
PRESENTE Y FUTURO DE LA GENÉTICA EVOLUTIVA

FRANCESC MESTRES

A pesar de la redundancia se puede afirmar que la genética evolutiva ha evolucionado de manera considerable en los últimos cien años. Se ha pasado de la época de los primeros modelos teóricos de Fisher, Haldane y Wright a la era de la secuenciación y de los análisis bioinformáticos. Ello ha abierto unas nuevas perspectivas insospechadas en el ámbito de la genómica. Bien podría decirse que esta disciplina se inauguró con los trabajos de Jacob y Monod (1961) sobre el operón lactosa de *E. coli*. Su trabajo de investigación es un buen ejemplo de análisis genético utilizando la colección de mutaciones que fueron capaces de obtener. Este procedimiento diferencia claramente a los genéticos de los bioquímicos. Pudo observarse que el modelo del operón era general en las bacterias, con diferentes variantes moldeadas por la selección natural. En un principio se supuso que dicho modelo era exportable a los organismos eucariotas. Sin embargo, éstos presentan genomas estructurados de manera más compleja. Existían secuencias de control múltiples, algunas delante de la secuencia del gen, otras detrás e incluso alguna interna. Las secuencias de control pueden estar muy alejadas de la secuencia del gen y muchas proteínas efectoras pueden controlar la expresión génica. Las sorpresas continuaron al descubrirse los genes partidos (Jeffreys and Flavell 1977), con alternancia de secuencias codificantes (exones) y no codificantes (intrones). Por lo tanto en muchos genes eucarióticos existía un procesado (“*splicing*”) del RNA recién transcrito para eliminar las secuencias no codificantes y, junto a modificaciones en los extremos, dar lugar al mRNA maduro. También se descubrió que en algunos casos podía haber diferentes puntos de “*splicing*”, produciéndose proteínas algo diferentes según las necesidades celulares. Con los años se fue viendo que existían más niveles de regulación de la expresión génica (un resumen básico puede encontrarse en el libro de texto de Pierce 2005). En todos ellos habría intervenido la selección natural, y por tanto cabe considerar con toda

Departament de Genètica, Facultat de Biologia, Universitat de Barcelona. Av. Diagonal, 645. 08028 – Barcelona, España. / fmestres@ub.edu.

Última contribución en *Ludus Vitalis*: “Conocimientos actuales y posibilidades de acción en genética evolutiva”, vol. XIV, num. 26, 2006, pp. 241-245.

seguridad que es un proceso altamente eficiente, a pesar de su complicación aparente para nuestra mentalidad. Cada vez se van detectando más casos en los que la selección natural positiva tiene un papel importante también a nivel molecular. Esta visión contrasta con la que se tenía en los años ochenta y buena parte de los noventa, en donde los cambios a nivel molecular eran considerados, por muchos genéticos evolutivos, como debidos fundamentalmente a la deriva genética, mientras que a la selección natural positiva se le otorgaba un papel insignificante. Esta escuela de pensamiento era la neutralista, encabezada por los trabajos pioneros de Kimura (1968). A esta gran controversia cabe atribuirle el actuar como catalizante de muchas investigaciones a nivel de la evolución molecular. Actualmente se admite la importancia de la deriva genética a dicho nivel, pero con el tiempo se ha revalorizado el efecto de la selección natural positiva en este ámbito y actualmente son frecuentes los artículos científicos con esta temática (por ejemplo, Tennessen 2005; Vermaak et al. 2005; Bakewell et al. 2007; Sabeti et al. 2007).

La genómica dio un gran paso adelante a partir de un proyecto indirectamente relacionado con la estructura del material genético, el de secuenciar el genoma humano completo. Cuando se presentó la propuesta hubo una gran controversia respecto a su utilidad y a sus posibles usos (Lee 1991). Sin embargo, actualmente se puede decir que, gracias al proyecto de secuenciación completa, el ser humano ha pasado a ser un organismo modelo (como la *Drosophila*, el maíz, o la bacteria *Escherichia coli*) en cuanto a que ha permitido conocer aspectos generales de la estructura genómica. En un principio se esperaba obtener e identificar todo el conjunto de genes del ser humano. Las primeras aproximaciones a la secuencia completa de nuestro genoma aparecieron en el año 2001 (Venter et al. 2001; Lander et al. 2001) y el anuncio de la obtención definitiva data del 2004 (*International Human Genome Sequencing Consortium* 2004). A pesar de ello aún no conocemos la lista completa de genes que codifican para proteínas en nuestro genoma. La primera gran sorpresa es que se había hecho la predicción de unos 100 000 genes; sin embargo, actualmente se estiman en poco más de 30 000 (Claverie 2001; Pennisi 2003; Stein 2004). ¿Cuál es la causa de esta diferencia? Ello es debido en gran parte a la distribución de los genes a lo largo de los cromosomas y, más importante aún, el estudio en detalle de regiones genómicas ha modificado nuestro concepto de gen (Pearson 2006; Reese and Guigó 2006). El gen académico que mostramos en nuestras clases (secuencias alternadas de exones e intrones, flanqueadas a inicio y fin del gen por secuencias de control) no siempre se cumple. En la realidad muchas veces hay secuencias a lo largo del genoma que pueden actuar como controladoras de la expresión génica. También, y en forma dispersa, existen secuencias potencialmente codificantes. Esta realidad ha producido que a los programas informáticos les cueste identificar dónde hay

genes. Por ello deben realizarse diferentes aproximaciones experimentales e incluso, como apuntaba el experto en genómica R. Guigó en el reciente Congreso de la Sociedad Española de Genética, en último término llevar a cabo el análisis de las secuencias de forma manual por un experto, no por un programa informático (Guigó et al. 2007). La genética evolutiva debería profundizar en los mecanismos que han llevado a lo largo del tiempo a las estructuras genómicas actuales y que debemos presuponer que son actualmente las más adaptativas. Para nosotros la estructura genómica es difícil de interpretar y comprender correctamente, y luego todavía será más complejo entender cómo la evolución ha aumentado la complejidad desde genomas simples a los más complejos. Mi profesor A. Prevosti, nos comentaba hace unos años que el genoma era un todo integrado. Entender las redes de interacciones de la expresión génica siempre se ha pensado que sería complicado. Con todo, al secuenciar todo el genoma esperábamos tener al menos todos sus componentes o piezas. Es decir, es como si no tuviésemos idea de electrónica, y entonces para entender como funciona un televisor, nuestro primer enfoque fuese abrirlo, sacar todas las piezas y catalogarlas, para después intentar ver como se interconectan y funcionan. En el fondo se esperaba que con la secuenciación del genoma humano fuéramos capaces de tener todas las piezas (los genes y las secuencias de control). Resulta sin embargo que en muchas situaciones no somos ni capaces de identificar qué son piezas o que no lo son. Existen así secuencias que no somos capaces de discernir si son funcionales o no. En mi opinión esta complejidad podría explicar también la poca información obtenida en los estudios clásicos sobre el fenómeno de la especiación basados en los cruzamientos interespecíficos. Aunque las dos especies que se crucen sean evolutivamente cercanas, muy probablemente sus genomas se habrán diferenciado lo suficiente como para que si se mezclan generen un híbrido no funcional (inviabile o estéril). Siguiendo con el símil electrónico es como si tomásemos la mitad de piezas de una marca de televisor y las uniésemos con la otra mitad procedente de otra marca. Lo más probable es que se produzca una gran disfunción. Por este motivo, cuando se intenta estimar el número de genes responsables de un determinado proceso de especiación se obtienen estimas elevadas y se habla de *loci* poligénicos. Esto nos conduce a la cuestión de los análisis de la genética cuantitativa. Hace pocos años, con la aparición de técnicas que permiten analizar muchos marcadores moleculares en un amplio espectro de especies, ha sido posible la localización de los QTL (*quantitative trait loci* o posiciones responsables de los caracteres cuantitativos) responsables del control genético de un carácter cuantitativo. Pero, ¿qué es exactamente lo que hemos localizado en una determinada zona de un cromosoma? Debido a la complejidad del genoma, diversos tipos de secuencias nucleotídicas pueden estar involucradas en el control

genético del carácter: una secuencia de control, una secuencia potencialmente codificante, un gen en el sentido más clásico, etc. Por tanto, puede ser que un QTL sea uno de estos elementos. Respecto a la genética cuantitativa, una última consideración de tipo evolutivo: En la actualidad se está volviendo al análisis de modelos que permitan interpretar los efectos de las interacciones y de los diferentes mecanismos en la evolución de estos caracteres (Yang 2004; Zeng et al. 2005; Álvarez-Castro and Carlborg 2007).

Otra vertiente actual es la de la genómica comparada. Esta aproximación científica es posible debido a que cada vez se tienen más genomas total o parcialmente secuenciados. Hace unos años, los científicos que estudian evolutivamente los genes del desarrollo (disciplina denominada *evo-devo*) observaron que los genes *Hox* estaban repetidos cuatro veces en los vertebrados (García-Fernández and Holland 1996; Holland 1997). Esta y otras observaciones condujeron a la elaboración de la hipótesis 2R, según la cual se produjeron dos duplicaciones completas del genoma en la base de los vertebrados (Wolfe 2001; Vandepoele et al. 2004). Es difícil discernir entre duplicaciones regionales y duplicaciones genómicas enteras, debido a que los rastros de las duplicaciones se borran con el tiempo (muchos genes devienen pseudogenes, otros se modifican sustancialmente cambiando de función, etcétera). Los estudios bioinformáticos parecen indicar que se pueden seguir tramos de los genes duplicados por todo el genoma, lo que parece avalar la hipótesis 2R (Van de Peer 2004; Dehal and Boore 2005). En este caso, la genética evolutiva debería investigar cómo se han producido estas duplicaciones de todo el genoma y cómo la selección natural (y las otras fuerzas evolutivas) han modelado los genes duplicados para hacer de ellos un sistema adaptativo exitoso.

Relacionado con lo anteriormente expuesto está también el fascinante estudio del genoma mínimo. Si bien el tema del origen de la vida es complejo de abordar y existe una amplia controversia sobre el escenario donde apareció la vida (Bada et al. 2007; Wächtershäuser and Huber 2007), un ámbito actual de estudio es tratar de identificar qué genes son los imprescindibles para tener un sistema vivo. Es decir, tratar de identificar todos aquellos genes que permiten a un organismo ser funcional y a la vez transmitir esta información a la descendencia. Existen diferentes aproximaciones para intentar resolver estos interrogantes (Koonin 2000; Luisi et al. 2002; Gil et al. 2004). También debe comentarse el interés del polifacético científico y hombre de negocios J. C. Venter de producir en el laboratorio la primera forma viviente. En la actualidad ha patentado el genoma de *Mycoplasma laboratorium* (especie derivada de *M. genitalium*, a la que se ha añadido un cromosoma sintético) y que se espera que produzca biocombustible y absorba dióxido de carbono (Glass et al. 2006).

Finalmente, me gustaría recordar que la genética evolutiva tiene una serie de vertientes de estudio relacionadas con el ser humano y que ya

expuse en un artículo anterior (Mestres 2006). Por ejemplo, será de gran interés en el futuro observar si las fuerzas evolutivas serán capaces de modificar las frecuencias de ciertos polimorfismos genéticos (algunos de ellos adaptativos) que se han visto alteradas por los movimientos migracionales masivos de la especie humana como consecuencia de la globalización de nuestro planeta.

Como miembro de la comunidad universitaria siento también una gran inquietud por los aspectos docentes de la genética evolutiva. ¿Se debe continuar enseñando los modelos evolutivos sencillos que permiten explicar los cambios elementales de las frecuencias génicas, cuando estamos actualmente en la época de la genómica? Creo que sí se debe enseñar la parte clásica de la genética de poblaciones, puesto que los nuevos conocimientos no anulan a los precedentes. Por ejemplo, en física se enseña mecánica clásica y también mecánica cuántica. Es obligación del docente universitario actualizar los temarios y, aunque se mantengan aquellos temas básicos de evolución, hay que añadir otros referentes a la evolución molecular y la genómica. Sólo de esta forma tendremos alumnos bien formados, con una buena base conceptual y una visión actual de la genética evolutiva.

BIBLIOGRAFÍA

- Álvarez-Castro, J. M. and Örfjan, C. (2007), "A unified model for functional and statistical epistasis and its application in quantitative trait loci analysis", *Genetics* 176: 1151–1167.
- Bada, J. L., Fegley jr., B., Miller, S. L., Lazcano, A., Cleaves, H. J., Hazen, R. M. and Chalmers, J. (2007), "Debating evidence for the origin of life on Earth", *Science* 315: 937–938.
- Bakewell, M. A., Shi, P. and Zhang, J. (2007), "More genes underwent positive selection in chimpanzee evolution than in human evolution", *Proc. Natl. Acad. Sci. USA* 104: 7489–7494.
- Claverie, J.-M. (2001), "Gene number: What if there are only 30,000 human genes?", *Science* 291: 1255–1257.
- Dehal, P. and Boore, J. L. (2005), "Two rounds of whole genome duplication in the ancestral vertebrate", *PLoS Biology* 3: e314.
- García-Fernández, J. and Holland, P. W. (1996), "Amphioxus Hox genes: insights into evolution and development", *Int. J. Dev. Biol. Suppl.* 1: 71s–72s.
- Gil, R., Silva, F. J., Peretó, J. and Moya, A. (2004), "Determination of the core of a minimal bacterial gene set", *Microbiol. Mol. Biol. Rev.* 68: 518–537.

- Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M., Hutchinson, C. A., Smith, H. O. and Venter, J. C. (2006), "Essential genes of a minimal bacterium", *Proc. Natl. Acad. Sci. USA* 103: 425–430.
- Guigó, R. and the ENCODE Genes and Transcripts Analysis Group (2007), "The ENCODE Project: uncovering the transcriptional complexity of the Human Genome". XXXVI Congreso de la Sociedad Española de Genética. León (España).
- Holland, P. W. (1997), "Vertebrate evolution: Something fishy about *Hox* genes", *Curr. Biol.* 7: R570–R572.
- International Human Genome Sequencing Consortium (2004), "Finishing the euchromatic sequence of the human genome", *Nature* 431: 931–945.
- Jacob, F. and Mond, J. (1961), "Genetic regulatory mechanism in the synthesis of proteins", *J. Mol. Biol.* 3: 318–356.
- Jeffreys, A. J. and Flavell, R. A. (1977), "The rabbit beta-globin gene contains a large insert in the coding sequence", *Cell* 12: 1097–1108.
- Kimura, M. (1968), "Evolutionary rate at the molecular level", *Nature* 217: 624–626.
- Koonin, E. V. (2000), "How many genes can make a cell: minimal-gene-set concept". *Annu. Rev. Genomics Hum. Genet.* 1: 99–116.
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C. et al. (2001), "Initial sequencing and analysis of the human genome", *Nature* 409: 860–921.
- Lee, T. F. (1991), *The Human Genetic Project*. Plenum Pub. Corp. N. Y.
- Luisi, P. L., Oberholzer, T. and Lazcano, A. (2002), "The notion of a DNA minimal cell: a general discourse and some guidelines for an experimental approach", *Helv. Chim. Acta* 85: 1759–1777.
- Pearson, H. (2006), "What is a gene?", *Nature* 441: 399–401.
- Pennisi, E. (2003), "Gene counters struggle to get the right answer", *Science* 301: 1040–1041.
- Pierce, B. A. (2005), *Genetics. A Conceptual Approach*. NY: W. H. Freeman and Co.
- Reese, M. G. and Guigó, R. (2006), "EGASP: Introduction", *Genome Biol.* 7 (Suppl 1): S1.
- Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E. et al. (2007), "Genome-wide detection and characterization of positive selection in human populations", *Nature* 449: 913–918.
- Stein, L. D. (2004), "End of the beginning", *Nature* 431: 915–916.
- Tennessen, J. A. (2005), "Molecular evolution of animal antimicrobial peptides: widespread moderate positive selection", *J. Evol. Biol.* 18: 1387–1394.
- Van de Peer, Y. (2004), "Computational approaches to unveiling ancient genome duplications", *Nature Reviews* 5: 752–763.
- Vandepoele, K., De Vos, W., Taylor, J. S., Meyer, A. and Van de Peer, Y. (2004), "Major events in the genome evolution of vertebrates: Paranome age and size differs considerably between ray-finned fishes and land vertebrates", *Proc. Natl. Acad. Sci. USA* 101: 1638–1643.
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J. et al. (2001), "The sequence of the human genome", *Science* 291: 1304–1351.
- Vermaak, D., Henikoff, S. and Malik, H. S. (2005), "Positive selection drives the evolution of *rhino*, a member of the heterochromatin protein 1 family in *Drosophila*", *PLoS Genet.* 1: e9.
- Wächtershäuser, G. and Huber, C. (2007), "Response", *Science* 315: 938–939.
- Wolfe, K. H. (2001), "Yesterdays's polyploids and the mystery of diploidization", *Nature Reviews* 2: 333–341.
- Yang, R.-C. (2004), "Epistasis of quantitative trait loci under different gene action models", *Genetics* 167: 1493–1505.
- Zeng, Z.-B., Wang, T. and Zou, W. (2005), "Modeling quantitative trait loci and interpretation of models", *Genetics* 196: 1711–1725.