

La huella genética de la selección natural

 F. Perfectti ¹, F.X. Picó ², J.M. Gómez ³

(1) Departamento de Genética. Universidad de Granada, E-18071 Granada, España.

(2) Estación Biológica de Doñana (CSIC). Pabellón del Perú, E-41013 Sevilla, España.

(3) Departamento de Ecología, Universidad de Granada, E-18071 Granada, España.

➤ Recibido el 5 de enero de 2009, aceptado el 27 de enero de 2009.

Perfectti, F., Picó, F.X., Gómez, J.M. (2009). La huella genética de la selección natural. *Ecosistemas* 18(1):10-16.

La selección natural es un proceso biológico que constituye uno de los principales motores de cambio evolutivo, y el origen de las adaptaciones fenotípicas. La selección que actúa sobre una determinada población se puede detectar ecológicamente cuantificando a nivel intrapoblacional la relación existente entre el fenotipo de los individuos y su éxito reproductivo diferencial. Sin embargo, detectar la actuación de la selección pretérita es imposible usando una aproximación exclusivamente ecológica. Para afrontar este problema, durante los últimos 40 años se han desarrollado una serie de métodos moleculares que infieren la actuación de la selección. Un primer grupo de pruebas usa como modelo nulo la teoría cuasi neutral de la evolución molecular, y asumen que ha ocurrido selección cuando el resultado difiere significativamente de lo esperable según evolución neutral. Algunos tests, como por ejemplo la D de Tajima, la relación d_N/d_S , o las pruebas de MK o HKA, pueden determinar incluso el tipo de selección operante. Sin embargo, raramente pueden cuantificar la intensidad de selección. Este inconveniente es exitosamente superado por el análisis de la selección basada en Campos Aleatorios de Poisson (PRF). Un segundo grupo de pruebas usa como método para inferir la acción pretérita de la selección natural la comparación entre la cantidad de diferenciación genética en caracteres cuantitativos (Q_{ST}) y la cantidad de diferenciación genética neutra (F_{ST}). Finalmente, unos de los más recientes y prometedores métodos que se están desarrollando para el estudio de la selección natural está relacionado con el avance tan espectacular que está registrando en la actualidad la genómica. Aún estamos lejos de saber la verdadera utilidad de esta herramienta para el estudio de la selección natural en poblaciones naturales, pero el pronóstico es bastante esperanzador.

Palabras clave: selección natural, deriva genética, evolución neutral, genómica.

Perfectti, F., Picó, F.X., Gómez, J.M. (2009). Genetic evidences of natural selection. *Ecosistemas* 18(1):10-16.

Natural selection is one of the main factors driving the evolutionary process and the only one leading to adaptations. Selection acting on a population can be detected by quantifying the relationship between individual phenotype and relative fitness. However, detecting the effect of past selection is not possible by only using an ecological approach. To cope with this issue, several molecular evolutionary genetics methods have been developed during the last 40 years. A first group of tests use the quasi neutral theory of molecular evolution as null model, and assume the occurrence of selection when the observed outcome significantly departs from the expectations under neutral evolution. Some tests, such as Tajima's D, d_N/d_S ratio, MK test or HKA test, can determine the kind of selection acting on populations, but are unable to quantify its intensity. The analysis of selection based on Poisson Random Fields (PRF) overcomes this caveat. A second group of analyses infer selection by comparing the amount of genetic variation in quantitative traits (Q_{ST}) against the amount of neutral genetic variation (F_{ST}). Lastly, one of the most recent and promising methods to explore natural selection is related with the superb advance in genomics. Although we are still far from fully appreciate the importance of this tool, we believe that the use of the genomics will produce a qualitative enhancement in our understanding of the importance of natural selection in the wild.

Key words: natural selection, genetic drift, neutral evolution, genomics.

El grado de diferenciación y divergencia entre poblaciones o especies relacionadas, que pueden observarse y cuantificarse a nivel cuantitativo y molecular, puede ser explicado por selección natural y/o deriva genética. Podemos decir que ambos procesos operan simultáneamente en todas las poblaciones aunque la importancia relativa de cada uno de ellos puede variar de una manera substancial. Tradicionalmente, entender bien los procesos que determinan la diferenciación poblacional y el peso que la selección natural puede tener frente a la deriva genética en dichos procesos ha sido de capital importancia para la biología evolutiva. La forma habitual para un ecólogo de detectar un evento de selección es establecer una relación estadística entre un rasgo fenotípico y la eficacia biológica, caracterizando así un suceso de selección fenotípica. La propia naturaleza de esta forma de proceder implica que hay que cazar los eventos de selección en el momento en que la selección natural está operando. Sin embargo, hay otra aproximación basada en el análisis de secuencias que busca las señales dejadas por la

selección natural en los genomas de los organismos. Esta aproximación está basada en la historia de acumulación de sustituciones en los linajes y no es una aproximación contemporánea, a diferencia de la aproximación ecológica basada en la asociación fenotipo-eficacia biológica. Además, y a diferencia de ciertas interpretaciones seleccionistas, tampoco se trata de inferir selección a partir de un modelo de optimización que prejuzga que hay adaptación, sino más bien justamente lo contrario: a partir de modelos que consideran la evolución neutral como la hipótesis nula, determinar qué desviaciones frente al modelo neutral ofrecen pistas para establecer los eventos de selección.

Detectando la huella genética de la selección natural.

La hipótesis nula: teoría cuasi neutral de la evolución molecular.

Kimura (1968) y King y Jukes (1969) desarrollaron la teoría neutral de la evolución molecular, pasando de una interpretación seleccionista de la evolución a una donde la mayor parte del cambio en las secuencias era no adaptativo y debido a la mutación. La probabilidad de reemplazo de un nucleótido por otro en una secuencia solo dependería de la tasa de mutación, jugando la selección un papel minoritario y reservado a la eliminación de las variantes deletéreas. Una importante revisión de la teoría fue propuesta por Tomoko Ohta (1973; 1992), que introdujo la teoría casi neutral de la evolución molecular, donde se considera que la evolución de mutaciones ligeramente deletéreas o ligeramente beneficiosas, casi-neutrales, va a depender fundamentalmente de la tasa de mutación, de los coeficientes de selección y del tamaño poblacional. Así, en linajes evolutivos con tamaños pequeños de población es posible encontrar un mayor número de cambios no neutrales, ligeramente deletéreas, mantenidos a baja frecuencia o incluso fijados por deriva genética.

A partir de este modelo teórico de evolución de secuencias se puede tener una hipótesis nula con la que comparar un juego de secuencias obtenidas, ya sea de poblaciones de una misma especie, o más comúnmente de especies diferentes pero con un ancestro común no muy lejano. Las desviaciones frente a este modelo pueden ser interpretadas como claros indicios de eventos de selección positiva ($s>0$) o, más frecuentemente, como evidencias de selección purificadora ($s<0$), aunque la selección equilibradora, o “balanceada”, u otros procesos demográficos también pueden ser detectados (**Figura 1**). A continuación haremos un repaso a los estadísticos más utilizados para demostrar desviaciones frente a la hipótesis nula de neutralidad.

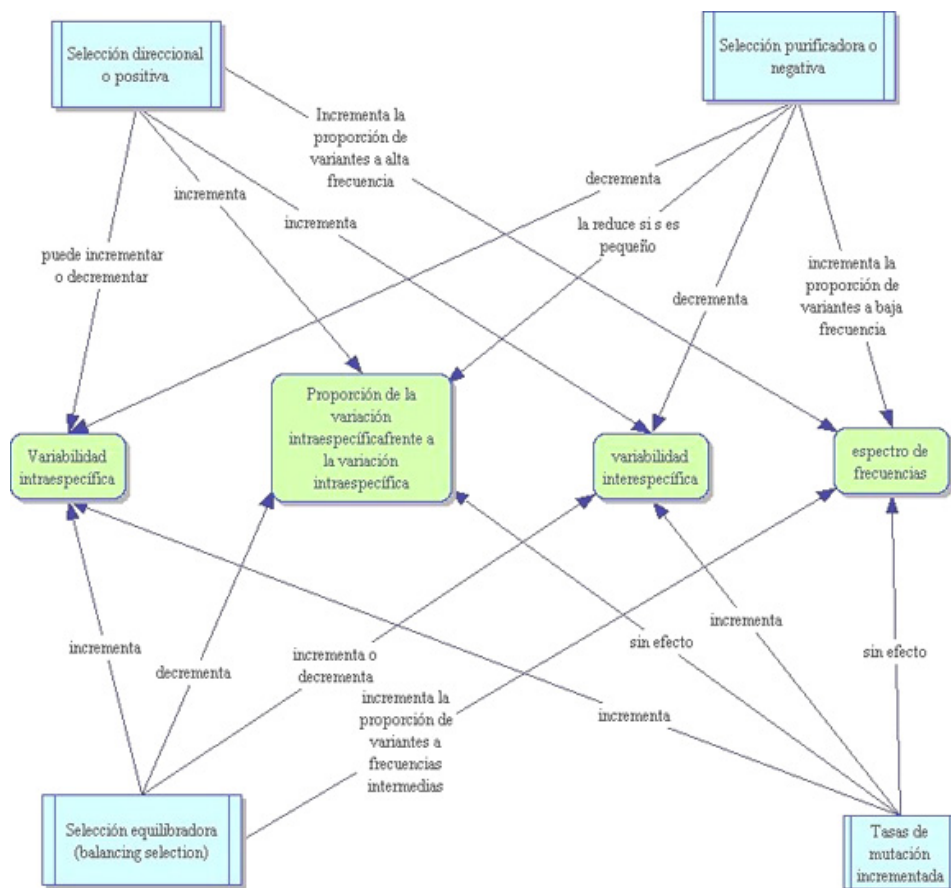


Figura 1. Esquema mostrando los efectos de diversos tipos de selección en la diversidad nucleotídica intra- e inter-específica (Nielsen 2005).

D de Tajima y sus derivados

Estos estadísticos se han desarrollado para determinar si el espectro de frecuencias de los polimorfismos encontrados en un juego de secuencias es explicable bajo la hipótesis de evolución puramente neutral. El estadístico D (Tajima 1989) compara dos estimas de la cantidad de variación genética (θ) de un grupo de secuencias. La primera es obtenida a partir del número total de sitios polimórficos en la muestra de secuencias y la segunda como la proporción de diferencias nucleotídicas entre las secuencias comparadas dos a dos. Ambas estimas deben ser iguales bajo una situación de evolución neutral (no selección, no recombinación, no subdivisión poblacional, no cambios en el tamaño poblacional). D se basa en la diferencia entre ambas estimas y tiene media cero y desviación 1. Si se detectan desviaciones significativas, éstas son indicadoras de no neutralidad, pero el factor concreto que las explica puede ser difícil de distinguir. Así, por ejemplo, valores positivos de D pueden indicar selección positiva, selección equilibradora o bien una reducción del tamaño poblacional.

Otros tests han incorporado información filogenética (por ejemplo una especie outgroup) para intentar estimar la dirección de los cambios y así tener una mayor potencia para detectar desviaciones frente a la hipótesis nula. El test de Fu y Li (Fu y Li 1993) es parecido al anterior en que también calcula un estadístico D como la diferencia de dos estimas de θ . Este test está basado en que diferentes tipos de mutaciones se acumulan de forma diferente en una filogenia. Así, las mutaciones puramente neutrales pueden acumularse a lo largo de toda la filogenia, pero las mutaciones ligeramente deletéreas lo harán únicamente en las ramas externas, o lo que es lo mismo, son más recientes, puesto que la selección habrá tenido más tiempo para eliminar las mutaciones deletéreas que aparecieron hace más tiempo.

Fay y Wu (2000) también desarrollaron un test basado en que el espectro de frecuencias esperado bajo condiciones de neutralidad debe estar enriquecido en mutaciones a baja frecuencia, siendo las mutaciones en alta frecuencia raras. Este test, que también requiere de información filogenética, es especialmente sensible a los arrastres o barridos selectivos. Hay que tener en cuenta que un evento de selección positiva eliminará polimorfismo no solo en la posición responsable (el nucleótido concreto), sino en una amplia región de posiciones ligadas. Así, un arrastre selectivo elimina variación muy rápidamente, pero también puede llevar a las variantes neutrales ligadas a tener frecuencias altas, dejando una señal que puede ser detectada por estos tests.

Detección de la selección usando d_N/d_S

Cuando se dispone de secuencias de genes codificadores de proteínas, a éstas se les pueden aplicar una serie de pruebas estadísticas muy útiles para valorar la importancia de la selección en la evolución de dichas secuencias. Quizás la más conocida y usada sea la razón d_N/d_S (Kimura 1977). Esta razón compara la tasa de sustituciones no sinónimas por sitio sinónimo (d_N) con la tasa de sustituciones sinónimas por sitio sinónimo (d_S). Habitualmente no se conoce el tiempo de divergencia entre las dos secuencias a comparar, y por tanto las tasas de cambio, y por ello se suelen utilizar las proporciones de sustituciones no sinónimas (por sitio no sinónimo) y sustituciones sinónimas (por sitio sinónimo). La interpretación de la razón d_N/d_S es directa asumiendo que las sustituciones sinónimas no se producen por selección y sólo reflejan la tasa neutral de sustitución. Así, $d_N/d_S = 1$ sería un indicador de evolución neutral y, por contra, valores superiores a la unidad indicarían selección positiva y valores inferiores selección purificadora. En la práctica, la estima de estos parámetros puede complicarse si la divergencia entre las secuencias ha sido grande. Además, ciertas sustituciones sinónimas podrían no ser neutrales y ser consecuencia de selección positiva. Hay ejemplos de que ciertas sustituciones sinónimas pueden afectar a la estabilidad del mRNA o a la mecánica del splicing o que incluso manifiesten cierta selección para minimizar los errores de traducción (Chamary y Hurst 2005; Parmley et al. 2006; Pal et al. 2006; Drummond y Wilke 2008). Las estimas genómicas de este parámetro varían ampliamente entre grupos. Así para *Drosophila* se reportan valores bajos (indicando una, altamente eficaz, selección purificadora) mientras que para mamíferos aparecen valores más altos que pueden indicar que la selección purificadora no es tan efectiva, posiblemente debido a un típico pequeño tamaño poblacional. Conviene recordar aquí que la selección purificadora será muy efectiva solo si $|s| \gg 1/4N_e$ y que, por tanto, en poblaciones pequeñas pueden fijarse mutaciones ligeramente deletéreas únicamente por deriva genética (Ohta 1987; Bromham y Penny 2003)

Prueba de MK

McDonald y Kreitman (1991) idearon un test que compara el polimorfismo intraespecífico con la divergencia interespecífica. El test se utiliza para comparar secuencias codificadoras de proteínas de dos especies muy emparentadas. Para ello se se construye una tabla de doble entrada donde se evalúa el número de sitios polimórficos sinónimos (P_S) y no sinónimos (P_N), y el número de sustituciones fijas sinónimas (D_S) y no sinónimas (D_N). Si la evolución es puramente neutral $P_N/P_S = D_N/D_S$. Sin embargo, si estas razones no son iguales, cabe asumir selección purificadora (si $P_N/P_S > D_N/D_S$) o positiva (si

$P_N/P_S < D_N/D_S$). Desgraciadamente, factores demográficos, como una expansión poblacional reciente, pueden producir una señal similar a la acción de la selección, complicando la interpretación del test.

Prueba de HKA

La prueba de Hudson-Kreitman-Aguadé (HKA, Hudson et al. 1987) utiliza secuencias de varios loci de especies altamente emparentadas, y pone a prueba la hipótesis neutral de que la divergencia (interespecifica) y el polimorfismo (intraespecifico) deben ser similares para cada locus. Se determinan teóricamente los valores de polimorfismo esperado en dos especies (o poblaciones) y el grado de divergencia esperado entre las dos especies y se comparan con los valores observados. Las desviaciones serán indicativas de que otros procesos adicionales a la deriva han afectado la evolución de esas secuencias. Puesto que la prueba se basa en obtener unos valores teóricos de polimorfismo y divergencia, este test es sensible a factores demográficos no contemplados en el modelo. Una ventaja de esta prueba es que es posible aplicarla a secuencias no codificadoras.

Análisis de la selección basada en campos aleatorios de Poisson

Las pruebas estadísticas anteriormente comentadas pueden determinar si las desviaciones frente a la hipótesis nula de evolución neutral son significativas e incluso el tipo de selección operante, pero más raramente determinan la intensidad de la selección. Sawyer y Hartl (1992) desarrollaron, a partir del modelo de evolución de Wright-Fisher, un marco teórico para estimar analíticamente la intensidad de la selección que afecta a una serie de variantes nucleotídicas: el modelo PRF (Poisson Random Field). Bajo una serie de asunciones (principalmente que las mutaciones surgen al azar en un curso temporal que sigue una distribución de Poisson, que cada mutación aparece en una posición diferente y que no hay ligamiento) es posible obtener estimas de la intensidad de selección. La potencia de esta aproximación es mayor cuando se utilizan datos de múltiples loci no ligados, como, por ejemplo, SNPs (polimorfismos de un solo nucleótido en su acrónimo en inglés) repartidos por todo el genoma. Esta aproximación se ha utilizado, por ejemplo, para cuantificar el papel de la selección en *Drosophila* y *Arabidopsis* (Bustamante et al. 2002), encontrando que en *Arabidopsis* se han acumulado mutaciones ligeramente deletéreas, posiblemente debido a su sistema de cruzamiento principalmente autógamo.

La unión de aproximaciones basadas en PRF con simulaciones por ordenador de procesos demográficos es actualmente la estrategia más completa para obtener información de la historia evolutiva de las poblaciones. La aplicación de los métodos ABC (approximate bayesian computation; Pritchard et al. 1999; Beaumont et al. 2002), junto con modelos PRF y la tecnología de secuenciación masiva de 2ª y 3ª generación, proporcionará una visión más realista del papel de la selección natural en el proceso evolutivo.

Diversos programas informáticos pueden ser muy útiles para dilucidar qué papel ha jugado la selección en el modelado de la diversidad nucleotídica. DnaSP (<http://www.ub.es/dnasp/>), VariScan (<http://www.ub.es/softevol/variscan/>), MEGA (<http://www.megasoftware.net/>), PAML (<http://abacus.gene.ucl.ac.uk/software/paml.html>), MKPRF (<http://cbsuapps.tc.cornell.edu/mkprf.aspx>) son, por citar unos pocos, programas muy utilizados para realizar varios de los tests anteriormente comentados.

Inferencia funcional y diversidad neutral

Con el avance de las técnicas moleculares que permiten comprender la base genética de la variabilidad de caracteres cuantitativos, el interés por determinar el papel de la selección natural y la deriva genética en el proceso de diferenciación poblacional se ha intensificado. No obstante, todavía estamos lejos de poder identificar nitidamente los genes que determinan caracteres cuantitativos de interés evolutivo y si la diferenciación poblacional en estos genes es debida a procesos selectivos o a procesos puramente demográficos, como la deriva genética. Por esta razón, la comparación entre la cantidad de diferenciación genética en caracteres cuantitativos (Q_{ST}) y la cantidad de diferenciación genética neutra (F_{ST}) representa un método apropiado para analizar el papel de la selección natural frente al de los procesos demográficos en la diferenciación poblacional de caracteres cuantitativos de interés evolutivo.

Cuando se comparan los índices Q_{ST} y F_{ST} se pueden obtener tres resultados. Primero, que $Q_{ST} > F_{ST}$. En este caso la diferenciación poblacional en caracteres cuantitativos es mayor que la diferenciación genética neutra (en función solamente de la deriva genética) y se asume que la selección natural direccional, que favorece a distintos fenotipos en distintas poblaciones, es la responsable de la diferenciación poblacional. Segundo, que $Q_{ST} < F_{ST}$. En este caso la diferenciación observada es menor que la que explica la diferenciación genética neutra y una explicación podría ser la selección estabilizadora, que favorece a los fenotipos más cercanos al promedio poblacional de manera que tiene un efecto homogeneizador. Finalmente, la tercera posibilidad es que $Q_{ST} \approx F_{ST}$. En este caso la cantidad de diferenciación debida a

procesos demográficos iguala a la diferenciación poblacional en caracteres cuantitativos pero no se puede decir que la deriva genética sea la única causa de dicha diferenciación.

Para estimar la cantidad de diferenciación genética en caracteres cuantitativos (Q_{ST}), lo más común es estimar los componentes de varianza entre y dentro de poblaciones mediante un análisis de la varianza (ANOVA) típico. En función del tipo de organismo de estudio y del grado de manipulación que estos permitan, los valores de Q_{ST} se pueden estimar a partir de datos puramente fenotípicos de poblaciones silvestres o de experimentos en condiciones controladas. En el primer caso, los Q_{ST} resultantes no permiten separar los componentes genéticos de los meramente ambientales, mientras que en el segundo caso los Q_{ST} se basan en los componentes genéticos solamente. En función del tipo de experimento, incluso es posible separar los efectos aditivos de los no aditivos, como los efectos maternos. Para estimar la cantidad de diferenciación genética neutra (F_{ST}) se pueden utilizar distintos marcadores moleculares neutros cuya variación entre y dentro de poblaciones se puede analizar mediante un análisis de varianza molecular (AMOVA).

Como en cualquier otra aproximación experimental, hay que tener en cuenta los principales problemas o limitaciones a la hora de comparar los valores de Q_{ST} y F_{ST} . En el caso de los Q_{ST} , los principales problemas son dos. El primero reside en cómo se estima la diferenciación genética en los caracteres cuantitativos de interés. Todo lo que no sea varianza genética aditiva (en particular, la varianza fenotípica debida a la respuesta plástica a distintos ambientes y la varianza debida a efectos maternos) puede enmascarar la verdadera divergencia poblacional en caracteres cuantitativos. El otro gran problema está en que generalmente la precisión y las propiedades estadísticas del índice Q_{ST} son bajas a no ser que se use un número elevado de poblaciones (la precisión del índice Q_{ST} mejora sensiblemente cuando el número de poblaciones se aproxima a 20). Las características del organismo de estudio van a permitir diseñar experimentos que incluyan un número alto de poblaciones y que estimen solamente los componentes de varianza genética aditiva para estimar la diferenciación poblacional de caracteres cuantitativos. En el caso de los F_{ST} , las estimas de diferenciación genética neutra dependerán del grado de polimorfismo de los marcadores que se utilicen. Cuanto mayor sea el polimorfismo de los marcadores menor será el grado de diferenciación poblacional genética neutra. Otro factor a tener en cuenta es que en la mayoría de casos se ignora, aunque se acabe asumiendo, si los marcadores están ligados a loci que sí están bajo selección o si las condiciones de equilibrio que se asumen para estimar los F_{ST} no se cumplen. Finalmente, las tasas de mutación intrínsecas a cada tipo de marcador también pueden afectar los F_{ST} , que nada tienen que ver con la interpretación puramente demográfica del índice F_{ST} .

A pesar de estas limitaciones, que siempre hay que tener en cuenta a la hora de diseñar un experimento para comparar valores de Q_{ST} y F_{ST} , los estudios de diferenciación poblacional para determinar el peso de la selección natural o de la deriva genética son claves para el avance de la biología evolutiva. Las estimas de Q_{ST} deben nutrirse de múltiples experimentos y observaciones de poblaciones naturales que al final nos permitan cuantificar y entender el grado de diferenciación poblacional en caracteres cuantitativos.

Por otro lado, las estimas de F_{ST} deben hacerse en base a una comparación entre tipos de marcadores para estimar el sesgo intrínseco a la naturaleza de cada tipo de marcador. No obstante, los avances en el campo de la genómica, que permiten trabajar a nivel de genoma incluyendo regiones genómicas codificantes, reguladoras y estructurales, van a cambiar definitivamente el concepto de la diferenciación poblacional genética neutra basado en los valores de F_{ST} .

Análisis genómico para el estudio del efecto de la selección

Gracias al avance de los últimos años que se ha experimentado en tres campos de investigación (biomedicina, genómica funcional aplicada a la agricultura y genética de poblaciones humanas), ha quedado patente la necesidad de desarrollar una línea de investigación en la que la genómica (y las demás “-ómicas” -ver Cieslak y Ribera 2009-) se aplique sistemáticamente al campo de la biología evolutiva. El principal objetivo es el de determinar el significado funcional de la variación genética que presentan las poblaciones de los organismos en sus ambientes naturales así como identificar los procesos evolutivos que generan o mantienen dicha variación. En este caso, la variación genética no hace referencia solamente a regiones neutras, sino que tiene que ver con la variación a nivel genómico con especial énfasis en aquellas regiones que pueden incluir genes de interés evolutivo. De todos modos, las técnicas que se desarrollan para estudiar dicha variación genética acaban generando marcadores con carácter funcional (regiones reguladoras de genes y regiones codificantes de proteínas) pero también neutro (regiones genómicas estructurales). Una de las consecuencias inmediatas de esta aproximación es que se pueden evitar las limitaciones severas que aparecen cuando se trabaja con un número reducido de marcadores supuestamente neutros para contestar a preguntas evolutivas.

Los estudios de genómica funcional aplicada a la biología evolutiva pivotan alrededor de los análisis de asociación entre la variación de caracteres cuantitativos de posible interés evolutivo y la variación genómica que presente el organismo de estudio. Para poder trabajar a nivel genómico se tienen que desarrollar un gran número de marcadores que cubran todas las regiones genómicas. Dos de las aproximaciones más comunes para la obtención de dichos marcadores incluyen, por una parte el desarrollo de SNPs (polimorfismos de un solo nucleótido en su acrónimo en inglés), y por otra el desarrollo de ESTs (marcadores de secuencia expresada en su acrónimo en inglés). Los SNPs son los polimorfismos más comunes que determinan la variación genética de un organismo. Por ejemplo, aproximadamente el 90% de la variación genética en humanos se debe a polimorfismos de tipo SNP.

Dado su gran número, podemos obtener SNPs en distintas regiones de un gen (SNPs funcionales) y en regiones genómicas estructurales (SNPs presuntamente no funcionales). En el caso de los ESTs, dado que se obtienen mediante la secuenciación de ARNm clonado, estos marcadores representan porciones de genes expresados y su carácter es mucho más funcional. A fecha de hoy y para un gran número de organismos representantes de toda la diversidad biológica, existen inmensas librerías de SNPs y ESTs (con miles o cientos de miles de marcadores en muchos casos) que sirven de base para el desarrollo de la genómica funcional aplicada a la biología evolutiva.

A parte de la naturaleza y cantidad de marcadores distribuidos por todo el genoma del organismo de estudio, sigue siendo capital tener un buen diseño experimental (que siempre que sea posible debe incluir el seguimiento de poblaciones de campo y la manipulación de poblaciones experimentales en condiciones controladas) para cuantificar la variación genética de caracteres cuantitativos de interés evolutivo. El reto ahora es el de desarrollar análisis de asociación entre caracteres cuantitativos y marcadores moleculares que nos permitan detectar las regiones genómicas responsables de la variación genética cuantitativa. Los métodos estadísticos para dichos análisis de asociación están en pleno desarrollo pues requieren en muchos casos el uso de herramientas como la biocomputación que permite trabajar con miles de marcadores a la vez. Los principales problemas a los que hay que hacer frente y que pueden sesgar los resultados de los análisis de asociación son la detección de falsos negativos/positivos (intrínsecos al uso de un número muy grande de marcadores), el grado de ligamiento entre marcadores muy relacionado con los grupos de ligamiento, y/o la estructura geográfica de la variación genética que presentan los datos como resultado de la historia de las poblaciones de estudio.

El desarrollo de proyectos de genómica funcional aplicados a la biología evolutiva requiere un cambio de mentalidad en todos los sentidos, incluyendo cuestiones meramente metodológicas, de formación de nuevos profesionales, y otras de tipo más económico y de estrategia pues, en el caso de trabajar con especies no modelo, la inversión inicial puede ser importante.

Ese es el cambio más importante que hay que hacer: decidir cómo y cuando invertir los recursos para generar herramientas que luego permitan responder a preguntas mucho más ambiciosas dentro del campo de la biología evolutiva.

El 1 de febrero de 1898, Hermon Bumpus recibió 136 gorriones recolectados tras una gélida tormenta que azotó Rhode Island (USA). Casi la mitad de ellos sobrevivió y Bumpus estudió las características físicas de todos ellos para analizar que rasgos habían intervenido en la supervivencia frente al frío (Bumpus 1899). Este trabajo se convirtió en un clásico de las estimas de selección natural y todavía esos datos son analizados de diferentes formas (eg, Pugesek y Tomer 1996). Es posible imaginar que gran información podría haber obtenido Hermon Bumpus de haber dispuesto de herramientas como la tecnología de secuenciación masiva y las aproximaciones teóricas y computacionales de la moderna teoría de evolución molecular. Con estos desarrollos, el papel de la selección natural en la evolución de las poblaciones pronto podrá ser valorado de forma precisa para muchos organismos, pasando de simples aproximaciones al papel de los diferentes procesos evolutivos a tener estimas robustas de su importancia en la historia evolutiva de las especies.

Referencias

- Beaumont, M.A., Zhang, W., Balding, D.J. 2002. Approximate Bayesian computation in population genetics. *Genetics* 162:2025-2035.
- Bromham, L., Penny, D. 2003. The modern molecular clock. *Nature Review Genetics* 4:216-224.
- Bumpus, H. 1899. The Elimination of the unfit as illustrated by the introduced sparrow, *Passer domesticus*. *Biological Lectures Delivered at the Marine Biological Laboratory, Woods Hole* 1898-4:209-226.
- Bustamante, C.D., Nielsen, R., Sawyer, S.A., Olsen, K.M., Purugganan, M.D., Hartl, D.L. 2002. The cost of inbreeding in *Arabidopsis*. *Nature* 416:531-534.
- Chamary, V., Hurst, L.D. 2005. Evidence for selection on synonymous mutations affecting stability of mRNA secondary structure in mammals. *Genome Biology* 6:R75.

- Cieslak, A., Ribera, I. 2009. Aplicaciones de proteómica en ecología y evolución. *Ecosistemas* 18(1):34-43.
- Drummond, D.A., Wilke, C.O. 2008. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* 134:341-352.
- Fay, J.C., Wu, C.I. 2000. Hitchhiking under positive Darwinian selection. *Genetics* 155:1405-1413.
- Fu, Y.X., Li, W.H. 1993. Statistical tests of neutrality of mutations. *Genetics* 133:693-709.
- Hudson, R.R., Kreitman, M., Aguadé, M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* 116:153-159.
- Kimura, M. 1968. Evolutionary rate at the molecular level. *Nature* 217:624-626.
- Kimura, M. 1977. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* 267:275-276.
- King, C.E., Jukes, T.H. 1969. Non-Darwinian evolution. *Science* 164:788-798.
- McDonald, J.H., Kreitman, M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351:652-654.
- Nielsen, R. 2005. molecular signatures of natural selection. *Annual Review of Genetics* 39:197-218.
- Ohta, T. 1973. Slightly deleterious mutant substitutions in evolution. *Nature* 246:96-98.
- Ohta, T. 1987. Very slightly deleterious mutations and the molecular clock. *Journal of Molecular Evolution* 26:1-6.
- Ohta, T. 1992. Theoretical study of near neutrality II. Effect of subdivided population structure with local extinction and recolonization. *Genetics* 130:917-923.
- Pal, C., Papp, B., Lercher, M.J. 2006. An integrated view of protein evolution. *Nature Reviews Genetics* 7:337-348.
- Parmley J.L., Chamary J.V., Hurst L.D. 2006. Evidence for purifying selection against synonymous mutations in mammalian exonic splicing enhancers. *Molecular Biology and Evolution* 23:301-309.
- Pugesek, B.H., Tomer, A. 1996. The Bumpus house sparrow data: a reanalysis using structural equation models. *Evolutionary Ecology* 10(4):387-404.
- Pritchard, J.K., Seielstad, M.T., Pérez-Lezaun A., Feldman M.W. 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution* 16:1791-1798.
- Sawyer, S.A., Hartl D.L. 1992. Population genetics of polymorphism and divergence. *Genetics* 132:1161-1176.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.