

Genes implicados en el desarrollo del esqueleto. Selección adaptativa en regiones reguladoras no-codificantes*

(Genes involved in skeletal development. Adaptive selection in non-coding regulatory regions)

López López, Saioa; Hervella Alfonso, Montserrat; Fontecha Martínez, Lara; Izagirre Arribalzaga, Neskuts; De La Rúa Vaca, Concepción; Alonso Alegre, Santos
Univ. del País Vasco/Euskal Herriko Unib. Fac. de Ciencia y Tecnología. Dpto. Genética, Antropología Física y Fisiología Animal. Sarriena, s/n. 48940 Leioa

BIBLID [1989-2012 (2010), 11; 27-39]

Recep.: 27.05.2009
Acep.: 28.09.2010

Uno de los rasgos que nos identifica como humanos es nuestro esqueleto. En este trabajo buscamos identificar aquellas zonas reguladoras de genes implicadas en la morfología esquelética y que presentan cambios en su secuencia de ADN que son específicos humanos. Con ello perseguimos identificar aquellos factores implicados en las adaptaciones esqueléticas humanas, como la bipedia o morfología craneal.

Palabras Clave: Evolución humana. Esqueleto. ADN no-codificante. Elementos reguladores. Análisis bioinformático.

Gure eskeletoa da gizaki bezala identifikatzen gaituen ezaugarrietariko bat. Lan honetan, eskeletoaren morfologian inplikatur dauden geneen eskualde erregulatzaileak, zeintzuk euren ADN sekuentzian aurkezten dituzten aldaketak gizakiaren berariazkoak diren, identifikatzea bilatzen dugu. Honekin, gizakiaren eskeletoaren moldaketan, hala nola bipedian edo garezurraren egitura, inplikaturako eragileak identifikatu nahi dugu.

Giltza-Hitzak: Gizakiaren eboluzioa. Eskeletoa. ADN ezkodifikagarria. Osagai erregulatzaileak. Análisi bio-informatikoa.

L'un des traits qui nous identifie en tant qu'humains est notre squelette. Dans ce travail nous cherchons à identifier les zones régulatrices des gènes impliquées dans la morphologie squelettique et qui présentent des changements dans leur séquence d'ADN qui sont des spécifiques humains. De cette façon nous cherchons à identifier les facteurs impliqués dans les adaptations squelettiques humaines, comme la bipédie ou morphologie crânienne.

Mots-Clés: Evolution humaine. Squelette. ADN non-codifiant. Éléments régulateurs. Analyse bioinformatique.

* Este trabajo ha contado con una ayuda a la investigación 2008 de Eusko Ikaskuntza.

1. ANTECEDENTES Y ESTADO ACTUAL DEL TEMA

En biología evolutiva humana, una de las áreas de mayor interés es la de identificar si nuestras características fenotípicas son el resultado de la selección natural (Clark; et al., 2003). En nuestra especie en particular, uno de los rasgos que nos identifica como humanos, aparte de las capacidades cognitivas, es la morfología de nuestro esqueleto. La evolución del bipedalismo hace unos 4 millones de años dio lugar a cambios morfológicos importantes en el esqueleto de los humanos, con una tendencia a la gracilización. Estudios comparativos de huesos entre chimpancés, gorilas, babuinos y humanos demuestran que en relación con el tamaño corporal, la fortaleza de los huesos de los brazos y piernas de esos primates es mayor que la de los humanos, lo que les permite también poseer una musculatura más vigorosa (Ruff, 2007).

Esto sugiere la posibilidad de que algunos de los genes implicados en el desarrollo esquelético pudieran haber sufrido ciertas sustituciones nucleotídicas, que supusieron una ventaja adaptativa para *Homo sapiens*. Identificar estos genes (y sus zonas reguladoras) implicados en la adaptación a la bipedia o a nuestra morfología craneal nos permitirá identificar aquellos factores que han posibilitado convertirnos en humanos y ampliará por lo tanto nuestro conocimiento sobre nuestra propia historia evolutiva. Por otra parte, dada la importancia evolutiva de dichos genes y zonas reguladoras, es posible que puedan ser asimismo regiones genómicas candidatas a asociación con enfermedades o malformaciones. Identificarlos puede así abrir una vía hacia la identificación de las mutaciones responsables de dichas enfermedades y, por lo tanto, de su posible tratamiento. Se confirmaría así la célebre frase de Dobzhansky: “Nada en biología tiene sentido si no es a la luz de la evolución”.

Es posible, en principio, tratar de identificar el efecto de la selección natural mediante análisis bio-informáticos de las secuencias nucleotídicas codificantes de los genes (Yang; Bielawski, 2000; Yang; Nielsen, 2002; Zhang; et al., 2005). De hecho, anteriormente hemos analizado por medio de PAML (<http://abacus.gene.ucl.ac.uk/software/paml.html>) la región codificante de 55 genes presuntamente relacionados con el desarrollo del esqueleto humano en busca de señales de esta selección adaptativa, y encontramos dos genes, *COL10A1* y *PRTN3* que mostraron indicios de selección adaptativa en el linaje humano exclusivamente.

El trabajo que ahora presentamos pretende continuar esta idea extendiendo la búsqueda de selección natural mediante métodos bioinformáticos a las zonas reguladoras (no codificantes) de genes candidatos.

El DNA no codificante de las zonas reguladoras es fundamental para determinar cuándo (en qué momento de la ontogenia, bajo qué circunstancias fisiológicas...), cuánto (qué nivel de expresión) y dónde (en qué tejidos sí y en qué tejidos no) se debe expresar un gen. Ya en 1975, King y Wilson (King; Wilson, 1975) propusieron que los cambios (sustituciones nucleotídicas) en elementos reguladores de la expresión génica podrían explicar las marcadas diferencias anatómi-

cas y de comportamiento entre los chimpancés y los humanos, dado que en las regiones codificantes el nivel de divergencia entre los genomas de ambas especies parecía ser “demasiado” bajo. Por otra parte, en muchos casos la divergencia entre las especies está muy ligada a los cambios cuantitativos en la expresión génica. En este sentido, según se ha estimado, el ADN no codificante regulador podría haber sufrido el doble de selección positiva que el codificante (Carroll, 2003).

La idea de que la variación reguladora no-codificante juega un papel determinante en la evolución morfológica está adquiriendo un ímpetu renovado (Carroll, 2005). Los cambios en las zonas reguladoras de los genes presentan un enorme atractivo debido a que, al contrario de lo que sucedería con la variación en las zonas codificantes, la variación en estas zonas solucionaría el problema de la pleiotropía antagonista, es decir, el hecho de que un mismo gen se exprese en diferentes tejidos o en distintos tiempos ontogénicos, lo cual puede tener diferentes intereses evolutivos. Ello permitiría a su vez un incremento de la modularidad del genoma, ya que se permitiría un control independiente de la transcripción génica en, por ejemplo, los diferentes tejidos. Este mecanismo permite asimismo la creación de nuevas combinaciones de redes reguladoras, es decir, la clase de evolución por corta-y-pegar ideada por Jacob (1977). Todo ello no implica que el papel de la variación codificante sea nulo, sino que encajaría mejor para aquellas situaciones de no antagonismo pleiotrópico, como en aquellos genes que se expresan en un solo tejido o tipo celular (ver revisión por Alonso; et al., 2007).

Con este trabajo, pretendemos por lo tanto profundizar en el estudio de adaptaciones esqueléticas a lo largo de nuestra historia evolutiva. La investigación de la relación entre cambios concretos en el genoma y los cambios morfológicos del esqueleto supone no sólo aportar conocimiento sobre nuestra historia evolutiva y de los rasgos que nos han hecho humanos, sino que por otra parte, la identificación de genes de especial relevancia evolutiva, puede asimismo proporcionarnos pistas sobre la importancia biomédica de los mismos.

2. METODOLOGÍA

La identificación de posible selección adaptativa en las regiones no codificantes reguladoras ha sido llevada a cabo siguiendo el siguiente esquema:

2.1. Identificación de genes candidatos, implicados en el desarrollo esquelético

Este punto se ha realizado buscando en la literatura la asociación entre palabras clave como “morfogénesis del esqueleto”, “desarrollo cartílago” y “osificación” y los genes citados. En total hemos seleccionado 53 genes (Tabla 1).

Tabla 1. Listado de genes candidatos agrupados por su función

Morfogénesis del esqueleto	Desarrollo del cartílago	Osificación
<i>ACVR2B</i>	<i>BMP1</i>	<i>ADRB2</i>
<i>BAPX1 (NKX3-2)</i>	<i>BMP2</i>	<i>AMBN</i>
<i>FOXC2</i>	<i>BMP3</i>	<i>AMELX</i>
<i>HOXA2</i>	<i>BMP4</i>	<i>AXIN2</i>
<i>MYF5</i>	<i>BMP5</i>	<i>CALCA</i>
<i>NDST1</i>	<i>BMP6</i>	<i>CALCR</i>
<i>OSR2</i>	<i>BMP7</i>	<i>CBFB</i>
<i>PRKRA</i>	<i>BMP8A</i>	<i>CD276</i>
<i>RYK</i>	<i>BMP8B</i>	<i>CHRD</i>
<i>SIX1</i>	<i>CTGF</i>	<i>CSF1</i>
<i>SIX4</i>	<i>EIF2AK3</i>	<i>CHRD1</i>
<i>TBX1</i>	<i>MGP</i>	<i>DMP1</i>
<i>TBX4</i>	<i>MMP13</i>	<i>DSPP</i>
<i>TFAP2A</i>	<i>PITX1</i>	<i>ENPP1</i>
<i>TGFBR1</i>	<i>RUNX2</i>	<i>FGF18</i>
<i>WNT7A</i>	<i>SOX5</i>	<i>FGF23</i>
	<i>SOX6</i>	<i>FGFR2</i>
	<i>SOX9</i>	<i>FOXC1</i>
	<i>THRA</i>	

2.2. Descarga desde las bases de datos genómicas

Hemos descargado del UCSC Genome Browser (<http://genome.ucsc.edu/cgi-bin/hgGateway>) 3kb de secuencia inmediatamente flanqueante 5' de los 53 genes candidatos (en humanos). En estas regiones no todo el ADN presenta función reguladora, aunque sí esperamos que en estas regiones se encuentren los principales elementos reguladores. Asimismo, se descargaron las secuencias ortólogas correspondientes a chimpancé y orangután mediante el uso de Blast (<http://www.ncbi.nlm.nih.gov/BLAST/>).

El hecho de que no fuera posible identificar las secuencias ortólogas para algunas de estas regiones flanqueantes conllevó que de los 53 genes seleccionados sólo hayamos podido analizar 50; los descartados fueron *MGP*, *MMP13* y *THRA*. La elección de 3kb se establece en función de la necesidad de facilitar el alineamiento de las secuencias. Se empleó el programa BioEdit para la edición y alineamiento de las secuencias ortólogas entre las tres especies. Se realizó un alineamiento automático inicial utilizando el software ClustalW. Posteriormente se realizó en todos los casos una comprobación manual del alineamiento. En ocasiones, el alineamiento es complejo y resulta poco fiable, por lo que es preferible desechar parte o la totalidad de la región.

2.3. Identificación de los elementos reguladores conocidos que caen dentro de las secuencias flanqueantes de los genes candidatos humanos

Mediante Match (<http://www.gene-regulation.com/pub/databases.html>), programa disponible de la página web de TRANSFAC (una base de datos de elementos reguladores), identificamos los posibles TFBSs (Transcription Factor Binding Sites), es decir, sub-secuencias cortas que pudieran funcionar como sitios de unión de elementos reguladores (factores de transcripción), presentes en las secuencias flanqueantes.

2.4. Análisis de la posible importancia evolutiva de las regiones no codificantes reguladoras

Para el análisis de la posible importancia evolutiva de las regiones no codificantes reguladoras a partir de la comparación de la secuencia de ADN de dichas regiones entre diferentes especies, no existe un software de referencia como en el caso del análisis de las regiones codificantes (PAML, por ejemplo. Yang; Nielsen, 2002). Por ello, hemos desarrollado nuestro propio software (en Perl) el cual nos permite:

2.4.1. obtener el número de nucleótidos divergentes entre las secuencias de ADN de las tres especies correspondientes a los TFBSs reguladores identificados por Match así como del resto de la secuencia flanqueante, en principio no reguladora.

2.4.2. identificar aquellas mutaciones (sustituciones nucleotídicas) que aparecen exclusivamente en humanos. Contabilizar cuántas de estas mutaciones específicamente humanas caen dentro de los TFBSs detectados por Match y cuántas fuera de estos elementos.

2.4.3. analizar estadísticamente si las secuencias de ADN correspondientes a dichos TFBSs (identificados por Match) están sobre-representadas en mutaciones específicamente humanas. Para calcular la significación estadística de esta sobre-representación hemos utilizado la distribución hipergeométrica. Dicha distribución nos permite calcular la probabilidad de que, si en una población de N ítems hay k ítems del tipo A y extraemos de esta población una muestra de tamaño n , saquemos x ítems de tipo A. Es decir: $h(x;N,n,k) = \frac{[k C_x][_{N-k} C_{n-x}]}{[N C_n]}$.

Si existe un mayor número de sustituciones dentro de la secuencia de estos TFBSs en comparación con el resto de la secuencia de la región flanqueante, podríamos inferir que dicho gen presenta una evolución reguladora. Es decir, las diferencias morfológicas posiblemente asociadas a ese gen en concreto podrían deberse no a cambios en su secuencia codificante, sino a la divergencia de aquellos elementos que participan en decidir cuándo, cuánto y dónde se produce su expresión.

Tabla 2. Loci con sobre-representación significativa ($p < 0,05$) de sustituciones exclusivas humanas en alguno de los TFBSs detectados en sus regiones flanqueantes

Gen	Num. TFBSs detectados en ~3kb de secuencia flanqueante 5'	TFBS sobre-representado en sustituciones específicas humanas	% de sustituciones específicas humanas en TFBSs	% de sustituciones específicas humanas en el resto de la secuencia flanqueante	p
<i>BMP2</i>	9	HLF	11,11	0,44	0,035
		VBP	11,11	0,44	0,035
<i>BMP4</i>	13	GAAT-1	11,11	0,45	0,039
<i>BMP8A</i>	9	PAX-4	5,00	0,76	0,032
<i>CALCR</i>	7	CP2	10,00	0,29	0,028
<i>ENPP1</i>	9	COMP1	4,35	0,48	0,038
		HAND1-E47 considerando todos los TFBSs en conjunto	13,33 2,21	0,44 0,43	0,002 0,015
<i>HOXA2</i>	6	E2F	14,28	0,33	0,023
		EVI-1 considerando todos los TFBSs en conjunto	7,14 2,29	0,33 0,31	0,045 0,026
<i>MYF5</i>	10	E2F	14,28	0,44	0,03
		considerando todos los TFBSs en conjunto	1,53	0,41	0,035
<i>RUNX2</i>	7	PAX-4	5,00	0,20	0,038
		considerando todos los TFBSs en conjunto	1,69	0,17	0,014
<i>SOX9</i>	9	USF	11,11	0,29	0,025
<i>TFAP2A</i>	4	USF	22,22	0,55	0,001
		considerando todos los TFBSs en conjunto	3,51	0,55	0,035

3. RESULTADOS

Tras analizar los 50 genes detallados en la Tabla 1 (descartando *MGP*, *MMP13* y *THRA*, por no poder identificar las secuencias ortólogas), hemos encontrado sobre-representación de sustituciones exclusivas del linaje humano estadísticamente significativa en 10 loci (Tabla 2). De ellos, 10 genes muestran sobre-representación para regiones reguladoras individuales, y 5 genes muestran sobre-representación cuando consideramos el conjunto de TFBSs detectados para cada gen. Estos 5 genes son:

- *ENPP1*, que codifica una fosfodiesterasa/pirofosfatasa ectonucleotídica. Esta proteína tiene una amplia especificidad y rompe una variedad de sustratos, incluidos los enlaces fosfodiéster y pirofosfato de los nucleótidos y

azúcares nucleotídicos. Algunas mutaciones descritas para este gen se asocian a enfermedades óseas, como la calcificación arterial infantil generalizada (GACI), una enfermedad autosómica recesiva severa caracterizada por un fenotipo hipermineralizante o raquitismo hipofosfatémico, y que se presenta como deformidades esqueléticas y retraso del crecimiento (Lorenz-Depiereux; et al., 2010). Por otro lado, también se han descrito polimorfismos asociados a la resistencia a la insulina y otras anomalías metabólicas relacionadas como cambios en el peso corporal, diabetes tipo 2 y complicaciones vasculares (Bacci, 2007).

- *HOXA2*, u homeobox A2, pertenece al conjunto de genes *HOX*, que desempeñan un papel muy importante en la morfogénesis de los embriones de vertebrados, ya que proveen información regional a lo largo del eje antero-posterior del cuerpo para la organización de los diferentes segmentos. Este gen codifica de hecho un factor de transcripción que se enlaza a ADN. Una mutación en este gen se ha relacionado con la microtia, una deformidad congénita del oído externo (Alasti; et al., 2008).
- *MYF5*, codifica el factor miogénico 5, perteneciente a la familia de factores de transcripción hélice-bucle-hélice, y que es capaz de activar el programa de diferenciación muscular (miogénesis). Mutaciones en este gen pueden provocar alteraciones en el desarrollo del músculo esquelético (Braun; Arnold, 1995).
- *RUNX2*, codifica una proteína nuclear esencial para la diferenciación de los osteoblastos y la morfogénesis esquelética. Asimismo, actúa como un andamio para factores reguladores implicados en la expresión. Algunas mutaciones en este gen han mostrado una asociación con la presencia de displasia cleidocraneal, un trastorno hereditario poco frecuente, en el que se produce una osificación defectuosa de los huesos craneales con fontanelas amplias y retraso en el cierre de las suturas, además de ausencia parcial o total de clavículas y anomalías dentales y vertebrales, entre otros. Otros estudios muestran asociación con las diferencias en la densidad mineral de los huesos y el riesgo de fractura ósea (Vaughan; et al., 2002), o la variación en la longitud del femoral (Ermakov; et al., 2005).
- *TFAP2A*, codifica el factor de transcripción de AP2-. La expresión de esta proteína es específica del tipo celular, y muestra una expresión temporal y espacial diferencial durante el desarrollo en diversos tejidos. Esta proteína es esencial durante la embriogénesis, como se ha demostrado a partir de estudios en ratón. La deficiencia en este factor de transcripción, *TFAP2A*, conlleva severas malformaciones del cierre craneal y de distintos órganos, y muerte en el nacimiento (Schorle; et al., 1996). La pérdida de actividad de esta proteína en general está relacionada con procesos de tumorigénesis, pues altera la proliferación e induce diferenciación prematura y/o apoptosis en varios tipos celulares (Hilger-Eversheim; et al., 2000).

De entre los TFBSs, los sitios de enlace para *PAX-4*, *E2F* y *USF* aparecen con sobre-representación de sustituciones exclusivas humanas en más de un gen. *USF* y *HAND1-E47* son los únicos sitios de enlace a factores de transcripción que presentan un valor p significativo incluso tras corrección por múltiples tests, es decir, tras tener en cuenta que para cada región se han analizado varios sitios de enlace a factores de transcripción (Tabla 2).

No hemos realizado una corrección por múltiples tests globales, dado que el elevado número de tests realizados (368) impondría un valor p excesivamente bajo ($\sim <0,0001$), de dudosa interpretación biológica.

4. DISCUSIÓN

Aunque se conocen muchos aspectos sobre la evolución a nivel morfológico del esqueleto de *Homo sapiens*, aún no se ha conseguido elucidar cuáles son los mecanismos genéticos que han actuado a nivel evolutivo. Estudios comparativos de las regiones codificantes de genes entre humanos y otros primates han puesto de manifiesto que no existen diferencias genéticas significativas que expliquen la aparición de los rasgos que nos hacen humanos, como los procesos cognitivos o de comportamiento (Nielsen; et al., 2005). Así, uno de los mecanismos más aceptados que explican las diferencias fenotípicas entre organismos estrechamente relacionados, como humanos y chimpancés, son los cambios a nivel de la regulación de la expresión génica (King; Wilson, 1975).

Los factores de transcripción son componentes celulares fundamentales en el control de la expresión génica. Ellos determinan cómo deben funcionar y responder las células ante estímulos internos y externos (Vaquerizas; et al., 2009). Numerosas enfermedades surgen por una disrupción en los sistemas reguladores de la expresión génica. De hecho, hay un número mayor de lo esperado de factores de transcripción (FT) que funcionan como oncogenes, es decir, que desencadenan un proceso tumoral cuando las mutaciones impiden que desarrollen correctamente su función. Además, un tercio de los desórdenes del desarrollo en humanos está causado por mutación en algún FT (mutación que convierte al FT en no funcional). Así, por ejemplo, fijándonos en los factores de transcripción cuyos TFBSs muestran sobre-representación significativa de sustituciones exclusivas humanas, para todos ellos se ha descrito alguna mutación causante de patologías o malformaciones. Aquellos FT con homeodominios (dominios de unión a ADN), como *PAX-4*, están asociados con procesos del desarrollo; en concreto *PAX-4* tiene un papel importante en la diferenciación y desarrollo de células pancreáticas, y ciertas mutaciones en esta proteína se relacionan con susceptibilidad a diabetes (Biason-Lauber; et al., 2005); Por otro lado, los factores de transcripción *USF* y *HAND1-E47* pertenecen a la familia de FT con dominios de unión a ADN del tipo hélice-bucle-hélice (HLH, *Helix-loop-helix*), importantes también en el desarrollo y función celular. Se conocen patologías asociadas a mutaciones en estos FT, como la hiperlipidemia combinada familiar, originada por mutaciones en *USF1* (Pajukanta; et al., 2004), o malformaciones cardíacas debidas a mutaciones en *HAND-1* (Reamon-Buettner; et al., 2009). Por

último, E2F es un factor de transcripción con estructura hélice-giro-hélice (HTH, *Helix-turn-helix*), implicado en la regulación de la expresión de genes del ciclo celular (Zheng; et al., 1999).

Por otro lado, se ha demostrado que la alteración en la actividad o en la especificidad de los FT es una fuente de diversidad fenotípica y adaptación evolutiva (Bustamante; et al., 2005; DE; et al., 2008). Los resultados obtenidos en este trabajo nos confirman que en algunas regiones 5'-flanqueantes (aguas arriba) de genes, es decir, en aquellas zonas donde se concentra la actividad reguladora de la expresión génica, el número de sustituciones nucleotídicas específicas humanas en las secuencias supuestamente correspondientes a TFBSs, es mayor de lo esperado en comparación con el resto de la secuencia (no TFBSs) de la misma región flanqueante. Esto concuerda con los resultados de un estudio reciente en el que se demuestra que existe una gran variación genética funcional en los lugares de unión a factores de transcripción con efectos en la expresión génica. En general, estiman que las diferencias observadas en los TFBSs entre humanos y chimpancés constituyen un 32%, mientras que la variación en las secuencias codificantes es de tan solo un 0,71% (Kasowski; et al., 2010).

Asimismo, queda por demostrar la funcionalidad de dichos elementos de enlace de FT. En humanos, los análisis bioinformáticos iniciales del genoma estimaron la presencia de unos 200-300 FT básicos, es decir, necesarios para la maquinaria transcripcional básica, y unos 2.000-3.000 FT cuya función es dependiente de que se enlacen a la secuencias de ADN específicas. Otras predicciones estiman unos 1.000-1.500 FT. Sin embargo, sólo 62 de ellos han sido confirmados experimentalmente como verdaderos FT. Los TFBSs constituyen un componente importante de la arquitectura reguladora, pero son pequeños, de un tamaño de entre 6 y 15 pares de bases, y por tanto difícil de diferenciar de secuencias muy similares que han emergido por azar (Gaffney; et al., 2008). Es por ello probable que una proporción sustancial de los TFBSs identificados sean falsos positivos. Aunque demostrar la funcionalidad de los TFBS escapa del ámbito de este proyecto, el estudio de los dominios de unión al ADN (DBD, DNA binding domains) de los factores de transcripción sí que puede dar pistas sobre su función así como de sus historias evolutivas (Charoensawan; et al., 2010).

Nuestros resultados apoyan la posibilidad de que al menos parte de las diferencias morfológicas esqueléticas que separan a los humanos del resto de los primates puedan ser debidas a diferencias en la regulación de la expresión génica. Se ha sugerido que el ADN regulador en los primates está bajo una presión selectiva más relajada con respecto a los roedores, probablemente por una disminución del tamaño efectivo de la población, lo que facilitaría la fijación de mutaciones deletéreas en estas regiones del ADN. De la misma manera, comparando regiones codificantes y no codificantes de humanos con otros primates se ha visto que en humanos la presión selectiva también se ha visto reducida dando lugar a una acumulación mayor de mutaciones ligeramente deletéreas (Gaffney; et al., 2008). Estudios recientes han atribuido este hecho a mayores tasas de evolución adaptativa en las regiones no codificantes del ADN (Prabha-kar; et al., 2006; Kim; Pritchard, 2007).

De los 5 genes en los que hemos encontrado sobre-representación significativa de sustituciones exclusivas del linaje humano, considerando todos los TFBS, 3 de ellos pertenecen al grupo de genes implicados en la morfogénesis del esqueleto (*HOXA2*, *MYF5* y *TFAP2A*), mientras que uno pertenece al conjunto de genes que participan en el desarrollo del cartílago (*RUNX2*) y otro en el proceso de osificación (*ENPP1*). Resulta interesante la sobre-representación de sustituciones exclusivas en este último, pues recientemente se han identificado polimorfismos en las regiones promotora, aguas arriba y aguas abajo del gen *ENPP1* asociadas a distintas características morfológicas craneofaciales (ERMAKOV; et al., 2010).

De estos resultados además se puede inferir que en la evolución del esqueleto son diversos los procesos biológicos que están sometidos a selección, e incluyen al menos la morfogénesis del esqueleto, el desarrollo del cartílago y la osificación. Los genes analizados en este estudio son aquellos que se ha observado que controlan los pasos claves del desarrollo esquelético, pero es probable sin embargo que el número de genes implicados sea muy superior. Consecuentemente, es de esperar que un trabajo más inclusivo y con mayor número de especies próximas evolutivamente, y con mayor calidad de resecuenciación, pueda aportar más información sobre los mecanismos genéticos responsables de los cambios morfológicos esqueléticos que nuestra especie ha sufrido a lo largo de la evolución.

5. CONCLUSIONES

Los resultados obtenidos nos permiten apoyar la hipótesis de que al menos parte de las diferencias morfológicas esqueléticas que separan a los humanos del resto de los primates pueden ser debidas a diferencias en la regulación de la expresión génica. Estas diferencias estarían causadas por sustituciones nucleotídicas en regiones del genoma que comprenden sitios de unión a factores de transcripción (TFBSs).

En este sentido, hemos encontrado pruebas de que en las regiones flanqueantes de al menos cinco genes existe sobre-representación significativa de sustituciones nucleotídicas de manera exclusiva en humanos. Esto sugiere que dichas sustituciones podrían haber conferido al linaje humano de una ventaja adaptativa.

No obstante, debemos considerar este estudio como aproximativo. Sería conveniente realizar un estudio a mayor escala, que incluyera más genes y más especies, para corroborar con mayor robustez estadística cuáles son las verdaderas fuerzas selectivas que han actuado o actúan sobre el desarrollo del esqueleto. En último término, sería necesario estudiar el efecto biológico de las sustituciones nucleotídicas humanas en dichos TFBSs mediante ensayos funcionales *in vitro* o *in vivo* (en ratón) con objeto de observar experimentalmente el efecto cada una de ellas en la regulación de la expresión de los genes adyacentes.

En este sentido, los elementos identificados proporcionan un punto de partida a partir del cual se podrán construir nuevas hipótesis sobre la regulación de la morfología del esqueleto humano, su evolución, e incluso de la patología de algunas enfermedades esqueléticas.

AGRADECIMIENTOS

Para la realización de este proyecto hemos disfrutado de una ayuda a la investigación de la Sociedad de Estudios Vascos - Eusko Ikaskuntza.

Además, el trabajo en nuestro laboratorio está subvencionado por el MICINN (CGL2008-04066/BOS) y el Gobierno Vasco (IT542-10 (GIC10/46)). S. L. disfruta de una beca predoctoral del Gobierno Vasco (BFI09.258).

BIBLIOGRAFÍA

- ALASTI, Fatemeh; et al. "A Mutation in HOXA2 Is Responsible for Autosomal-Recessive Microtia in an Iranian Family". *Am J Hum Genet*, vol. 82, 2008; pp. 982-991.
- ALONSO, Santos; IZAGIRRE, Neskuts; DE LA RÚA, Concepción. "Anthropological genomics: investigating the human phenotype and its evolution". En: SANTOS, Cristina; LIMA, Manuela (eds.). *Recent advances in Molecular Biology and Evolution: Applications to Biological Anthropology*. Kerala, India: Research Signpost, 2007; pp. 37-64.
- BACCI, Simonetta; DE COSMO, Salvatore; PRUDENTE, Sabrina; TRISCHITTA, Vincenzo. "ENPP1 gene, insulin resistance and related clinical outcomes". *Curr Opin Clin Nutr Metab Care*, vol. 10, 2007; pp. 403-409.
- BIASON-LAUBER, Anna; et al. "Association of childhood type 1 diabetes mellitus with a variant of PAX4: possible link to beta cell regenerative capacity". *Diabetología*, vol. 48, 2005; pp. 900-905.
- BRAUN, Thomas; ARNOLD, Hans H. "Inactivation of Myf-6 and Myf-5 genes in mice leads to alterations in skeletal muscle development". *EMBO J*, vol 14, 2005; pp.1176-1186.
- BUSTAMANTE, Carlos; et al. "Natural selection on protein-coding genes in the human genome". *Nature*, vol. 437, 2005; pp. 1153-1157.
- CARROLL, Sean B. "Genetics and the Making of *Homo sapiens*". *Nature*, vol. 422, 2003; pp. 849-857.
- . "Evolution at Two Levels: On Genes and Form". *PLoS ONE*, vol. 3, 2005; p. e425.
- CHAROENSAWAN, Varodom; WILSON, Derek; TEICHMANN, Sarah A. "Lineage-specific expansion of DNA-binding transcription factor families". *Trends in Genetics*, vol. 26, 2010; pp. 388-393.
- CLARK, Andrew; et al. "Inferring Nonneutral Evolution from Human-Chimp-Mouse Orthologous Gene Trios". *Science*, vol. 302, 2003; pp. 1960-1963.

López, S.; [et al.]: Genes implicados en el desarrollo del esqueleto. Selección adaptativa...

DE, Subhajyoti; LOPEZ-BIGAS, Nuria; TEICHMANN, Sarah. "Patterns of evolutionary constraints on genes in humans". *BMC Evol Biol*, vol. 8, 2008; p. 275.

ERMAKOV, Sergey; et al. "Variation in femoral length is associated with polymorphisms in RUNX2 gene". *Bone*, vol. 38, 2006; pp. 199-205.

ERMAKOV, Sergey; ROSENBAUM, Michael G.; MALKIN, Ida; LIVSHITS, Gregory. "Family-based study of association between ENPP1 genetic variants and craniofacial morphology". *Ann Hum Biol*, 2010, doi:10.3109/03014461003639231.

GAFFNEY, Daniel J.; BLECKMAN, Ran; MAJEWSKI, Jacek. "Selective constraints in experimentally defined primate regulatory regions". *PLoS Genet*, 2008, vol. 4, e10000157.

HILGER-EVERSHEIM, Kristina; MOSER, Markus; SCHORLE, Hubert; BUETTNER, Reinhard. "Regulatory roles of AP-2 transcription factors in vertebrate development, apoptosis and cell-cycle control". *Gene*, vol. 260, 2000; pp. 1-12.

JACOB, Françoise. "Evolution and Tinkering". *Science*, vol. 196, 1977; pp. 1161-1166.

KASOWSKI, Maya; et al. "Variation in transcription factor binding among humans". *Science*, vol. 328, 2010; pp. 232-235.

KIM, Su Y; PRITCHARD, Jonathan K. "Adaptive evolution of conserved noncoding elements in mammals". *PLoS Genet*, vol. 3, 2007; pp. 1572-1586.

KING, Mary C; WILSON, Alan C. "Evolution at two levels in humans and chimpanzees". *Science*, vol. 188, 1975; pp. 107-116.

LORENZ-DEPIEREUX Bettina; et al. "Loss-of-function ENPP1 mutations cause both generalized arterial calcification of infancy and autosomal-recessive hypophosphatemic rickets". *Am J Human Genet*, vol. 12, 2010; pp. 267-272.

NIELSEN, Rasmus; et al. "A scan for positively selected genes in the genomes of humans and chimpanzees". *PLoS Biol*, 2005, vol. 3: e170.

PAJUKANTA, Päivi; et al. "Familial combined hyperlipidemia is associated with upstream transcription factor 1 (USF1)" *Nat. Genet*, vol. 36, 2004; pp. 371-376.

PRABHAKAR, Shyam; NOONAN, James P; PAABO, Svante; RUBIN, Edward M. "Accelerated evolution of conserved noncoding sequences in humans". *Science*, vol. 314, 2006; pp. 786-786.

RUFF, Christopher B. "Gracilization of the modern human skeleton". *Am Sci*, vol. 94, 2006; pp. 508-514.

SCHORLE, Hubert; MEIER, Pascal; BUCHERT, Michael; JAENISCH, Rudolf; MITCHELL, Pamela J. "Transcription factor AP-2 essential for cranial closure and craniofacial development". *Nature*, vol. 381, 1996; pp. 235-238.

VAQUERIZAS, Juan M; KUMMERFELD, Sarah K; TEICHMANN, Sarah A; LUSCOMBE, Nicholas M. "A census of human transcription factors: function, expression and evolution". *Nat Rev Genet*, vol. 10, 2009; pp. 252-263.

- VAUGHAN, Tanya; et al. "Alleles of RUNX2/CBFA1 gene are associated with differences in bone mineral density and risk of fracture". *J Bone Min Res*, vol. 17, 2002; pp. 1527-1534.
- YANG, Ziheng; BIELAWSKI, Joseph. "Statistical Methods for Detecting Molecular Adaptation". *Tree*, vol. 3, 2000; pp. 496-503.
- YANG, Ziheng; NIELSEN, Rasmus. "Codon-Substitution Models for Detecting Molecular Adaptation at Individual Sites Along Specific Lineages". *Mol Biol Evol*, vol. 19, 2002; pp. 908-917.
- ZHANG, Jianzhi; NIELSEN, Rasmus; YANG, Ziheng. "Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level". *Mol Biol Evol*, vol. 22, 2005; pp. 2472-2479.
- ZHENG, Ning; FRAENKEL, Ernest; PABO, Carl O; PAVLETICH, Nikola P. "Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP". *Genes Dev*, vol. 13, 1999; pp. 666-674.