación con ambos fenotipos mostró una significación más sólida (rs4727338) se encuentra en una región intrónica del gen *FLJ42280*, señalándolo como un *locus* de susceptibilidad para la osteoporosis (Figura 1). Otros trabajos de GWA demostraron que otros SNPs intrónicos del mismo gen (rs7781370, rs10429035 y rs4729260) también se asociaban con la DMO^{8,9}. *FLJ42280* es un gen muy poco estudiado, cuya relación con la biología del hueso se desconoce.

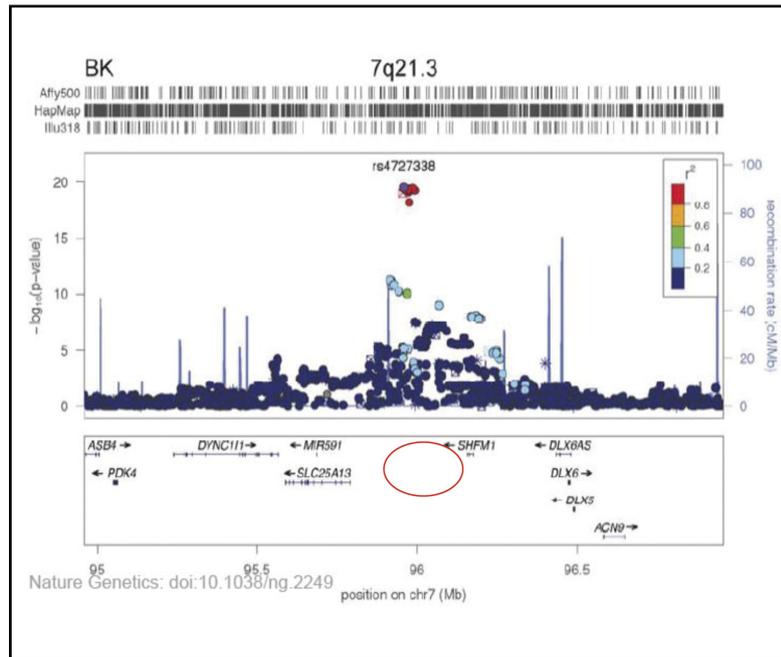
En este contexto, el objetivo de este trabajo fue dar sentido a esta asociación mediante la determinación de cuál es la variante causal. ¿Es rs4727338 el SNP causal o hay otro SNP en desequilibrio de ligamiento con él que sea el verdadero SNP funcional? Para ello, hemos explorado la variabilidad genética de la región genómica donde se encuentra el gen *FLJ42280* y hemos abordado la funcionalidad de estas variantes por enfoques diferentes. En primer lugar, por resecuenciación de la región en mujeres con DMO extremadamente alta o extremadamente baja para buscar variantes con una distribución desequilibrada entre los dos grupos; en segundo lugar, estudio bioinformático de la superposición de las variantes halladas con señales funcionales definidas en el proyecto ENCODE (*The Encyclopedia of DNA Elements*) y finalmente evaluación del posible papel como eQTLs de algunas de las variantes halladas.

Material y métodos

Selección de la muestra de estudio

La muestra de este estudio está formada por 100 mujeres de la cohorte BARCOS¹⁰. Esta cohorte está compuesta por unas 1.500 mujeres postmenopáu-

Figura 1. Región genómica 7q21.3 con las señales de asociación a DMO de los SNPs testados por Estrada *et al.*⁷. La coordenada entre la posición genómica del SNP rs4727338 (eje x) y el $-\log_{10}$ del valor de p (eje y) de su asociación con la DMO (eje y) está marcada con un cuadro de línea roja. Dicho SNP muestra la mayor significación en esta región. Los puntos de colores son las coordenadas del resto de SNPs estudiados en la región. Cada color indica un grado distinto de desequilibrio de ligamiento entre cada SNP y el SNP rs4727338. El gen *FLJ42280* no está mostrado porque, en el momento en que se realizó el meta-análisis de GWAs, este gen todavía no estaba anotado en el genoma. Su localización entre *SLC25A13* y *SHFM1* se indica con un óvalo rojo. Esta figura es una modificación de la que se presenta en Estrada *et al.*⁷



sicas españolas monitorizadas en el Hospital del Mar de Barcelona. Las mujeres diagnosticadas de osteomalacia, de enfermedad de Paget, de algún trastorno metabólico o endocrino, o que estuvieran siguiendo una terapia de sustitución hormonal o tratadas con fármacos que pudieran afectar la masa ósea, fueron excluidas de la cohorte. Las mujeres con una menopausia temprana (antes de los 40 años) también fueron excluidas. La información recogida para cada muestra fue la DMO, la edad, la edad de menarquía, la edad de menopausia, los años desde la menopausia, el peso y la estatura. De cada paciente se obtuvieron muestras de sangre y consentimientos informados escritos, según las regulaciones del Comité Ético de Investigación Clínica del Parque de Salud Mar. La DMO (g/cm^2) fue medida en el cuello del fémur y en la columna lumbar. Se utilizó un densitómetro de rayos X de energía dual para realizar las medidas.

Se seleccionaron dos grupos de 50 muestras con valores de DMO extremos, según el valor del Z -score. Concretamente, los grupos consistieron en las 50 muestras con el Z -score más alto (rango: de 2,98 a 0,73) y las 50 muestras con el Z -score más bajo (rango: de -2,41 a -4,26) de la cohorte BARCOS.

Preparación de las muestras genómicas

El ADN de cada mujer se extrajo a partir de muestras de sangre periférica. La concentración y la calidad de las muestras de ADN (*ratios* 260/280 y 260/230) se midieron por espectrofotometría en un aparato NanoDrop ND-1000 (NanoDrop Products). Para determinar la integridad del ADN, se analizaron 5 μl de cada muestra mediante electroforesis en gel de agarosa al 1%. Finalmente, las muestras se normalizaron a una concentración de 100 $\text{ng}/\mu\text{l}$.

Long-Range PCR (LR-PCR)

Se dividió una región genómica de 28 kb (que contiene el gen *FLJ42280* [22 kb] junto con 3,8 kb de región flanqueante a 5' y 2 kb de región flanqueante a 3') en 7 fragmentos solapantes (Figura 2). Los tamaños y las coordenadas de estos 7 fragmentos y los pares de cebadores utilizados para amplificarlos, se muestran en la tabla 1. Los fragmentos, de entre 2 y 5 kb, se amplificaron mediante LR-PCR. Cada reacción de LR-PCR incluyó: 100 ng de ADN genómico, 5 μl de tampón *Ex Taq* "Magnesium +" (20 mM Mg^{2+} ; Takara) x10, mezcla de dNTPs (a 2,5 mM cada uno), *Ex Taq* polimerasa (5 U/ μl) y cebadores (20 μM), en un volumen final de 50 μl .

Las reacciones se llevaron a cabo en un termociclador GeneAmp[®] PCR System 2700 (Applied Biosystems). Cada fragmento requirió unas condiciones de tiempo de elongación y temperatura de hibridación distintas. El número total de amplificones fue de 700 (100 muestras x 7 fragmentos). Se comprobó la cantidad y calidad de todos los amplificones mediante electroforesis en gel de agarosa al 1% p/v en tampón TBE x1.

Purificación y cuantificación de las muestras

Para eliminar los restos de los reactivos de la PCR, los productos de PCR se purificaron utilizando placas de filtro de 96 pocillos con un tamaño de poro adecuado (Pall Corporation). Se aplicó el vacío (Vacuum Manifold, Merck Millipore) y el ADN retenido en el filtro se resuspendió en 35 μl de agua milliQ. A continuación, los productos de PCR se cuantificaron mediante Quant-iT PicoGreen dsDNA Reagent and Kit (Life Technologies), según las instrucciones del fabricante. Brevemente, se construyó una curva estándar de concentraciones mediante medidas de emisión de fluorescencia a 520 nm después de haber excitado el ADN a 480 nm. La curva se utilizó posteriormente para calcular la concentración del ADN de las muestras.

Normalización de las concentraciones de las muestras y pooling

Las muestras en las placas se normalizaron a una concentración de 5 ng/μl y, a continuación, se mezcló en un solo tubo 5 μl de cada muestra de una placa (un tubo por placa) mediante el The epMotion® 5075 Liquid Handling Workstation (Eppendorf). Así, se obtuvieron 14 tubos con 250 μl cada uno, dos para cada fragmento de PCR (DMO alta y DMO baja). Los 14 pools se concentraron hasta 5 veces utilizando el Genevac EZ-2 evaporator (Genevac SP Scientific) y se cuantificó cada tubo mediante Qubit® 2.0 Fluorometer (Life Technologies). Finalmente, los fragmentos de PCR se mezclaron equimolarmente en dos tubos, uno para la DMO alta y otro para la DMO baja.

Secuenciación masiva en paralelo

La secuenciación masiva de las muestras se llevó a cabo en el Servicio de Genómica de los Centros Científicos y Tecnológicos de la Universidad de Barcelona, utilizando el GS 454 Junior System (Roche). Brevemente, se fragmentó el ADN por nebulización, se prepararon dos librerías marcadas con adaptadores (secuencias de 10 nucleótidos) distintos, una para cada grupo, que se mezclaron en un solo tubo. Seguidamente, se amplificó la mezcla por PCR de emulsión y la librería final se cargó en una placa *picotiter* dónde se llevó a cabo la pirosecuenciación. Se realizaron 4 carreras de secuenciación, correspondientes a 140 Mb de datos finales (35 Mb/carrera). Este volumen de datos proporciona una cobertura teórica de 40x para cada muestra inicial.

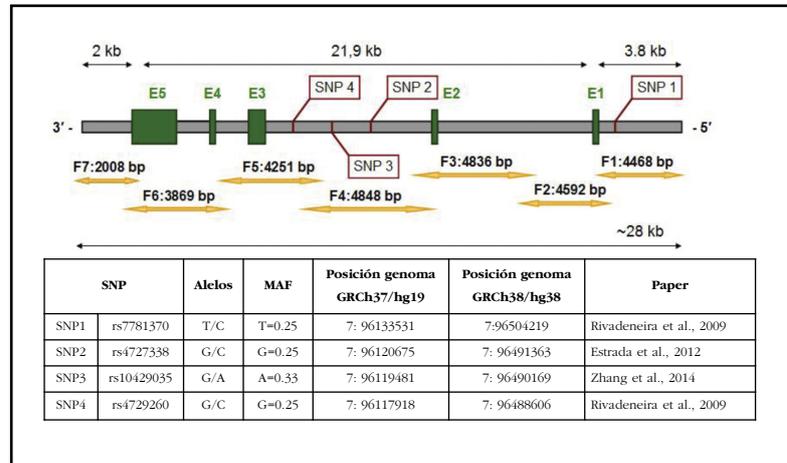
Procesamiento de los datos de secuenciación y selección de variantes

Las lecturas obtenidas de la secuenciación fueron preprocesadas en base a su calidad y se alinearon contra el genoma de referencia (GRCh37) utilizando el programa GS Mapper (Roche). Las lecturas se indexaron y filtraron utilizando SAMtools. Se detectaron las variantes presentes en los dos grupos mediante GATK utilizando parámetros de filtrado estándar¹¹. Las variantes encontradas se priorizaron según los siguientes criterios: se seleccionaron las variantes con una cobertura de al menos 1.000 lecturas, presentes en un 1% de las lecturas y con una *strand bias* bajo. El número de lecturas de las variantes que pasaron los filtros fue normalizado por la cobertura y las variantes fueron clasificadas entre comunes (con una frecuencia mayor al 5%) y raras o de baja frecuencia (con una frecuencia menor al 5%).

Análisis funcional y estadístico de las variantes

Se compararon las frecuencias de cada variante entre los dos grupos mediante un test exacto de

Figura 2. En la región genómica de *FLJ42280* estudios previos de GWA han identificado 4 SNPs que muestran asociación con la DMO, cuyos detalles se muestran en la tabla incluida en la figura. Para conocer mejor la variabilidad de este *locus*, la región se subdividió en 7 fragmentos solapantes (flechas amarillas) para su resecuenciación en mujeres con muy alta o muy baja DMO



Fisher, aplicando la corrección de Bonferroni para comparaciones múltiples. El análisis funcional de las variantes consistió en mirar si estaban descritas en bases de datos como dbSNP y 1000 Genomas y, en caso afirmativo, buscar su MAF en la población europea e ibérica. Además, para las variantes exónicas se observó qué cambio de aminoácido suponían y su severidad predicha por SIFT, PolyPhen y Provean. Para las variantes intrónicas, se analizó la región que contiene la variante: lugares de hipersensibilidad a la DNasa, unión de factores de transcripción, metilación del ADN, modificaciones de histonas, regiones reguladoras, etc. Todos estos datos fueron obtenidos de bases de datos y repositorios como Ensembl, UCSC Genome Browser, ENCODE, BioMart, MatInspector. También se utilizó HaploReg para buscar anotaciones de regulación. Finalmente, todas las variantes encontradas fueron analizadas con el *Variant Effect Predictor* de Ensembl y UCSC y con el *SNP function prediction* del Instituto Nacional de Ciencias de la Salud Ambiental (*National Institute of Environmental Health Sciences*) de los EE.UU.

Análisis del desequilibrio de ligamiento

Para calcular el desequilibrio de ligamiento entre todas las variantes de la región genómica de *FLJ42280* se utilizaron los genotipos de los SNPs presentes en la región y los haplotipos de los individuos de HapMap fase 3. Para calcular tal desequilibrio y generar un gráfico se utilizó el software HaploView.

Análisis de eQTLs

Los SNPs que dieron significativos en los distintos GWAs y el SNP rs4613908 fueron evaluados como posibles eQTLs mediante dos aproximaciones: utilizando el portal del proyecto GTEx y utilizando los genotipos de esos SNPs en individuos de HapMap y los niveles de expresión de genes en *cis* en los

Tabla 1. Amplicones utilizados para secuenciar la región *FLJ42280*

Frag	Primers	Tamaño (pb)	Coordenadas genómicas
1	1F TTGACCTGAATACTGCCGC 1R GCCAAATGAATGTGGACAAG	4.468	7:96136619-96132152
2	2F CACTGCTGGGTCTTAGATTGG 2R GCATGTGTGCATGATGTTGG	4.592	7:96132302-96127711
3	3F TGCAAGTTTCCCTCAATTCATC 3R TCCCTCTCATCTGTGCAACAC	4.836	7:96127863-96123028
4	4F TTAGGTGAGTAGAAAGCAATGGC 4R CTGGGTGGCTATAGACCTGAATAG	4.848	7:96123158-96118311
5	5F GCGGCACTGTGAGAGTACATC 5R CCTGGTGAAATGGGAACA	4.251	7:96118477-96114227
6	6F CTGACACTTTGGCAGCACC 6R GGGATTGTTGAAGCTGACCC	3.869	7:96114348-96110480
7	7F CAACCATCACAACCCATAGAC 7R CCTGAGCAAGTCTCGTAAGTG	2.008	7:96110702-96108695

mismos individuos. Concretamente, se obtuvieron los genotipos de los SNPs de 210 individuos no emparentados de la fase 1 y 2 de HapMap y los niveles de expresión de los genes *SHFM1*, *SLC25A13* y *DLX5* de una línea celular linfoblástoides de los mismos individuos obtenidos.

Resultados

Variantes halladas y pistas sobre su función

La región genómica de *FLJ42280* (28 kb) se secuenció masivamente en dos *pools* de DNA correspondientes a las 50 mujeres con mayor DMO y las 50 mujeres con menor DMO de la cohorte BARCOS (ver detalles en Material y Métodos), a una alta profundidad (alrededor 3.600x por grupo). Se comparó el número y la frecuencia de las variantes que se encontraron en cada grupo. Se identificó un total de 110 variantes, de las cuales 18 eran nuevas y 59 fueron variantes raras o de baja frecuencia (Tabla 2). Se observó que el número de variantes de baja frecuencia entre los dos grupos extremos era equilibrado. Así mismo, se observó que las diferencias de frecuencia de todas las variantes estaban por debajo de la potencia estadística del diseño, aunque 9 mostraron una tendencia.

Para cada variante, se analizó su superposición con elementos funcionales anotados en el genoma por el proyecto ENCODE. Cuatro de las variantes solaparon con posibles secuencias potenciadoras de la transcripción (o *enhancers*) de osteoblastos y de ellas una [SNP rs4613908; MAF(CEU)=0,39] solapó con un enhancer activo en osteoblastos (Figura 3).

Análisis de desequilibrio de ligamiento

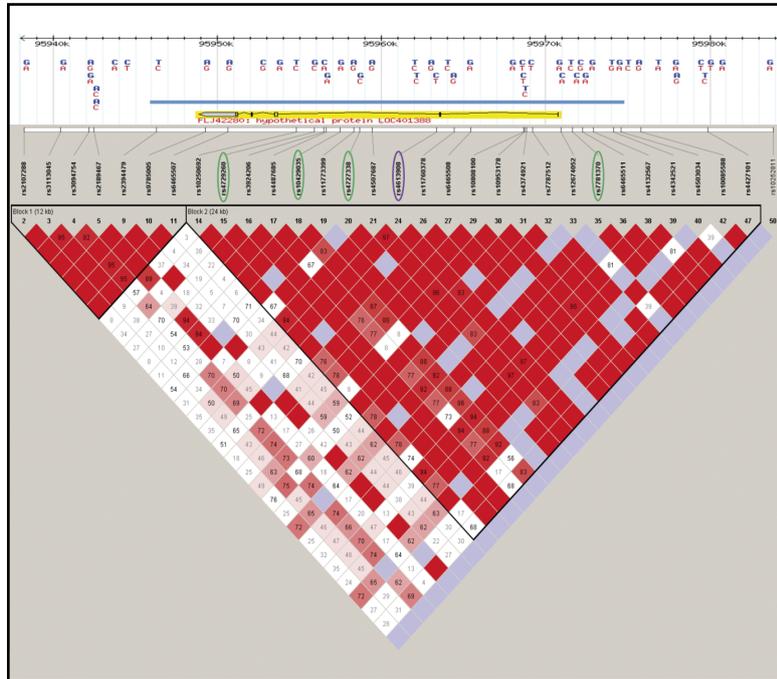
También se estudió el desequilibrio de ligamiento entre todas las variantes comunes en esta región. Se creó un gráfico de desequilibrio de ligamiento (LD)

utilizando HaploView e información de haplotipos del proyecto HapMap (Figura 4) y se observó que hay un gran bloque de LD que incluye casi todo el gen (a excepción de la región 3' UTR) y que por la parte *upstream* del gen se extiende 5 kb más allá de la región resecuenciada. También se constató que los SNPs rs4613908 y rs4727338 (meta-análisis de GWAs de Estrada *et al.*) presentan una gran desequilibrio de ligamiento entre ellos.

Análisis de eQTLs

Para completar el análisis funcional, se realizó un análisis de eQTLs. Disponiendo de los genotipos de los cuatro SNPs asociados a la DMO, y del SNP rs4613908 de 210 individuos del proyecto HapMap y de los niveles de expresión génica de un *array* genómico en líneas linfoblástoides de estos mismos individuos, se determinó si los diferentes alelos o genotipos de los SNPs correlacionaban con los niveles de expresión génica de los genes situados en la región genómica de *FLJ42280*. Ninguno de los SNPs mostró influencia sobre los niveles de expresión de los genes *SHFM1*, *SLC25A13* o *DLX5* (en el *array* no hay información de niveles de expresión de *FLJ42280*). También se accedió a la base de datos GTEx para recabar información de eQTL para los mismos SNPs y el resultado fue negativo para todos ellos. Finalmente, se realizaron búsquedas de anotaciones de regulación en HaploReg. Este último análisis confirmó que la secuencia que rodea el SNP rs4613908 está altamente conservada entre los mamíferos y que en varios tipos celulares, incluyendo los osteoblastos primarios, contiene marcas de cromatina típicas de secuencias potenciadoras (H3K4me1, H3K27ac). Por otra parte, HaploReg destacó la alteración de motivos reguladores de este SNP y de rs10429035, pero no mostró ningún efecto de estos SNPs sobre la expresión génica.

Figura 4. Gráfico del desequilibrio de ligamiento presente en la región genómica de *FLJ42280*. Se observa un bloque grande (24 kb) que incluye gran parte del gen excepto su región 3'UTR. Los 4 SNPs que se han asociado a DMO en los distintos GWAs están señalados con óvalos verdes. El SNP que solapa con el potenciador activo en osteoblastos está señalado con un óvalo morado. La línea horizontal azul indica la región genómica resecuenciada en este estudio



encuentra en uno de estos potenciadores (eDLX#18), situado a unas 500 kb de *DLX5*. El potenciador eDLX#18 se ha descrito como activo en los arcos branquiales en estadios embrionarios¹³.

Existen evidencias de que *DLX5* está involucrado en la determinación de la DMO¹⁵, lo que nos hace proponer la hipótesis de que el potenciador eDLX#18 también es activo como un potenciador para *DLX5* en osteoblastos de adultos y que nuestro SNP de interés es un eQTL en osteoblastos. Será muy necesario comprobar esta hipótesis mediante análisis de expresión de *DLX5* en osteoblastos primarios y genotipación de rs4613908 de los mismos.

Conflicto de intereses: Los autores declaran no tener conflicto de intereses en relación con este trabajo.

Agradecimientos: Este trabajo se ha llevado a cabo gracias a la ayuda económica brindada por una beca FEIOMM de 2014.

Bibliografía

- Johnell O, Kanis JA, Oden A, Johansson H, De Laet C, Delmas P, et al. Predictive Value of BMD for Hip and Other Fractures. *J Bone Miner Res.* 2005;20:1185-94.
- Kanis JA, Oden A, Johnell O, Johansson H, De Laet C, Brown J, et al. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. *Osteoporos Int.* 2007;18:1033-46.
- Peacock M, Turner CH, Econs MJ, Foroud T. Genetics of Osteoporosis. *Endocr Rev.* 2002;23:303-26.
- Deng HW, Chen WM, Recker S, Stegman MR, Li JL, Davies KM, et al. Genetic determination of Colles' fracture and differential bone mass in women with and without Colles' fracture. *J Bone Miner Res.* 2000;15:1243-52.
- Hardy J, Singleton A. Genomewide Association Studies and Human Disease. *N Engl J Med.* 2009;360:1759-68.
- Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. *N Engl J Med.* 2010;363:166-76.
- Estrada K, Styrkarsdóttir U, Evangelou E, Hsu YH, Duncan EL, Ntzani E, et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nat Genet.* 2012;44:491-501.
- Rivadeneira F, Styrkarsdóttir U, Estrada K, Halldórsson BV, Hsu YH, Richards JB, et al. Twenty bone-mineral-density loci identified by large-scale meta-analysis of genome-wide association studies. *Nat Genet.* 2009;41:1199-206.
- Zhang LL, Choi HJ, Estrada K, Leo PJ, Li J, Pei YF, et al. Multistage genome-wide association meta-analyses identified two new loci for bone mineral density. *Hum Mol Genet.* 2014;23:1923-33.
- Bustamante M, Nogués X, Enjuanes A, Elosua R, García-Giralt N, Pérez-Edo L, et al. *COL1A1*, *ESR1*, *VDR* and *TGFB1* polymorphisms and haplotypes in relation to BMD in Spanish postmenopausal women. *Osteoporos Int.* 2007;18:235-43.
- DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43:491-8.
- Shamseldin HE, Faden MA, Alashram W, Alkuraya FS. Identification of a novel *DLX5* mutation in a family with autosomal recessive split hand and foot malformation. *J Med Genet.* 2012;49:16-20.
- Birnbaum RY, Everman DB, Murphy KK, Gurrieri F, Schwartz CE, Ahituv N. Functional characterization of tissue-specific enhancers in the *DLX5/6* locus. *Hum Mol Genet.* 2012;21:4930-8.
- Rasmussen MB, Kreiborg S, Jensen P, Bak M, Mang Y, Lodahl M, et al. Phenotypic subregions within the split-hand/foot malformation 1 locus. *Hum Genet.* 2016;135:345-57.
- Ulsamer A, Ortuño MJ, Ruiz S, Susperregui AR, Osses N, Rosa JL, et al. BMP-2 Induces Osterix Expression through Up-regulation of *Dlx5* and Its Phosphorylation by p38. *J Biol Chem.* 2008;283:3816-26.