

## CAZANDO EL GENOMA DE UN CAZADOR: LA GENÓMICA AL RESCATE DEL LINCE IBÉRICO

José A. Godoy. Estación Biológica Doñana, CSIC. C/Américo Vespucio, s/n. 41092 Sevilla. [godoy@ebd.csic.es](mailto:godoy@ebd.csic.es)

La aparición de las tecnologías de secuenciación de nueva generación ha ampliado sustancialmente tanto el rango de especies que pueden ser caracterizadas a un nivel genómico como el de problemas a los que aplicar estos nuevos recursos. Después de que se culminase el gran hito de la secuenciación de genoma humano, que requirió colosales esfuerzos y una gran inversión económica con las tecnologías existentes en esa época, la secuenciación de genomas completos fue extendiéndose lentamente a organismos modelo para la investigación (e.g. ratón, mosca del vinagre) y algunos organismos con algún interés económico (cultivos, ganado, ...). Pero con las tecnologías actuales, la secuenciación de un genoma nuevo partiendo de cero puede costar poco más de cien mil euros y la re-secuenciación de genomas ya caracterizados ronda los pocos miles de euros. Por ello, ahora empiezan a hacerse factibles empresas que antes resultaban imposibles, como la secuenciación de genomas individuales para medicina personalizada, o la secuenciación de diez mil genomas de vertebrados para conocer el origen de las fascinantes adaptaciones que han evolucionado en estos organismos (Hausser *et al.* 2009). Aun así, no deja de ser llamativo que la primera especie de vertebrado en ver su genoma secuenciado en España no es un modelo de investigación biomédica y carece aparentemente de interés económico: el lince ibérico, *Lynx pardinus*.

La verdad es que ninguna de las dos afirmaciones anteriores es totalmente cierta: el lince tiene un enorme valor ecológico y simbólico, y constituye un buen modelo de investigación en el ámbito de la biología de la conservación. El lince ibérico es una especie clave del ecosistema en que habita, el monte mediterráneo, y su presencia añade valor ecológico al entorno y contribuye de manera fundamental a su conservación. Pensando en aspectos muy directamente relacionados con actividades humanas como la caza, el lince mantiene a raya a otros depredadores más dañinos para el hombre como el zorro o el meloncillo, por lo que ha de verse más como un aliado que como un competidor en este ámbito. Es uno de los múltiples endemismos de la península ibérica, que se ha diferenciado de sus hermanos los linceos boreales (*Lynx lynx*) en su adaptación al clima mediterráneo y su especialización en el consumo de conejos. Es muy complicado –y algo arriesgado– establecer un valor económico concreto a la existencia de una especie, pero el lince ibérico es sin duda un patrimonio de nuestra diversidad natural, para la que se ha convertido en símbolo y estandarte. Y este patrimonio está en grave riesgo de perderse.

El lince ibérico es el felino más amenazado del planeta y uno de las cinco especies de mamíferos europeos catalogados como en peligro crítico por la Unión Internacional para la Conservación de la Naturaleza (IUCN). La especie ha sufrido un intenso declive durante la última mitad del siglo XX, debido a la destrucción de su hábitat natural, el monte mediterráneo, y la intensa persecución directa a la que fueron sometidos los predadores, incluso con incentivos económicos de programas estatales para el control de “alimañas”. A estas causas, que han afectado también a muchas otras especies, se ha sumado en el caso del lince la reducción de su presa principal, el conejo, por sucesivas epidemias virales. La acción combinada de estos factores relegó la especie a sólo dos

poblaciones aisladas (Doñana y Andújar) con menos de 100 individuos en total en 2002. Los esfuerzos de conservación iniciados en esa época, financiados por la Unión Europea a través de sucesivos proyectos LIFE, han conseguido evitar la extinción por el momento, estabilizando en primera instancia las dos poblaciones remanentes e incrementando posteriormente el censo global hasta superar los 300 ejemplares en libertad en 2011 (Simon *et al.* 2012). Al mismo tiempo, se ha conseguido establecer una población cautiva que supone una salvaguarda de la especie y que está aportando ejemplares para el refuerzo de las poblaciones silvestres y para actuaciones de reintroducción en áreas de presencia histórica (Vargas *et al.* 2009). Estas ya han comenzado en Andalucía (Gualmellato, Córdoba, y Guarrizas, Jaen) y están programadas en los próximos años en Extremadura, Castilla-La Mancha, Murcia y Portugal bajo el nuevo proyecto LIFE “Iberlince” (<http://www.lifelince.org>).

El éxito de los esfuerzos que se están realizando para la conservación del lince ibérico pueden, sin embargo, quedar comprometidos por la amenaza oculta de la erosión genética. Las poblaciones que se hacen pequeñas y se aíslan de otras poblaciones pierden diversidad genética, acumulan genes defectuosos, y se hacen endogámicas. El resultado es que estas poblaciones ven reducida su viabilidad, al quedar afectadas negativamente la supervivencia y la reproducción, y su capacidad para adaptarse a cambios ambientales. Estos cambios genéticos sobrevenidos pueden precipitar la extinción de la especie, incluso cuando las causas últimas del declive hayan desaparecido, debido a la retroalimentación positiva que se establece entre el deterioro demográfico y deterioro genético, una dinámica que se ha descrito como “el vórtice de la extinción” (Allendorf *et al.* 2012; Frankham *et al.* 2010). Y esto es lo que puede estar sucediendo con el lince ibérico. Las dos poblaciones de lince que han sobrevivido hasta la fecha muestran signos de erosión genética y varios indicios apuntan a que puede estar comprometida la reproducción y la supervivencia, sobre todo en Doñana (Casas-Marce *et al.* 2013; Palomares *et al.* 2012; Ruiz-Lopez *et al.* 2012). No es fácil cuantificar el lastre que la genética supone para la recuperación de la especie, pero este puede ser evaluado mediante estudios genéticos y minimizado con una adecuada gestión genética. Para esto, los marcadores moleculares como los microsatélites han sido hasta ahora las herramientas básicas (Godoy 2009), pero los nuevos enfoques de la genómica prometen un salto cualitativo en la manera de estudiar la genética de las especies amenazadas. El salto de unos pocos a marcadores neutrales a miles de variantes con posición y efectos funcionales conocidos no sólo aumentará de manera espectacular la precisión y resolución de los análisis, sino que dará acceso a la hasta ahora inexplorada fracción funcional del genoma (Allendorf *et al.* 2010).

De estos antecedentes surgió la propuesta del proyecto “Desvelando el genoma del lince ibérico”, financiada por la Fundación General CSIC, con una contribución especial del Banco de Santander, en su convocatoria de Proyectos Cero sobre “Especies Amenazadas” del año 2010 (Gabaldón 2012; Godoy 2010) (<http://www.lynxgenomics.es>). La solicitud inicial fue impulsada por cuatro institutos de tres instituciones distintas: la Estación Biológica Doñana y el Centro de Investigaciones Biológicas, ambos del CSIC, el Centro Nacional de Análisis Genómico, el Centro de Regulación Genómica. El objetivo general del proyecto es el de generar recursos genómicos para la investigación y la conservación de la especie y el aprovechamiento de estos recursos para estudiar la evolución y la historia demográfica de la especie, así como para evaluar el impacto del declive sobre su variación genómica.

A continuación se resumen los objetivos particulares del proyecto y los resultados hasta la fecha:

*Secuenciación y ensamblaje de un genoma de referencia.* Un primer objetivo del proyecto es el de generar el genoma de referencia de la especie. El objetivo es reconstruir la secuencia del conjunto del genoma, de todos y cada uno de los cromosomas, de la manera más completa y contigua posible. El uso de tecnologías de secuenciación de segunda generación supone una ventaja evidente debido a su alta productividad y su bajo coste por base secuenciada, pero las secuencias se obtienen en pares de bloques de 100 bases contiguas provenientes de cualquier lugar al azar del genoma y que se encuentran separadas por una distancia conocida. Ensamblar estas secuencias en fragmentos mayores aprovechando los solapamientos entre estas lecturas cortas es todo un reto que exige eficientes algoritmos matemáticos y potentes ordenadores. Para asistir en esta tarea, el proyecto ha optado por la generación de pares de secuencias separadas entre si por distintas distancias (500, 4.000, 8.000, 40.000 bases), una estrategia habitual en la secuenciación *de novo* de genomas. Además, hemos usado una estrategia complementaria consistente en reducir la complejidad del problema, secuenciando y ensamblando una pequeña fracción del genoma cada vez. Para ello se han generado colecciones de fragmentos de 40,000 bases insertados en vectores de clonación (f se han secuenciado y ensamblado 96 conjuntos aleatorios independientes con 1.200 de estos fragmentos cada uno. Por último, se ha aprovechado la información obtenida de la secuenciación de RNAs y la secuencia genómica del gato doméstico para generar un borrador con una longitud total de ca. 2,5 gigabases (la longitud estimada del genoma del lince es de aproximadamente 2,8 gigabases), en el que más de la mitad del genoma está representado en fragmentos contiguos de más de 1,5 megabases. Estos resultados se consideran bastante buenos, especialmente cuando se considera la moderada inversión realizada en la generación de datos, lo que sugiere que la estrategia de secuenciación de conjuntos aleatorios de fósmidos puede convertirse en una estrategia ventajosa para la secuenciación de genomas complejos.

*Anotación del genoma.* El ensamblaje proporciona un mapa mudo del genoma al que, para que sea realmente útil, hay que añadir leyendas que identifiquen la localización de genes y otros elementos funcionales y estructurales. Los genes se reconocen utilizando tres estrategias distintas: i) identificando patrones de secuencia característicos de los genes (predicción *ab initio*), ii) por sus similitudes con genes identificados en otras especies y iii) por coincidir con las secuencias obtenidas de RNAs, las moléculas intermediarias en la expresión de los genes. La aplicación de estas estrategias ha permitido identificar algo más de veinte mil genes putativos. El trabajo en curso pretende identificar la función de cada uno de estos genes y asignarles atributos funcionales como los procesos celulares en los que participa, la localización intracelular de su producto o su función molecular (i.e. ontología génica).

*Genómica comparada.* La comparación del genoma del lince con el del gato doméstico, el perro y otros mamíferos más lejanamente emparentados, permitirá identificar las características compartidas por distintas especies y aquellas que son exclusivas de una especie en particular. Este análisis puede revelar por tanto los cambios evolutivos que subyacen a las peculiaridades y adaptaciones características que han ido diferenciando unas especies de otras y, en particular, las que definen al lince ibérico como especie. En este proceso se identifican los genes que han estado sometidos a una selección natural

intensa en la especie, así como las familias génicas que han sufrido expansión o contracción en el linaje de los linces.

*Genómica poblacional.* El proyecto ha secuenciado con una menor cobertura diez individuos de lince ibérico adicionales y un lince boreal, la especie más cercana al lince ibérico. A partir de estos datos se han identificado más de un millón de variantes de secuencia que están segregando en las poblaciones de lince, incluyendo polimorfismos de un solo nucleótido (*Single Nucleotide Polymorphisms*, SNPs), inserciones y deleciones y otras variantes complejas. El análisis de estos datos aportará una imagen de muy alta resolución de los patrones genómicos contemporáneos y estimas precisas de diversidad genética, consanguinidad y diferenciación entre poblaciones. Los resultados confirman los obtenidos con unos cuantos marcadores moleculares, al indicar una importante erosión genética en las poblaciones de lince. Pero lo que es más importante, también revelarán qué regiones genómicas y funciones génicas han quedado más afectadas y ayudarán a identificar la base genética responsable de caracteres deletéreos que pueden ser responsables de la pérdida de viabilidad que comienza a documentarse en la especie. Además, a partir de los datos obtenidos de genomas individuales se podrán reconstruir los cambios demográficos sufridos por la especie a lo largo de su historia y estimar cuándo y cómo se separó del lince boreal. En un aspecto más aplicado, cada una de las variantes encontradas aporta un marcador genético para el seguimiento de individuos y poblaciones y para la gestión genética de la especie. El proyecto seleccionará 1500 de estos marcadores para caracterizar el conjunto de la población cautiva y muestras representativas de las poblaciones silvestres.

El del lince ibérico se convierte así en el primer genoma animal complejo en ser secuenciado íntegramente en España y uno de los primeros genomas de especies amenazadas en ser secuenciado en el mundo. Los recursos generados serán fundamentales para la investigación de la especie y los análisis realizados arrojarán luz sobre su origen, su evolución y su historia demográfica. Lo que es más importante, nos permitirán caracterizar, con un nivel de detalle antes inimaginable, los cambios genéticos que la especie ha sufrido como consecuencia del declive poblacional. Es de esperar que esta información contribuya de manera significativa a la conservación de una especie que es emblema de la biodiversidad en España y en Europa. Además, lo que aprendamos será relevante para la conservación de las numerosas especies que se encuentran en una situación similar y la experiencia acumulada será sin duda útil para la caracterización de nuevos genomas. El objetivo de salvar la biodiversidad requiere esfuerzos enormes y la contribución de múltiples áreas de conocimiento, pero en esta empresa contamos ahora con la genómica como un nuevo y poderoso aliado.

## Referencias

- Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics* **11**, 697-709.
- Allendorf FW, Luikart G, Aitken SN, ebrary Inc (2012) Conservation and the genetics of populations. In: *E-Libro*, pp. xviii, 602, [608] p. John Wiley & Sons, Chichester, West Sussex.
- Casas-Marce M, Soriano L, López-bao JV, Godoy JA (2013) Genetics at the verge of extinction: insights from the Iberian lynx. *Molecular Ecology*, n/a-n/a.

- Frankham R, Ballou JD, Briscoe DA (2010) *Introduction to conservation genetics*, 2nd edn. Cambridge University Press, Cambridge ; New York.
- Gabaldón T (2012) El lince ibérico entra en la era de la genómica. *Lychnos*, 14-19.
- Godoy JA (2009) La genética, los marcadores moleculares y la conservación de especies. *Ecosistemas* **18**, 23-33.
- Godoy JA (2010) La genómica y la conservación del lince ibérico. *Lychnos*, 14-18.
- Hausler D, O'Brien SJ, Ryder OA, *et al.* (2009) Genome 10K: A Proposal to Obtain Whole-Genome Sequence for 10 000 Vertebrate Species. *Journal of Heredity* **100**, 659-674.
- Palomares F, Godoy JA, Lopez-Bao JV, *et al.* (2012) Possible Extinction Vortex for a Population of Iberian Lynx on the Verge of Extirpation. *Conservation Biology* **26**, 689-697.
- Ruiz-Lopez MJ, Ganan N, Godoy JA, *et al.* (2012) Heterozygosity-Fitness Correlations and Inbreeding Depression in Two Critically Endangered Mammals. *Conservation Biology* **26**, 1121-1129.
- Simon MA, Gil-Sanchez JM, Ruiz G, *et al.* (2012) Reverse of the Decline of the Endangered Iberian Lynx. *Conservation Biology* **26**, 731-736.
- Vargas A, Breitenmoser C, Breitenmoser U (2009) Iberian Lynx Ex situ Conservation: An Interdisciplinary Approach, p. 556. Fundación Biodiversidad in collaboration with: International Union for Conservation of Nature, Species Survival Commission (IUCN /SSC ) Cat Specialist Group, Madrid.