TESIS DOCTORAL 2020

María Lucena Pérez

LINCES
EN EL PALEÁRTICO
VARIACIÓN GENÓMICA EN POBLACIONES PASADAS Y PRESENTES

Estación Biológica Doñana CSIC

UNIVERSIDAD PABLO DE OLAVIDE SEVILLA AÑO 1997

Tesis doctoral

Linces en el paleártico: Variación genómica en poblaciones pasadas y presentes

Memoria presentada por

María Lucena Pérez
para optar al título de Doctora
por la Universidad Pablo de Olavide.

Fdo. María Lucena Pérez

Con el visto bueno de

Dr. José Antonio Godoy López
Director y Tutor.
Investigador en la Estación Biológica de Doñana del Consejo Superior de Investigaciones Científicas.
Profesor asociado en el Departamento de Biología Molecular e Ingeniería Bioquímica de la Universidad Pablo de Olavide.

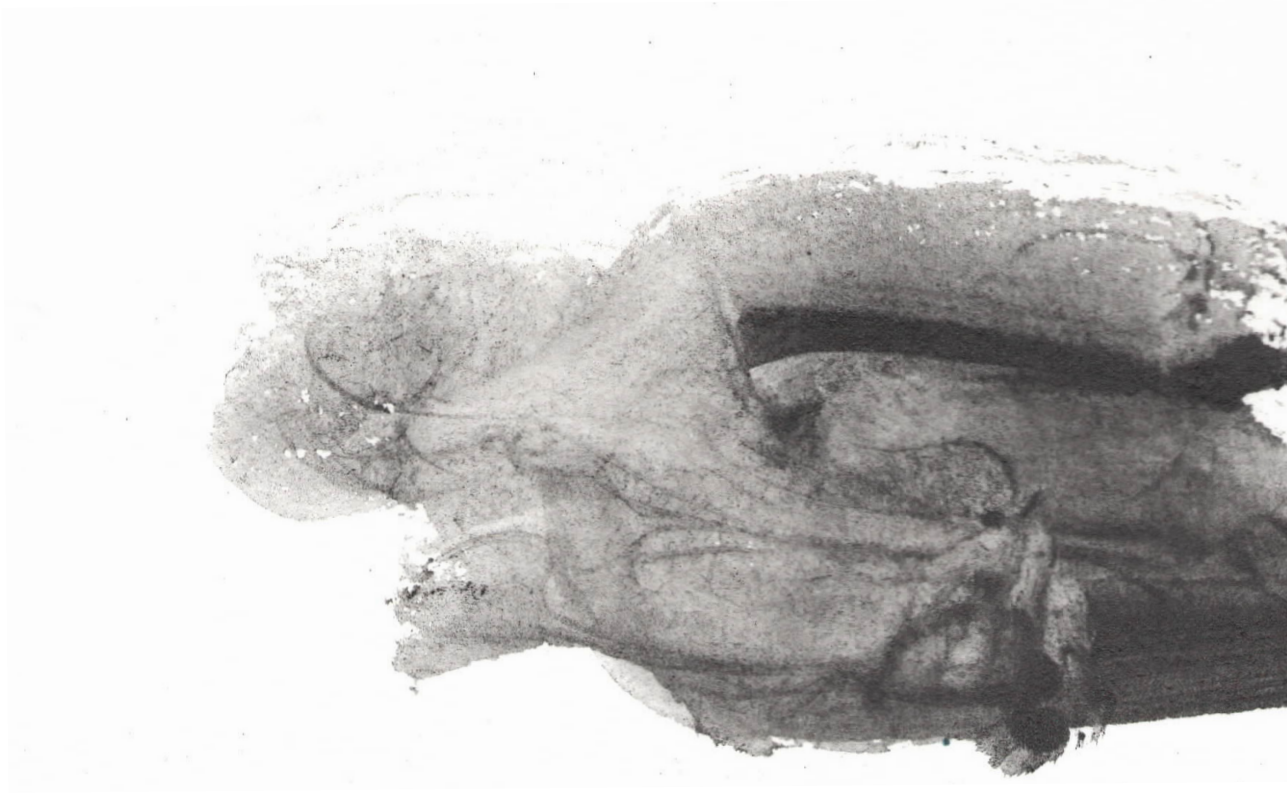
Los trabajos que contiene esta memoria fueron financiados a través de la Dirección General de Investigación Científica y Técnica del Ministerio español de Ciencia e Innovacion (MICINN), anteriormente Ministerio de Economía y Competitividad (proyectos CGL2013-47755-P, CGL2017-84641-P y CGL2014-55266-P); por el Centro Nacional de Ciencia de Polonia (proyectos 2014/15/B/ NZ8/00212), por la Fundación de Ciencia Rusa (projecto No. 18-14-00093), por la Fundación BBVA y por la Comisión Europea en el marco de acciones Marie Skłodowska-Curie del septimo programa para la investigación y el desarrollo tecnológico a través del proyecto BIOGEAST (PIRSESGA-2009-247652).

La autora, María Lucena Pérez, recibió financiación del programa de ayudas de Fundación "La Caixa" (ID 100010434) para contratos de formación predoctoral en Centros de Excelencia Severo Ochoa (SEV-2012-0262) según el acuerdo LCF/BQ/SO14/52250035. Además, recibió financiación del programa de creación de redes de investigación de la Fundación Europea de Ciencia (European Science Foundation Reseach Networking Programmes - ConGenOmics) destinada a una estancia internacional de tres meses en la Universidad de Potsdam (Alemania), y una beca de mobilidad (Godfrey Hewitt Mobility Award) de la Sociedad Europea de Biología Evolutiva (European Society of Evolutionary Biology) para una estancia breve de un mes en el Instituto de Genética del University College de Londres (UCL) (Reino Unido).

Citación Recomendada

Lucena-Pérez, María (2020). Linces en el paleártico : Variación genómica en poblaciones pasadas y presentes. Tesis doctoral. Universidad de Pablo de Olavide, Sevilla, España.

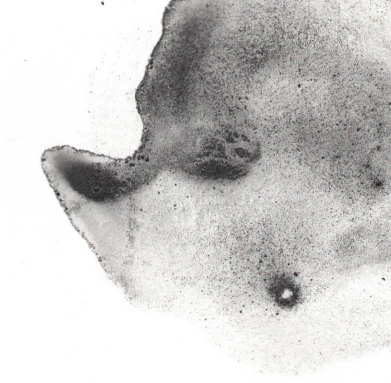
Diseño maquetación por Manolo Domínguez Ramírez



A mi familia,
la que me tocó,
la que he formado.

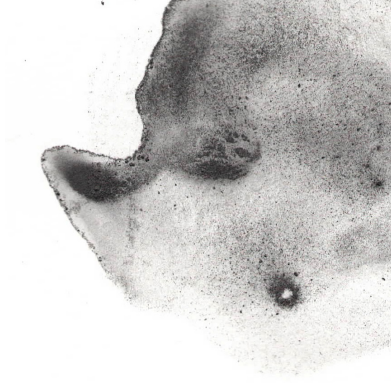
A papá.

ÍNDICE



Págs.

RESUMEN



Distintas fuerzas evolutivas, como migración, mutación, flujo génico, selección y deriva, actúan en las poblaciones provocando cambios en frecuencias alélicas que se acaban consolidando como divergencias nucleotídicas, algunas de las cuales pueden sustentar divergencias adaptativas. La acumulación de estos cambios evolutivos termina diferenciando genéticamente las poblaciones y generando linajes y, eventualmente, especies. A lo largo de este continuo, eventos naturales, como las fluctuaciones climáticas y antrópicos alteran la demografía de estas especies y el equilibrio entre las fuerzas evolutivas y generan patrones genómicos, que nos informan sobre la historia evolutiva del linaje. En esta tesis, usamos nuevos desarrollos metodológicos en el área de la genómica y el ADN antiguo para ahondar en el conocimiento de la historia evolutiva de las dos especies de lince del paleártico, el lince boreal (*Lynx lynx*) y el lince ibérico (*Lynx pardinus*), que a pesar de presentar rasgos de vida similares, tienen historias demográficas, en apariencia, muy dispares. El lince boreal se considera una especie de preocupación menor en cuanto a conservación y tiene una distribución muy amplia en Eurasia, mientras que el lince ibérico, distribuido en un par de poblaciones al sur de la península ibérica, ha sido considerado el felino más amenazado del mundo. Específicamente, en esta tesis, vamos a describir los patrones genómicos de poblaciones presentes y pasadas de ambas especies para reconstruir su historia evolutiva, e investigar el rol que distintas fuerzas evolutivas, cambios demográficos y factores extrínsecos han jugado en el moldeado de sus genomas a lo largo del tiempo.

En el **capítulo 1** el uso de 80 genomas completos de lince boreal distribuidos a lo largo de su área de distribución en Asia y Europa, nos permite caracterizar los patrones biogeográficos de la especie hoy día y relacionarlos con eventos naturales y antrópicos que han ocurrido a lo largo del final del Pleistoceno y el Holoceno. Concretamente, descubrimos que las poblaciones de esta especie han tenido una historia común hasta hace aproximadamente 100 mil años. A partir de este punto,

empezaron a divergir las poblaciones de Asia y Europa y la especie entró en un declive continuado y generalizado, en el que las poblaciones más occidentales mantuvieron tamaños efectivos menores que las poblaciones orientales. Estos declives, y el aislamiento de las poblaciones que conllevaron, estuvieron detrás de la diferenciación genética que se observa a día de hoy entre poblaciones del centro de Europa que se encuentran física y ecológicamente muy próximas. Por el contrario, las poblaciones asiáticas son relativamente homogéneas, mostrando un patrón de aislamiento por distancia, a pesar de los hábitats tan diferentes que engloban. Nuestros resultados apuntan que tanto las divergencias mitogenéticas como nucleares podrían tener su origen en las fluctuaciones climáticas acaecidas durante el Pleistoceno, y en los impactos antropogénicos que se intensificaron progresivamente durante el Holoceno.

A lo largo del **capítulo 2** el análisis de una muestra de lince boreal residente en la península ibérica hace 2 mil años, junto con el reanálisis de fragmentos mitocondriales de muestras antiguas, aporta luz a los patrones biogeográficos de la especie en el pasado en el oeste de Europa. Específicamente, los resultados sugieren que la población, que entró en Iberia en el cambio del Pleistoceno al Holoceno y que vivió hasta su extinción en el siglo XX, formaba parte de un linaje próximo y posiblemente ancestral al que se encuentra hoy en Cárpatos y los países bálticos. Esta muestra presenta la diversidad más baja reportada hasta la fecha para lince boreal sugiriendo que la población sufría un empobrecimiento genético que pudo precipitar su extinción, seguramente debido a la combinación de una serie de eventos fundadores y una presión antrópica intensificada durante el Holoceno. Además, la distribución de haplotipos mitocondriales en el espacio y el tiempo apuntan a la coexistencia de distintos linajes, que pudieron convivir en Europa aunque con ciertas fluctuaciones en su distribución. Mientras que un linaje más relacionado con Balcanes y Cáucaso fue predominante en el Pleistoceno En el suroeste de Europa, otro linaje, relacionado con Europa central y Asia lo fue durante el Holoceno.

En el **capítulo 3** el uso de muestras antiguas, de 2 a 4 mil años de antigüedad nos permite caracterizar genéticamente la población que existió en la península ibérica en ese periodo, y los cambios que ha sufrido en su historia reciente. Específicamente, encontramos que esta población era relativamente homogénea

y cercana a la población de Andújar y que tenía una diversidad genética más baja que las poblaciones actuales de la especie. El aumento de la diversidad con el tiempo ha sido concomitante con un proceso de flujo génico con lince boreal que se registra por igual en las dos poblaciones remanentes de lince ibérico. Esta señal de introgresión es mayor con linces boreales occidentales que con linces orientales, pero similar entre linces de distintas poblaciones occidentales. La muestra de lince boreal residente en la península ibérica hace dos mil años y analizada en el capítulo anterior, muestra, al contrario de lo esperado, menos alelos compartidos con lince ibérico que las poblaciones actuales de lince boreal, indicando que este linaje no estuvo implicado en este proceso de introgresión y que fue posiblemente otro linaje hibrido y ancestral de las poblaciones contemporáneas la fuente de la introgresión. Además, encontramos que este intercambio fue aumentando a lo largo del tiempo, de modo que las muestras más recientes, muestran un mayor número de alelos compartidos. Este intercambio, aunque seguramente está detrás del aumento de diversidad en la especie, pudo haber tenido consecuencias desconocidas en cuanto a eficacia biológica y demografía.

En el **capítulo 4** se comparan poblaciones que han sufrido un cuello de botella muy pronunciado, frente a otras que se han mantenido demográficamente más estables tanto de lince boreal como de ibérico, para investigar los patrones que dejan en el genoma los cuellos de botella. Encontramos que, como esperamos, la intensidad y/o duración del cuello de botella está negativamente correlacionada con la diversidad de las poblaciones. Las correlaciones, esperadas y conocidas, de la diversidad con numerosas variables se observan en las poblaciones no sometidas a cuello de botella intenso, pero son más débiles en poblaciones que han sufrido cuello de botella intenso. Además, zonas del genoma que se encuentran bajo selección natural, relacionadas con la estructura proteica pero también con la regulación, así como el cromosoma X, muestran una acumulación de alelos a baja frecuencia en las poblaciones que han sufrido un cuello de botella, de forma que incluso llegan a superar en diversidad a las poblaciones demográficamente más estables y grandes. Esta acumulación se concentra en zonas de alta mutación, en las que algunas variantes segregando a muy baja frecuencia aumentarían en frecuencia tras el cuello de botella hasta hacerse perceptibles, lo que parece estar relacionado con una relajación de la selección purificadora.

ABSTRACT

Different evolutionary forces, such as migration, gene flow, selection and drift, shape populations by changing their allele frequencies that eventually leads to nucleotide divergences, some of which might be adaptive. These evolutionary changes accumulate and differentiate populations, eventually generating lineages and species. Along this continuum, natural events, such as climatic fluctuations, and human activities change the species demography and disrupt the equilibrium of the evolutionary forces, generating genomic patterns that inform of the evolutionary history of the lineage. In this thesis, using new methodological advances in genomics and ancient DNA, we deepen into the evolutionary history of the two paleartic lynx species, Eurasian lynx (*Lynx lynx*) and Iberian lynx (*Lynx pardinus*). These two species, despite their similar life traits, show demographic histories apparently very different. The Eurasian lynx is considered a Least Concern species by the IUCN, and has a wide distribution in Eurasia, while the Iberian lynx distribution encompasses only a couple of populations in the south of the Iberian Peninsula, and was considered the most endangered felid in the world. Here, we specifically describe genomic patterns of the present and past populations from both species to reconstruct their evolutionary history and discuss the role that different evolutionary forces, demographic change and extrinsic factors have played in shaping their genomes throughout time.

In **chapter 1**, the use of 80 whole genome sequences from Eurasian lynx, distributed along its range in Asia and Europe, allowed us to characterize biogeographic patterns of the species and link them to natural and anthropogenic changes that occurred during the Late Pleistocene and the Holocene. Precisely, our results show that Eurasian lynx populations shared a common history until 100 kya, when Asian and European populations started to diverge and both entered a period of continuous and widespread decline, with western populations maintaining lower effective sizes than eastern populations. Population declines and increased isolation in more recent times likely drove the genetic differentiation between geographically and ecologically close westernmost European populations. By contrast, and despite the wide range of habitats covered, populations are quite homogeneous genetically across the Asian range, showing a pattern of isolation by distance. Mitogenomic and nuclear divergences and population declines can be mostly attributed to climatic fluctuations during the Pleistocene, and to anthropogenic impacts, which intensified during the Holocene.

Throughout **chapter 2**, the analysis of a 2 kyo Eurasian lynx sample from the Iberian Peninsula, together with the reanalysis of mitochondrial fragments from ancient samples, shed light on the past biogeographic patterns of the species in Western Europe. Precisely, our results suggest that the Iberian population, which entered in the Peninsula during the Pleistocene-Holocene boundary and remained there until the XX century, is part of an extinct European lineage closely related and probably ancestral to current Carpathian-Baltic lineages. Also, this sample holds the lowest Eurasian lynx diversity reported so far for the species, suggesting a genetic impoverishment of the population that could precipitated its extinction, likely due to a combination of historical factors, such as a founder effect and human impacts that intensifying during the Holocene. Mitogenomic lineages distribution in space and time supports the long-term coexistence of several lineages of Eurasian lynx in Western Europe with fluctuating ranges. While the mitogenomic lineage related to the one currently found in Balkans and Caucasus was predominant during the Pleistocene, another lineage related to the one currently distributed in Central Europe prevailed during the beginning of the Holocene.

In **chapter 3**, the usage of 2-4 kyo samples allowed us to genetically characterize the Iberian lynx populations that inhabited the peninsula during that time, and the changes that they went through till the present. The ancient population, which is relatively homogeneous and closer to Andújar population, harboured a lower diversity than the highly eroded contemporary populations. The increase in diversity in the species is concomitant with an inferred gene flow event with European Eurasian lynx that is equally recorded in the two remnant Iberian lynx populations. The signal of introgression into Iberian lynx is higher from current western than eastern Eurasian lynx populations, but is similar across the former. The 2 kyo Eurasian lynx from the Iberian Peninsula, analysed in the previous chapter, showed less shared allele with current Iberian lynx than modern European Eurasian lynx populations, indicating that this particular Eurasian lynx lineage was not involved in the recent admixture process. Also, we found an increasing signal of introgression through time, with more recent, although ancient, individuals showing more shared alleles with Eurasian lynx. The admixture with Eurasian lynx could have increased the extremely low genetic diversity of the species, but with unknown consequences on fitness and demography.

Throughout **chapter 4** we compared bottlenecked populations with demographically more stable ones from the two different species, Iberian and Eurasian lynx, to unveil how demographic bottlenecks shape genomic patterns. As expected overall genomic diversity correlated negatively with bottleneck intensity and/or duration. Correlations of genetic diversity with several known correlates were observed in non-bottlenecked populations, but were weaker in bottlenecked populations. Also, the more selectively constrained features related to protein structure but also regulation, and the X chromosome showed an accumulation of low frequency alleles in bottlenecked populations, even resulting in higher θ_W diversity than in non-bottlenecked populations. This accumulation seems to be related to both a higher mutational input in those regions creating a large collection of low-frequency variants, a few of which increase in frequency during the bottleneck, and the relaxation of purifying selection.

INTRODUCCIÓN

LOS PROCESOS QUE MOLDEAN LOS GENOMAS: DE LA HETEROGENEIDAD A LO LARGO DEL GENOMA A LA HETEROGENEIDAD ENTRE POBLACIONES

La historia evolutiva de las especies está escrita en el genoma de los individuos que las conforman. Las fuerzas evolutivas que actúan sobre las poblaciones generan una heterogeneidad en los patrones genéticos que estas presentan. Además, como algunos de estos procesos tienen una acción más local en el genoma (como la selección, o la mutación), mientras que otros tienen tiene un efecto global (como la deriva o el flujo genético), van a generar también una heterogeneidad a lo largo del genoma.

De forma muy resumida y tomando como referencia trabajos clásicos de evolución molecular y genética de poblaciones, podemos considerar que las secuencias de ADN están sometidas a la acción de la mutación que va a aportar el sustrato para los demás procesos evolutivos mediante cambios en la secuencia de ADN (Dobzhansky, 1937). Esta tasa de mutación (μ) va a ser variable dependiendo de la región del genoma en la que nos encontremos, estando correlacionada con variables genómicas como la recombinación o el contenido GC (Galtier et al., 2001; Lercher & Hurst, 2002). La mayoría de las nuevas mutaciones que aparecen se suponen neutrales o casi-neutrales, mientras que muchas serán deletéreas y pocas ventajosas (Kimura, 1983; Ohta, 1992).

Como la mutación, la recombinación es una fuerza creativa en el sentido de que introduce variación genética nueva en las poblaciones al crear durante la meiosis nuevas combinaciones de alelos existentes en un mismo cromosoma (Creighton & McClintock, 1931; Felsenstein, 1974). Es una fuerza que rompe, por tanto, el ligamiento, es decir la tendencia a que alelos cercanos en el genoma cosegreguen. Debido a su mecanismo de acción, la recombinación tiene, además, un efecto mutagénico en el genoma (Duret & Arndt, 2008). Aunque la tasa de recombinación es variable a lo largo del genoma, a grandes rasgos está relativamente conservada entre especies cercanas (Smukowski & Noor, 2011).

El flujo génico va a producir la entrada de alelos de una población en otra. El impacto del flujo génico va a depender en primer lugar de la cantidad de migración efectiva (*m*), y de las diferencias en frecuencias alélicas entre las dos poblaciones (Slatkin, 1985). En general, se considera una fuerza homogeneizadora entre poblaciones. Cuando las poblaciones que intercambian flujo son de distintas especies se suele hablar de procesos de hibridación o introgresión.

Sobre esta variación preexistente, la selección actúa de forma local en el genoma aumentando la frecuencia del alelo que tenga una ventaja relativa con respecto al otro, cuantificada por su coeficiente de selección (*s*) (Endler, 1986; Fisher 1930; Wright 1932). Cuando la selección actúa aumentando la frecuencia de aquellas mutaciones que resultan beneficiosas se le conoce como selección positiva, mientras que cuando su acción reduce la frecuencia de aquellas que resulten perjudiciales se habla de selección purificadora. La selección balanceadora, en cambio, va a provocar que los alelos se mantengan a frecuencias intermedias, ya sea porque el heterocigoto presenta una ventaja frente a los homocigotos (sobredominancia) o porque las poblaciones se ven sometidas a selección positiva que es cambiante en el tiempo o en el espacio.

Por último, la deriva genética es el muestreo aleatorio de alelos que se dan en una población finita de una generación a la siguiente (Kimura, 1968). Su intensidad es inversamente proporcional al tamaño efectivo poblacional, y por tanto va a ser una fuerza muy potente en poblaciones pequeñas, mientras que su efecto será casi imperceptible en poblaciones muy grandes.

En las poblaciones todos estos procesos coocurren e interactúan. Por ejemplo, el aumento o disminución en frecuencia de alelos que entran en una población por flujo génico va a depender del valor adaptativo de estos alelos en esa población, y este a su vez de las divergencias en ambiente y adaptaciones locales de las poblaciones donantes y receptora. Una interacción muy interesante es la que surge de la coocurrencia de deriva y selección. Aquellas variantes potencialmente deletéreas cuyo coeficiente de selección sea menor que la inversa del tamaño efectivo poblacional van a comportarse como neutras (Kimura, 1962). Por tanto, la eficacia de la selección depende directamente del tamaño efectivo poblacional, i.e. de la fuerza con la que actúe la deriva genética. Esto conlleva que en poblaciones pequeñas variantes potencialmente deletéreas puedan aumentar en frecuencia y acabar eventualmente fijándose, mientras que estas mismas variantes son eliminadas eficientemente por selección en poblaciones grandes. Este proceso se conoce como relajación de la selección purificadora (Charlesworth, Morgan & Charlesworth, 1993).

La demografía va a afectar, pues, a los patrones genómicos, en primer lugar, mediante la deriva genética y la interacción de esta con el resto de fuerzas evolutivas pero, además, la distribución de esas poblaciones en el espacio también puede alterar los patrones genómicos, por ejemplo, facilitando o dificultando la ocurrencia de eventos de flujo génico entre poblaciones.

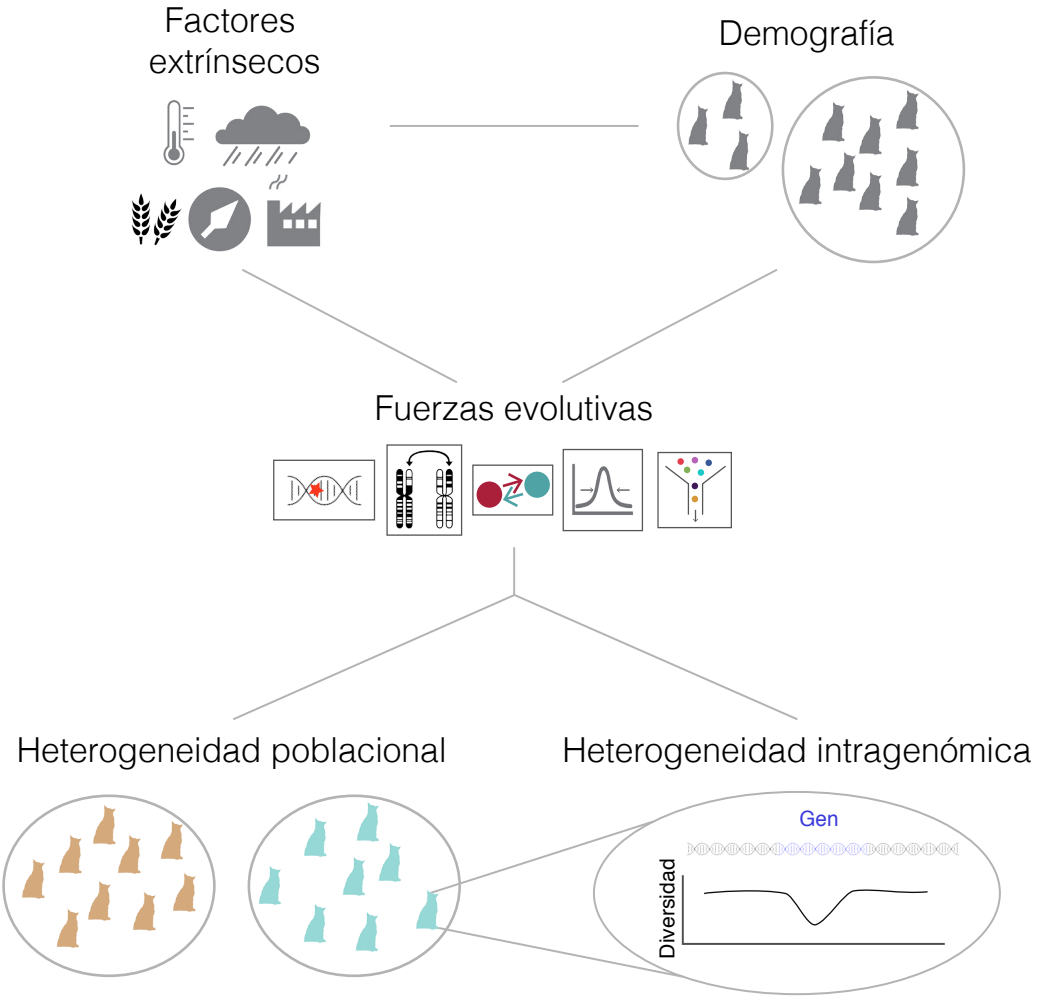


Fig. 1. Representación esquemática de los cómo se generan patrones genómicos a nivel interpoblacional y a lo largo del genoma mediante la interacción de fuerzas evolutivas, cambios demográficos y cambios en el ambiente.

Todos estos procesos que actúan a una escala microevolutiva, en forma de cambio en las frecuencias alélicas de la población, realmente tienen lugar a lo largo de un continuo temporal. Con el tiempo, estos cambios microevolutivos se acaban consolidando como divergencias nucleotídicas que finalmente podrían acabar definiendo linajes y especies.

A lo largo de este continuo, la demografía y las propias fuerzas evolutivas se ven impactadas por sucesos extrínsecos. Eventos naturales, como por ejemplo las oscilaciones climáticas planetarias acaecidas durante el Pleistoceno (Hewitt, 2000; Nadachowska-Brzyska et al., 2015), alteran la demografía de las especies, su área de distribución y los tamaños efectivos de sus poblaciones, y también pueden hacer que aparezcan nuevas presiones selectivas. Además de los eventos naturales, recientemente, la actividad antrópica está produciendo cambios en el entorno que acaban afectando a la biodiversidad. De hecho se estima que la actividad humana está detrás de la extinción de miles de especies animales, y el declive de gran parte de las poblaciones restantes (Dirzo et al., 2014). Además de cambios en la demografía, los humanos están provocando también cambios en las presiones selectivas que impactan en las especies a largo plazo (Otto, 2018; Pelletier & Coltman, 2018).

Por tanto, los genomas que observamos hoy día no son más que la consecuencia de la actuación de todos estos procesos extrínsecos, demográficos y evolutivos y su interacción. Es decir, son un reflejo de su historia evolutiva. Además de una ventana al pasado, el estudio de los genomas nos permite hacer predicciones y, en alguna ocasión, testar la interrelación que existe entre los cambios observables en el ADN y estos eventos, y evaluar el impacto que cambios futuros tendrán sobre la persistencia de estas especies y poblaciones.

INFIRIENDO PROCESOS A PARTIR DE PATRONES

El conocimiento sobre los procesos viene, en última instancia, informado por los patrones que dejan en el genoma. Sin embargo, el estudio de estos patrones y por consiguiente la inferencia del proceso(s) subyacente(s), ha adolecido de una serie de limitaciones debido al abordaje con el que tradicionalmente se han estudiado. Por un lado, el número de marcadores genéticos y, en menor medida, el número de individuos con el que hacer las inferencias suele ser escaso, lo que limita el estudio, básicamente, a patrones de diversidad genética, estructura y diferenciación. Por otro, estos marcadores suelen ser neutrales por lo que quedan fuera del estudio todos los procesos selectivos, que son los que en última

instancia determinan la eficacia biológica y la viabilidad poblacional y los que informan sobre adaptación.

El desarrollo y progresivo abaratamiento de técnicas de secuenciación masiva de nueva generación (en sus siglas en inglés NGS) ha puesto al alcance de especies no modelo los análisis genómicos, incluso a niveles poblacionales (Ellegren, 2014). Disponer de un genoma anotado y poder resecuenciar individuos de distintas poblaciones de la especie supone un salto a nivel cuantitativo, pero más importante aún, a nivel cualitativo con respecto al pasado (Helyar et al., 2011; Ellegren, 2014). Lo que hace unos años se hacía, en el mejor de los casos, con unas decenas de marcadores microsatélites genotipados o secuencias de cientos de pares de bases de ADN mitocondrial en decenas de individuos, puede ahora a hacerse con millones de pares de bases y decenas de miles de variantes en decenas o cientos de individuos. Los patrones de diversidad, estructura y diferenciación se pueden ahora determinar de una manera más precisa y con mucho más poder estadístico. Además, el contar con una alta densidad de marcadores repartidos por todo el genoma permite capturar información que anteriormente quedaba fuera de alcance. En concreto, una alta densidad de marcadores permite hacer inferencias demográficas pasadas y estimas precisas del tamaño poblacional presente de la especie, ya sea mediante la distribución de los tiempos de coalescencia a lo largo del genoma (Li & Durbin, 2011), el espectro de frecuencias (Liu & Fu, 2015) o su grado de ligamiento (Hill, 1981; Barbato et al., 2015). Por último, los estudios genómicos nos permiten acceder al componente de la variación no neutral, eliminando la principal barrera para el estudio de la selección natural en poblaciones silvestres (Vitti, Grossman & Sabeti, 2013).

Toda esta información, ahora accesible, permite inferir mucho mejor los procesos subyacentes a estos patrones, pero incluso cuando esta información cuenta con el mayor grado de detalle posible, la información inferida a partir de genomas de poblaciones contemporáneas nos da una imagen parcial de la historia evolutiva de la especie. Por ejemplo, a pesar de poder inferir ciertos procesos, usando sólo poblaciones contemporáneas no podemos describir la dinámica de los patrones genómicos en el tiempo y en el espacio. De forma más evidente, los patrones de linajes extintos y sus relaciones con los linajes contemporáneos quedarán fuera del estudio. Bajo esta perspectiva, un muestreo que abarque distintas épocas y espacios emerge como una aproximación mucho más poderosa para inferir los procesos que han ocurrido a lo largo de la historia evolutiva de la especie. Sin embargo, el reto que supone en muchos casos acceder a muestras alejadas en el tiempo, extraer su material genético y secuenciarlo ha limitado, hasta ahora, el acceso a este tipo de información.

Por suerte, de un tiempo a esta parte, junto con el desarrollo de técnicas de secuenciación masiva, se han desarrollado nuevas y mejores técnicas moleculares de extracción de ADN y preparación de librerías para secuenciación a partir de material degradado, en muchos casos subfósil, aliviando estas limitaciones y permitiendo acceder a información de poblaciones antiguas, no solo a nivel de zonas neutrales o mitocondriales, sino a nivel de genoma completo (Dabney et al., 2013; Gansauge & Meyer, 2013).

El rápido avance de la genómica a nivel de laboratorio también ha propiciado un progreso completamente abrumador de las herramientas disponibles para analizar estos nuevos datos. Estos adelantos conllevan a su vez una serie de retos, pero también de oportunidades. La mayoría de estas herramientas requieren por parte del usuario un conocimiento sobre computación y programación que incluso a día de hoy queda en muchos casos fuera de su formación como biólogo (Kumar & Dudley, 2007). Además, muchos programas se desarrollan en el marco de grupos de investigación pequeños, generalmente formados por científicos sin mucha experiencia en el mundo de la computación, sin financiación dedicada para ello y de una forma *ad hoc*, por lo que suelen adolecer de problemas de estabilidad o fiabilidad (Mangul et al., 2019; Siepel, 2019). Sin embargo, aunque sin duda estas limitaciones están lastrando el avance del campo, la creciente formación computacional de los biólogos que se enfrentan a estos retos, junto con las propias características del análisis de estos datos, generalmente secuencias de comandos escritos, abre también una oportunidad única a que la ciencia se vuelva más reproducible y más democrática (Groves & Godlee, 2012).

INFERIR EL PASADO PARA AFRONTAR EL FUTURO

Los avances anteriormente descritos han propiciado que mejore el conocimiento sobre muchas especies de interés. Dentro de estas, y debido al escenario de cambio global en el que nos encontramos, las especies en peligro de extinción son un grupo especialmente relevante, puesto que por un lado nos aportan un escenario evolutivo único y por otro nos brindan, en muchos casos, la oportunidad de informar y planear las acciones de conservación basándonos en conocimiento científico (Ryder, 2005; Williams, Nevill & Krauss, 2014).

El lince como modelo

El género lince está formado por un linaje independiente de felinos que surgió entre 8 y 10 millones de años al separarse del linaje de gatos asiáticos (Catupuma y Pardofelis) (Li et al., 2016). En la actualidad el género cuenta con cuatro especies: el lince rojo (*Lynx rufus*) y el lince canadiense (*Lynx canadensis*) en Norteamérica, el lince boreal (*Lynx lynx*) en Eurasia y el lince ibérico (*Lynx pardinus*) en la península ibérica (Sunquist & Sunquist, 2002). Aunque tradicionalmente se ha asignado la categoría de especies hermanas al lince boreal y al lince ibérico, existen ciertas incongruencias entre filogenias mitocondriales, que soportan al lince ibérico como especie hermana de lince canadiense, y genomas nucleares, que mayoritariamente sí agrupan al lince boreal con el ibérico (Li et al., 2016). Sin embargo, cuando estas inferencias nucleares se realizan basándose en regiones sin recombinación el lince ibérico aparece como una especie hermana al grupo canadiense-boreal (Li et al., 2019). Estas incongruencias sugieren la ocurrencia en el género lince, y en especial en el lince ibérico, de eventos frecuentes de hibridación, que ya se habían inferido a partir de unas pocas secuencias de genoma completo (Abascal et al., 2016) y que caracterizan, por otra parte, a la mayor parte de especies de felinos (Li et al., 2019).

Las especies del género lince presentan ciertos rasgos morfológicos y de ciclo de vida comunes. Son felinos de tamaño medio, con garras poderosas y colas cortas, con un pelaje moteado que va desde un gris amarillento hasta tonos más rojizos. Los linces suelen vivir entre 10 y 15 años en libertad y alcanzar la madurez sexual en torno a los 2-3 años (Fritts & Sealander, 1978; Axnér et al., 2009; Göritz et al., 2009; Nilsen et al., 2010), y con un tiempo de generación algo mayor en torno a 4-5 años (Lucena-Perez et al., 2018). Aunque generalmente habitan zonas boscosas o de matorral, la diversidad de hábitats, incluso en poblaciones de la misma especie, es grande (Breitenmoser et al., 2015; Rodríguez & Calzada, 2015; Kelly, Morin & López-González, 2016; Vashon, 2016). Con mayor o menor grado de especialización dependiendo de la especie de lince considerada, sus presas suelen ser lagomorfos, aunque la especie de mayor tamaño, el lince boreal, suele depredar sobre animales de mayor porte como ungulados (Breitenmoser et al., 2015; Rodríguez & Calzada, 2015; Kelly, Morin & López-González, 2016; Vashon, 2016). Sin embargo, a pesar de estas similitudes, las cuatro especies de lince son muy diferentes en cuanto a distribución y estado de conservación. Esto es así especialmente en las dos especies de lince que van a ser objeto de estudio de esta tesis doctoral, el lince boreal y el lince ibérico.

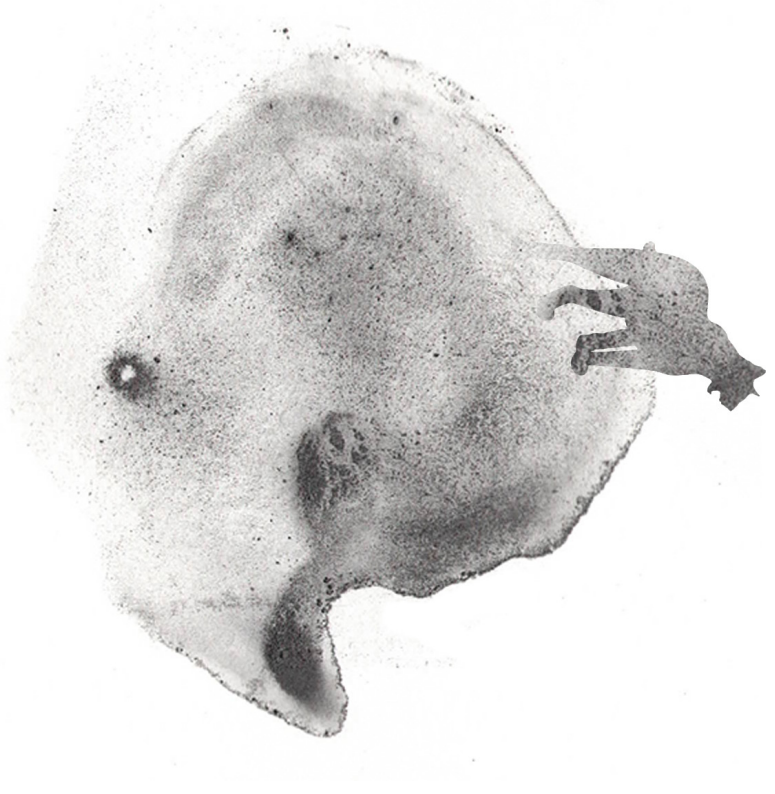
Fig. 2. Fotografías de las cuatro especies del género *Lynx*. De izquierda a derecha y de arriba a abajo, lince ibérico (*Lynx pardinus*), lince boreal (*Lynx lynx*), lince canadiense (*Lynx canadensis*) y lince rojo (*Lynx rufus*). Las fotos están realizadas, en orden, por: Programa de Conservación Ex-situ del lince ibérico, Martin Mecnarowski, Keith Williams y Bill W Ca.

El lince boreal se distribuye por gran parte de Eurasia abarcando, de forma discontinua, desde el Lejano Oriente ruso hasta la península escandinava, y desde la frontera del Polo Norte hasta el Mar Mediterráneo. La extensión de su área de distribución, que engloba una gran diversidad de zonas climáticas y hábitats, junto con algunas evidencias que indicaban tamaños poblacionales grandes, ha llevado a considerar al lince boreal como una especie de preocupación menor en cuanto a conservación (Breitenmoser et al., 2015). A pesar de ser la especie de lince, y probablemente uno de los felinos, con mayor área de distribución, esta se ha visto mermada, especialmente durante los últimos siglos. Según registros históricos y evidencias fósiles el lince boreal, que se originó aparentemente en Asia, se extendió hasta el oeste de Europa, incluyendo la península ibérica, durante el Pleistoceno tardío y el comienzo del Holoceno (Kratochvil, 1968; Von Arx et al., 2004; Breitenmoser et al., 2015; Boscaini

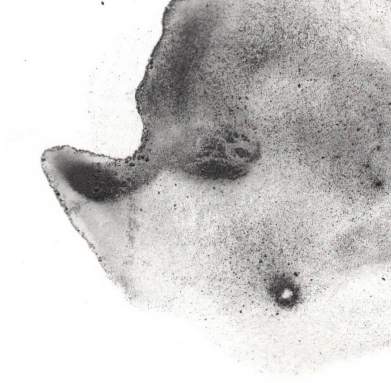
et al., 2016; Rodríguez-Varela et al., 2016). En Europa occidental las poblaciones sobrevivieron hasta hace unos siglos cuando empezaron a desaparecer primero de las Islas Británicas durante la Edad Media (Hetherington, Lord & Jacobi, 2006), y después del resto de la zona oeste de Europa continental durante buena parte del siglo XIX y XX (Kratochvil, 1968; Clavero & Delibes, 2013). A día de hoy, son precisamente las poblaciones remanentes más occidentales las que presentan un estado genético más erosionado, tras su aislamiento del resto de la distribución y la actuación de la deriva genética (Hellborg et al., 2002; Schmidt et al., 2009; Ratkiewicz et al., 2012; Förster et al., 2018). Para las poblaciones que ocupan la zona este de Europa, el Cáucaso y Asia la información disponible es relativamente escasa pero en general se consideran poblaciones abundantes y demográficamente sanas (Breitenmoser et al., 2015).

Por otro lado, el lince ibérico es la especie del género que se encuentra más amenazada (Rodríguez & Calzada, 2015). La especie que se encuentra hoy restringida a la mitad sur de la península, aparece por primera vez en el NE de la península iberica en el Pleistoceno Temprano (Boscaini et al., 2016), ocupando a partir de ese momento y durante buena parte del Holoceno, zonas del norte de la península ibérica, sur de Francia e incluso el noroeste de Italia, donde convivió con el lince boreal (Rodríguez-Varela et al., 2015; Boscaini et al., 2016). Durante su historia evolutiva el lince ibérico ha sufrido una serie de cuellos de botella, que empezaron hace unos 50000 años (Abascal et al., 2016). El último gran declive de la especie se inició hace unos tres o cuatro siglos y ha sido reconstruido gracias a reportes de campo (Rodríguez & Delibes, 2002), a datos genéticos de material de museo y subfósil (Casas-Marce et al., 2017) e inferencias basadas en datos de genomas completos (Abascal et al., 2016). Durante este último declive se sabe que las poblaciones de lince transitaron desde una metapoblación bien conectada y de un tamaño poblacional más o menos estable hacia un progresivo aislamiento, primero de las poblaciones en los límites de la distribución, y una extinción generalizada, hasta quedar exclusivamente dos poblaciones, Andújar de mayor tamaño en el centro de la distribución, y Doñana, marginal y de muy pequeño tamaño. En 2002, la especie contaba con aproximadamente 100 individuos y se llegó a considerar el felino más amenazado del mundo (Nowell & Jackson, 1996). Esta sucesión de cuellos de botella y su bajo tamaño poblacional han tenido un gran impacto en cuanto a su diversidad genética y estructura, como se ha comprobado mediante estudios de microsatélites y genes mitocondriales primero, y genomas completos recientemente (Palomares et al., 2012; Casas-Marce et al., 2013, 2017; Abascal et al., 2016). A día de hoy, el lince ibérico es una de las especies con menor diversidad genómica global (Abascal et al., 2016).

El tándem que forman lince boreal y lince ibérico constituye un buen modelo para estudiar la dinámica de especies con rasgos de vida similares, pero historias demográficas en apariencia muy dispares. Además, gracias al trabajo que se ha llevado a cabo hasta la fecha en ambas especies, pero sobre todo el lince ibérico, disponemos de gran número de recursos que son cruciales para responder los interrogantes que nos planteamos en esta tesis. Por un lado, contamos con un genoma de referencia de lince ibérico ensamblado y anotado (Abascal et al., 2016) y un genoma de gato con una contigüidad que llega a nivel de cromosoma (Buckley et al., 2020), y se conoce la sintenia entre ambos genomas. Por otro lado, los programas de seguimiento y gestión a los que se somete a las dos especies, aunque desiguales, brindan una oportunidad única para la consecución de muestras contemporáneas que se completan con un buen conjunto de muestras antiguas en el caso de lince ibérico que se prolongan atrás en el tiempo hasta los miles de años. Por último, la posibilidad de informar acciones de conservación, en una especie muy manejada, y en otra en la que la información es escasa y está poco estandarizada, supone una oportunidad única de contribuir a una mejora en la conservación de estas dos especies.



OBJETIVOS



Con este marco teórico en mente esta tesis pretende describir los patrones genómicos de poblaciones presentes y pasadas con el objetivo de reconstruir la historia evolutiva de lince boreal (*Lynx lynx*) y lince ibérico (*Lynx pardinus*), y determinar el rol que distintas fuerzas evolutivas, cambios demográficos y factores extrínsecos han jugado en el moldeado del genoma de las dos especies a lo largo del tiempo.

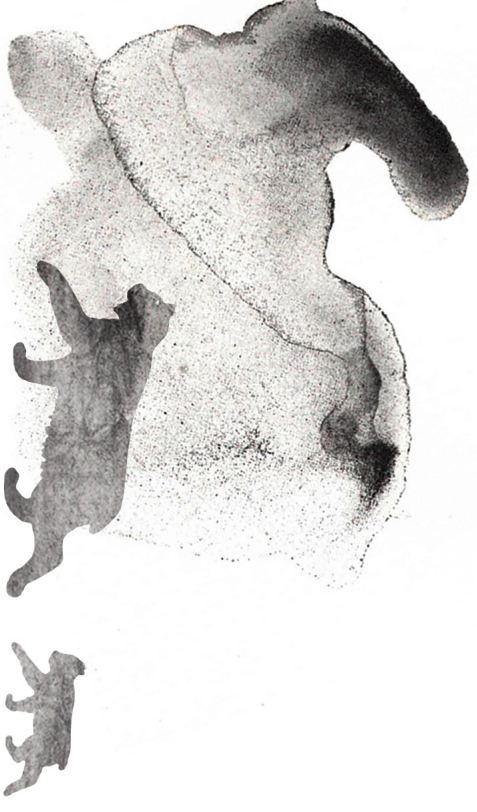
Concretamente, durante la tesis transitamos de una escala más macroevolutiva, estudiando distintos linajes y poblaciones presentes o pasadas que los conforman, hasta una escala más microevolutiva, en la que describimos los cambios a lo largo del genoma de diversas poblaciones ante procesos relativamente recientes y de origen antrópico. Los objetivos específicos son los siguientes:

1. Describir los patrones genómicos de poblaciones de lince boreal a lo largo de prácticamente toda su área de distribución e inferir los procesos subyacentes a los mismos, tanto demográficos como evolutivos, relacionándolos con procesos extrínsecos con los que convivió la especie a lo largo de su historia evolutiva.

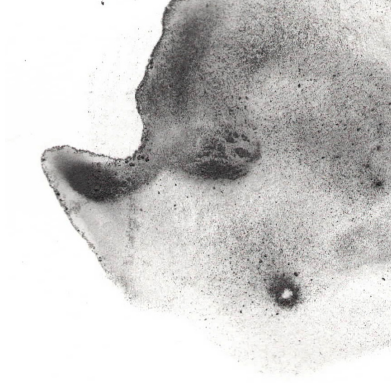
2. Describir el linaje de lince boreal que habitó la península ibérica, usando para ello el genoma completo de una muestra datada en torno a 2000 años de antigüedad, y relacionarlo con las poblaciones presentes, y con la información pasada disponible, y arrojar luz sobre la dinámica de la especie a lo largo del tiempo en Europa occidental.

3. Describir cuál ha sido la dinámica de la diversidad y la estructura en el lince ibérico durante los últimos miles de años, usando para ello genomas completos de muestras milenarias y contemporáneas, e investigar el rol que ciertos procesos evolutivos, en concreto la hibridación, han podido tener en el modelado de los patrones a lo largo del tiempo.

4. Caracterizar los cambios que provoca un declive demográfico muy reciente en los niveles de diversidad genética a lo largo del genoma, comparando muestras pertenecientes a poblaciones sometidas a cuellos de botella recientes frente a poblaciones más estables de lince boreal e ibérico, y determinar el rol que procesos como la mutación, la recombinación o la selección han podido jugar en el moldeado de estos patrones intragenómicos.



MATERIALES Y MÉTODOS



POBLACIONES DE ESTUDIO Y DISEÑO EXPERIMENTAL

Durante el desarrollo de esta tesis se han usado, dependiendo del objetivo, diferentes poblaciones de estudio que en algunos casos se repiten a lo largo de los capítulos (Fig. 1). En esta sección paso a resumir qué poblaciones incluimos en cada capítulo y cómo se ha diseñado el muestreo dependiendo del objetivo a abordar.

En el primer capítulo de esta tesis se analizaron 96 muestras de lince boreal, 80 de ellas a nivel de genoma completo, tanto nuclear como mitocondrial, y las 16 restantes a nivel de mitogenoma. Estas muestras se encuentran distribuidas a lo largo del área de distribución de la especie y, de acuerdo a su procedencia, las agrupamos en las siguientes poblaciones: 1) NE-Polonia (Białowieża and Knyszyn Primeval Forests); 2) Balcanes; 3) Cárpatos; 4) Letonia; 5) Noruega; 6) Kirov, Rusia; 7) Urales, Rusia; 8) Tuva (República de Tyva), Rusia; 9) Yakutia (República de Sakha), Rusia; 10) Primorsky Krai, Rusia; 11) Mongolia. Las localidades se escogieron basándonos principalmente en dos parámetros, disponibilidad de muestras y cobertura espacial, intentando tener una buena representación, en términos de número de individuos, de cada una de ellas.

Para el segundo capítulo secuenciamos a genoma completo una muestra antigua de lince boreal residente en la península ibérica, concretamente en la zona del País Vasco, de 2000 años de antigüedad. Esta muestra se analizó junto con las muestras mencionadas anteriormente, nuevas muestras disponibles de los Balcanes y el Cáucaso que se van a describir por primera vez en un estudio actualmente en preparación (Lucena-Pérez et al. *in prep.*), y secuencias de mitocondrial de muestras antiguas distribuidas por Europa generadas por estudios anteriores (Rodríguez-Varela et al., 2015, 2016).

En el tercer capítulo el foco se centra en el lince ibérico y, con una aproximación similar conceptualmente a la anterior, se analizan tres genomas completos de individuos que vivieron en distintos puntos de la península entre 2000 y 4300 años atrás, junto con las

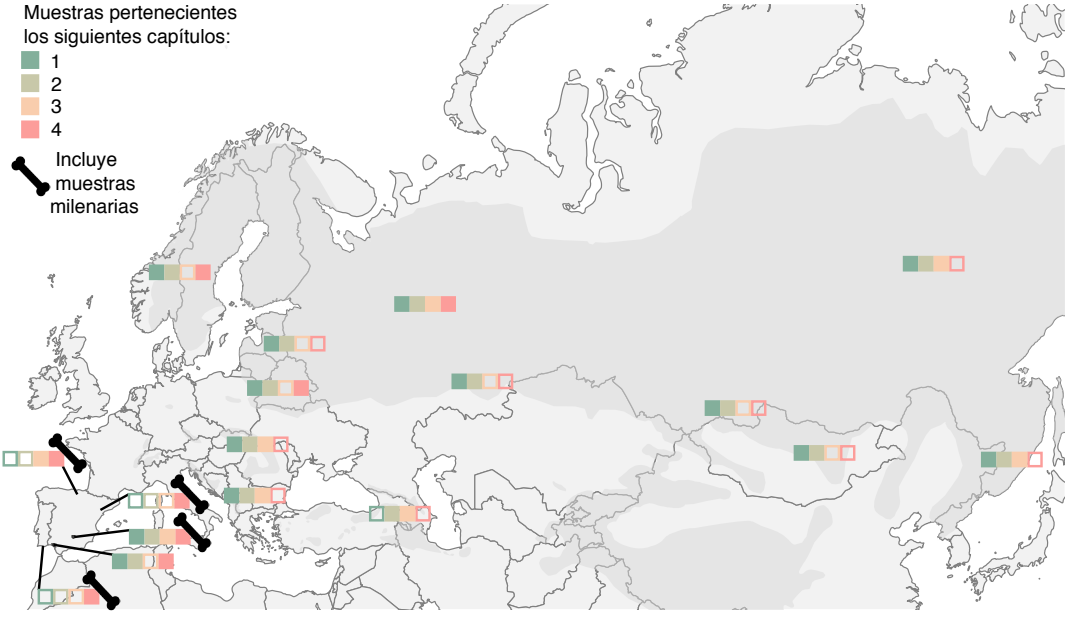


Fig. 1. Poblaciones muestreadas en esta tesis para genoma completo y capítulos en los que se han empleado representados por distintos colores. Los cuadrados sombreados por completo indican inclusión en ese capítulo mientras que un cuadrado vacío indica que no ha sido empleada esa población en ese capítulo. Aunque la mayoría de las muestras de la población de Balcanes y la totalidad de las muestras de Caucaso no se han secuenciado específicamente para esta tesis, se han incluido estas poblaciones puesto que se usan datos de geomas completos de estas poblaciones en varios de los capítulos.

dos poblaciones contemporáneas que persistieron hasta el presente. Concretamente, se usaron muestras de genoma completo secuenciadas anteriormente por el grupo (Abascal et al., 2016), junto con nuevos datos de genoma completo, que en total suman 12 individuos para la población contemporánea de Doñana, y 18 para Sierra Morena. Para uno de los análisis llevados a cabo en este capítulo se usaron también muestras de seis poblaciones de lince boreal.

Para el cuarto capítulo se eligieron, en base a los resultados del primer capítulo y del conocimiento previo, muestras de poblaciones pertenecientes a una misma especie, lince boreal o lince ibérico, pero que tuviesen historias demográficas muy recientes diferentes. Concretamente seleccionamos para su comparación una población demográficamente estable, dentro de lo posible, que se pudiese asemejar a la población ancestral previa al cuello de botella sufrido en los últimos siglos, y otra(s) población(es) derivada(s) que hubiese(n) sufrido un declive poblacional marcado y un aislamiento reciente. Las poblaciones de lince boreal elegidas fueron Kirov, como población estable, y NE-Polonia y Noruega como poblaciones derivadas, todas ellas

pertenecientes al linaje europeo de lince boreal, y por otro lado, Andújar, como población relativamente estable, y Doñana como población derivada.

Todas las muestras han sido aportadas por colaboradores del proyecto, bajo los pertinentes permisos especificados en cada capítulo, o estaban disponibles gracias a trabajos previos del grupo.

La naturaleza de las muestras ha sido variable dependiendo principalmente de la oportunidad. Así, tenemos muestras muy óptimas para la extracción de ADN como músculo fresco y sangre bien preservada, mientras que en otras ocasiones la calidad de las muestras era muy inferior. En concreto, durante el desarrollo de esta tesis se ha trabajado con piel y tejidos blandos de individuos fallecidos, con un grado de conservación variable, encontrados de forma oportunista durante programas de seguimiento o estudios de campo. Por último, la aproximación temporal ha obligado a trabajar con muestras muy degradadas, huesos subfósiles de miles de años, y material de museo, principalmente piel. Con estas últimas muestras no se consiguió generar datos genómicos de calidad suficiente y, por tanto, no llegaron a formar parte de esta tesis.

EXTRACCIÓN DE ADN, SECUENCIACIÓN Y MAPEO

Cada una de las muestras usadas en esta tesis recibió un tratamiento adecuado a su grado de conservación en las instalaciones apropiadas. A continuación se pasa a exponer muy brevemente en que consistieron estos distintos protocolos usados, de los que se pueden encontrar más detalles en cada uno de los capítulos.

En el caso de muestras contemporáneas, las técnicas de extracción de ADN fueron principalmente dos tipos: extracción con fenol-cloroformo y extracción con cuentas paramagnéticas recubiertas de sílica (NucleoMag® Tissue, MACHEREY-NAGEL GmbH & Co. KG), y se llevaron a cabo en las instalaciones del Laboratorio de Ecología Molecular (LEM) de la Estación Biología de Doñana. La preparación de librerías de muestras contemporáneas se llevó a cabo en las instalaciones del CNAG (España) y Macrogen (Corea), donde posteriormente se secuenciaron, por técnicos especialistas siguiendo los protocolos establecidos por Illumina en cada momento para cada plataforma. Brevemente, el ADN extraído se fragmentó, y tras una selección de tamaño se le repararon los extremos para permitir la ligación de adaptadores que incorporan las secuencias necesarias para la amplificación en

puente, la secuenciación y, en el caso que varias muestras se secuenciaran juntas en un único carril, identificadores únicos de muestra. Tras la preparación de las librerías las muestras se secuenciaron en maquinaria Illumina de distinto tipo dependiendo del tipo de librería. Básicamente las plataformas de secuenciación usadas han sido HiSeq2000/4000 y HiSeq X Ten. En todos los casos, el análisis de datos inicial se llevo a cabo siguiendo el protocolo de Illumina. La mayoría de muestras contemporáneas se secuenciaron a una profundidad de cobertura de 5 a 10x con HiSeq2000/4000, mientras que algunas muestras concretas representantes de poblaciones de interés se secuenciaron a 20x con HiSeq X Ten.

La extracción y procesado de ADN de muestras muy degradadas, o subfósiles, tuvo lugar en el laboratorio especializado para tratamiento de este tipo de muestras de la Universidad de Potsdam (Alemania) que está a cargo del Dr. Michael Hofreiter. Para ello se seleccionaron 20 muestras que habían dado positivo en amplificación nuclear en un test preliminar para un estudio anterior (Casas-Marce et al., 2017). Estas muestras se extrajeron siguiendo un protocolo con guanidinio (Rohland & Hofreiter, 2007) en el que se incluyeron las modificaciones necesarias para la recuperación de fragmentos de ADN de pequeño tamaño que se describe en Dabney et al. (2013). En algunos casos este paso de extracción fue precedido por un tratamiento del polvo de hueso con lejía para eliminar ADN exógeno de la superficie (Korlević et al., 2015). Posteriormente, se prepararon librerías de cadena simple (Gansauge & Meyer, 2013) que permiten recuperar todas las cadenas de ADN que se han desnaturalizado por el paso del tiempo. Durante la preparación de librerías se añadió un tratamiento de las muestras con uracil-DNA glicosilada y endonucleasa VIII para eliminar el daño post-mortem. Una vez preparadas las librerías se procedió a una primera evaluación del contenido endógeno de las muestras con secuenciación a pequeña escala en un equipo Illumina MiSeq.

De las 70 librerías procesadas, 4 de ellas, tres de lince ibérico y una de lince boreal, tuvieron un contenido endógeno suficientemente elevado y pudieron ser secuenciadas a genoma completo con profundidad de cobertura entre 2.5 y 4x en el Museo de Historia Natural Sueco en un equipo Illumina HiSeq X.

Tras la secuenciación de todas las muestras, tanto contemporáneas como antiguas, se procedió a la evaluación de la calidad usando FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Posteriormente, se mapearon usando BWA-mem (Li, 2013) en el caso de análisis que únicamente se incluyeran muestras contemporáneas o BWA-aln (Li & Durbin, 2009) si se analizaban muestras contemporáneas junto

con muestras antiguas. Los genomas de referencia nucleares usados en esta tesis son el genoma de lince ibérico (Abascal et al., 2016), y el genoma del gato doméstico (*Felis catus*) (GCF_000181335.3) (Buckley et al., 2020). Para los análisis de genoma mitocondrial, las secuencias se mapearon al mitogenoma de lince boreal generado en Abascal et al. (2016). Las lecturas mapeadas se filtraron por calidad usando samtools (Li et al., 2009), se clasificaron por grupos de lectura (*read groups*) y se marcaron como duplicados cuando correspondía usando Picard (https://broadinstitute.github.io/picard), y se realinearon usando GATK (McKenna et al., 2010).

ANÁLISIS DE LOS DATOS

Dependiendo del objetivo concreto a abordar, el análisis de los datos varió, tanto por el programa usado como por las regiones del genoma analizadas, como se especifica en cada capítulo. No obstante, hay un programa, *ANGSD*, que se repite consistentemente a lo largo de todos los capítulos y que tiene una filosofía analítica concreta algo diferente a la aproximación tradicional que usa polimorfismos de nucleótidos simples (*SNPs*).

ANGSD es una suite de aplicaciones que trabaja directamente con datos de mapeo y que integra la incertidumbre en el llamado de genotipos. En lugar de asumir como cierto el genotipo de mayor verosimilitud, *ANGSD* trabaja directamente con la verosimilitud de todos los genotipos posibles (Korneliussen, Albrechtsen & Nielsen, 2014). Esta aproximación tiene la ventaja de ser más precisa al dar medidas de diversidad cuando se trabaja con datos de cobertura media-baja, que son los que presentan mayor incertidumbre en el llamado de genotipos debido a errores de secuenciación y mapeo, pero especialmente debido al muestreo aleatorio de lecturas cuando el genoma es diploide, lo que puede provocar que se pierda con frecuencia algunos de los alelos (Korneliussen, Albrechtsen & Nielsen, 2014). Por eso, debido a que en esta tesis los datos analizados son en su mayoría de cobertura media-baja (5x a 10x) se ha optado por esta aproximación metodológica. El contar con una metodología que se había probado adecuada para el análisis de datos de coberturas medias-bajas nos permitió, además, aumentar el número de individuos secuenciados por población como recomiendan trabajos recientes (Fumagalli, 2013). Usando *ANGSD* y las diversas suites y programas asociados se ha calculado diversidad genética (π y θ Watterson), se ha medido estructura (usando PCA y análisis de ancestría individual) y diferenciación (poblacional usando FST e interindividual usando distancias genéticas).

Además, en esta tesis se han usado programas que nos han permitido inferir la historia demográfica de las distintas poblaciones, como *PSMC* (Li & Durbin, 2011), *Stairway plot* (Beichman, Huerta-Sanchez & Lohmueller, 2018) o *SNeP* (Barbato et al., 2015) dependiendo del periodo a reconstruir. Por otro lado, las relaciones evolutivas entre poblaciones se han inferido usando *Treemix* (Pickrell & Pritchard, 2012). Por último, se ha usado el estadístico D (D-stat) basado en la cantidad de sitios ABBA vs. BABA para inferir procesos de mezcla entre poblaciones (Green et al., 2010; Durand et al., 2011). Además, se han calculado medidas de diversidad y reconstruido y datado las relaciones filogenéticas en genomas mitocondriales (Bouckaert et al., 2014; Stamatakis, 2014; Leigh & Bryant, 2015).

Todos estos programas se han ejecutado, usando secuencias de comandos personalizados escritos en su mayoría en *bash*, en la terminal del servidor de genómica de la EBD (ICTS-RBD) o usando la infrastructura del Centro Tecnolóxico de Supercomputación de Galicia (CESGA). Las representaciones gráficas y la mayoría de análisis estadísticos o matemáticos se han realizado usando *R* (R Core Team, 2019) y paquetes asociados al mismo.

Por último, como hemos mencionado anteriormente, dependiendo del objetivo se han analizado distintos compartimentos genómicos. Así para los análisis de filogeografía en el capítulo uno y dos se optó por analizar solo las zonas intergénicas y mitogenoma, distinguiendo los autosomas y el cromosoma X, mientras que en otros capítulos se analiza todo el genoma.

10 CAPÍTULO

GENOMIC PATTERNS IN THE WIDESPREAD EURASIAN LYNX SHAPED BY LATE QUATERNARY CLIMATIC FLUCTUATIONS AND ANTHROPOGENIC IMPACTS

Maria Lucena-Perez

Elena Marmesat, Daniel Kleinman-Ruiz, Begoña Martínez-Cruz, Karolina Węcek, Alexander P. Saveljev, Ivan V. Seryodkin, Innokentiy Okhlopkov, Mikhail G. Dvornikov, Janis Ozolins, Galsandorj Naranbaatar, Milan Paunovic, Mirosław Ratkiewicz, Krzysztof Schmidt, José A. Godoy

ABSTRACT

Disentangling the contribution of long-term evolutionary processes and recent anthropogenic impacts to current genetic patterns of wildlife species is key for assessing genetic risks and designing conservation strategies. Here, we used 80 whole nuclear genomes and 96 mitogenomes from populations of the Eurasian lynx covering a range of conservation statuses, climatic zones and subspecies across Eurasia to infer the demographic history, reconstruct genetic patterns and discuss the influence of long-term isolation and/or more recent human-driven changes. Our results show that Eurasian lynx populations shared a common history until 100 kya, when Asian and European populations started to diverge and both entered a period of continuous and widespread decline, with western populations, except Kirov, maintaining lower effective sizes than eastern populations. Population declines and increased isolation in more recent times likely drove the genetic differentiation between geographically and ecologically close westernmost European populations. By contrast, and despite the wide range of habitats covered, populations are quite homogeneous genetically across the Asian range, showing a pattern of isolation by distance and providing little genetic support for the several proposed subspecies. Mitogenomic and nuclear divergences and population declines starting during the Late Pleistocene can be mostly attributed to climatic fluctuations and early human influence, but the widespread and sustained decline since the Holocene is more probably the consequence of anthropogenic impacts which intensified during the last centuries, especially in western Europe. Genetic erosion in isolated European populations and lack of evidence for long-term isolation argue for the restoration of lost population connectivity.

INTRODUCTION

Climatic oscillations and geological events have influenced the range of species and the size and connectivity of their populations, driving divergence and admixture processes that give rise to the biodiversity patterns we see today (Endler, 1977; Avise et al., 1987). More recently, human-driven habitat alteration, fragmentation, and destruction, among other drivers of biodiversity loss, are fuelling the decline and subdivision of populations into small and isolated fragments where random genetic drift becomes the main evolutionary force. The result is often the loss of genetic variation, an increase in inbreeding in the population, and the genetic differentiation among populations (Benazzo et al., 2017; Srbek-Araujo et al., 2018; Thatte et al., 2018). Recent, human-driven genetic divergence among populations must be considered together with the effects of long-term evolution in isolation, which enable adaptive divergence and, eventually, speciation, as possible factors shaping current genetic patterns (Frankham et al., 2002; Allendorf, Luikart & Aitken, 2013). It is thus important that the delimitation of conservation units and the design of conservation strategies are informed by good knowledge of the demographic and evolutionary processes that have acted upon the species across space and time.

Recently developed high throughput sequencing approaches can significantly expand our ability to obtain genomic scale information and infer evolutionary processes in non-model species in a cost-effective way. Also, the availability of new reference genomes is helping to overcome most of the limitations of classical genetic markers and to expand the range of questions that can be addressed, including the assessment of the relative influence of current (human-driven) and long-term evolutionary processes (Murchison et al., 2012; Li et al., 2014; Abascal et al., 2016; Feng et al., 2019).

The Eurasian lynx (*Lynx lynx*) is one of the most broadly distributed felids in the world, representing a suitable but understudied model for exploring the long-term, as well as recent anthropogenic impacts, patterns of variation in the genome. The species' range extends from Central Europe to the Asian Far East, encompasses a wide range of habitats (shrubland, forest, desert, rocky areas, and grassland) and climates (Mediterranean, temperate, boreal; from sea level to 5,500 m), and includes populations with varied recent demography, some of which were led to near extirpation in the last century by anthropogenic impacts and extermination policies, followed by varied rates of recovery. The fossil and historical records indicate that

the Eurasian lynx was already present in Europe during the Pleistocene (Sommer & Benecke, 2006), and that its westernmost range reached the Iberian Peninsula (Clavero & Delibes, 2013; Rodríguez-Varela et al., 2016) and Great Britain (Hetherington, Lord & Jacobi, 2006). The species was extirpated from most of central, western and southern Europe during the 20th century, and the remaining central European populations are severely fragmented and isolated. Previous genetic studies of these European populations using microsatellite markers and mtDNA sequences have found the lowest levels of diversity and strong population differentiation within Europe, but also relatively high levels of gene flow among populations across the central part of its range (Hellborg et al., 2002; Schmidt et al., 2009; Ratkiewicz et al., 2012; Förster et al., 2018). In contrast, the range of the species in Asia is often described as continuous, with demographically healthy and well-connected populations (Rueness et al., 2014), although the information is often scattered or completely lacking. The only genetic study that covered most of the distribution range of the species, by resorting to museum specimens, found three different mitochondrial clades and a clear structuring along an east-west gradient (Rueness et al., 2014). However, the low resolution imposed by the few microsatellites and the small mitochondrial region used hampered robust conclusions on the phylogeographic relationships among populations and on the ultimate drivers of the observed genetic differentiation. While the Eurasian lynx has historically been divided into many subspecies based mostly on morphological characteristics (Kitchener et al. 2018), so far genetic studies have not been able to provide sufficient resolution data to resolve intraspecific taxonomy.

With the power of genomics and the recent availability of a reference genome from the closely related Iberian lynx (*Lynx pardinus*; Abascal et al., 2016), we analysed the genetic variation of the Eurasian lynx across most of its geographic range to assess the relative influence of evolutionary history and recent demographic declines and fragmentation. Specifically, we addressed to what extent long-term isolation and/or recent human-driven changes have impacted the lynx populations by analysing: i) the history of population size, divergence and admixture among lynx populations; ii) the current patterns of genetic structure and diversity across its distributional range. Additionally, we discuss the level of genetic support for the proposed subspecies and the implications for the conservation of its most endangered populations.

MATERIALS AND METHODS

Sampling

We sampled 80 *L. lynx* across the distribution range of the species, including five out of the six subspecies proposed by the IUCN Cat Specialist Group: *L. l. lynx*, *L. l. balcanicus*, *L. l. carpathicus*, *L. l. isabellinus*, and *L. l. wrangeli* (Kitchener et al., 2018) (Fig. 1; Table S1). Also, one *Lynx rufus* (bobcat) from Jerez Zoo (Spain), and one *Lynx canadensis* from Ostrava Zoo (Czech Republic) were sampled to be used as an outgroup and to identify the ancestral state of detected variants (supplementary methods).

The design of the study was intended to sample a minimum of six individuals from eleven a priori populations defined on the basis of only geography: 1) North-Eastern Poland (Białowieża and Knyszyn Primeval Forests); 2) Balkans; 3) Carpathian Mountains; 4) Latvia; 5) Norway; 6) Kirov region, Russia; 7) Ural Mountains, Russia; 8) Tuva (the Republic of Tyva), Russia; 9) Yakutia (Republic of Sakha), Russia; 10)

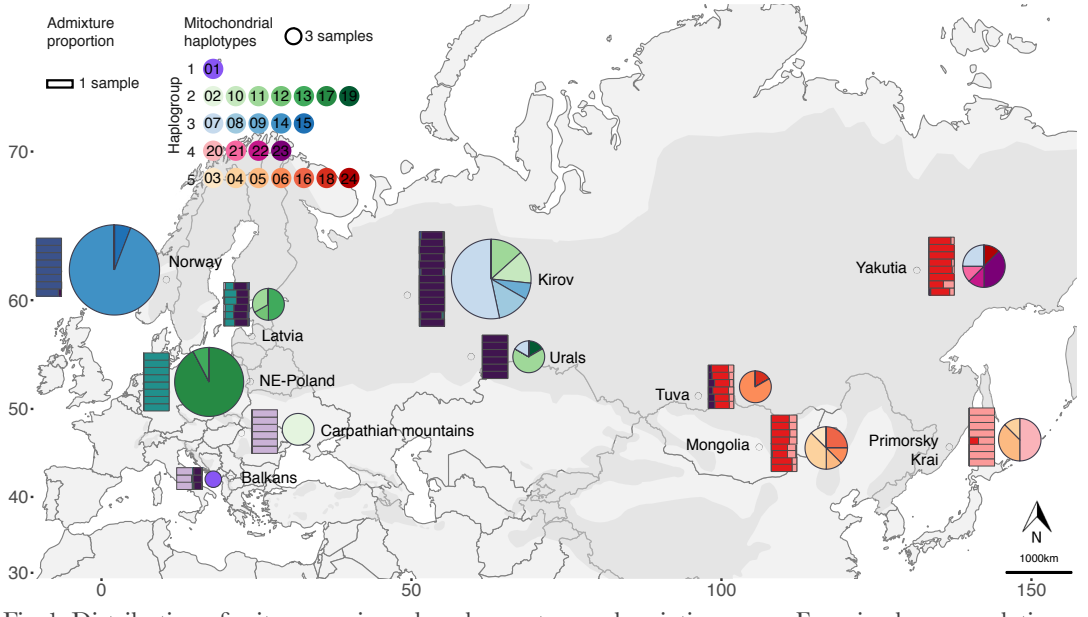


Fig. 1. Distribution of mitogenomic and nuclear autosomal variation across Eurasian lynx populations. Pie charts represent the frequency of each of the 24 identified mitochondrial genome haplotypes in each population (right), and rectangles depict the ancestry of individuals in each of six genetic clusters, as estimated with NGSadmix (left).

Primorsky Krai, Russia; 11) Mongolia (for detailed description of the populations see supplementary methods & Table S2; for a detailed description of sampling methods see supplementary methods). These populations represent different climatic and land cover zones, but also they differ with respect to demographic history and recent human exposure. Based on available records of recent demographic status (see supplementary methods & Table S2), Norway, NE-Poland, Carpathians and Balkans populations are the remnants of a process of anthropogenic range contraction in Europe initiated in the 16th century driven by habitat alteration and direct persecution which intensified by the turn of the 19th and 20th centuries by the implementation of extermination policies in several countries. The remnant central European populations recovered from their rather extreme bottlenecks following legal protection enacted during the 20th century, but have remained relatively isolated until today. In more eastern and northern parts of Europe, Latvia, Kirov region of Russia, and the Urals populations remained moderately large and/or well interconnected during this process. In Asia, we sampled Tuva, Yakutia, Primorsky Krai and Mongolia, which are considered part of a large contiguous range that has been much less affected by habitat alteration (see supplementary methods & Table S2).

DNA extraction, sequencing & mapping

Samples consisted of good quality tissue or blood, except for the Balkan samples, which were poorly preserved specimens that yielded signatures of low quality and extensive contamination (supplementary methods). All samples were digested overnight using proteinase K and gDNA was extracted using silica-coated paramagnetic beads (NucleoMag® Tissue, MACHEREY-NAGEL GmbH & Co. KG). Depending on the sequencing strategy (depth targeted and quality of the gDNA), library preparation and sequencing of the samples differed (details in supplementary methods). Briefly, gDNA was sheared, size selected, end-repaired, and adenylated following the appropriate Illumina protocol. After ligating indexed paired-end adapters, DNA fragments were amplified via PCR if required, and the quantity, quality and size of the libraries were assessed. Finally, libraries were sequenced using Illumina HiSeq2000 or Illumina HiSeq X-10, in CNAG or MACROGEN facilities, respectively. In all cases, samples

were sequenced using Illumina protocols, and primary data analysis was carried out with the standard Illumina pipeline. We performed a quality control of our data, and we trimmed and mapped our sequences to a 2.4 Gb *Lynx pardinus* nuclear reference genome, which diverges from the Eurasian lynx by an average of ~0.00122 substitutions per site (Abascal et al. 2016; http://denovo.cnag.cat/genomes/iberian_lynx/), using BWA-MEM (Li 2013) (details in supplementary methods).

To reconstruct mitogenome sequences we used raw reads coming from both this whole-genome (WG) project and a separate capture-based study (16 additional individuals, Marmesat et al. *in prep.*). Reads were mapped to the *L. lynx* mitochondrial reference genome generated by (Abascal et al., 2016) using BWA-MEM (Li, 2013) with default parameters. We called SNPs using Freebayes (Garrison & Marth, 2012) and constructed a consensus for each mitochondrial genome using the FastaAlternateReferenceMaker command in GATK (McKenna et al., 2010) (Table S1; and further details in supplementary methods).

In summary, we generated WG re-sequencing data for 80 *L. lynx* individuals, 76 at low-medium depth (4-13x) and four at high depth (19-28x), and mitogenome data for 96 *L. lynx* individuals with an average depth of 137.7x (Table S1). For analyses including all individuals, data at medium-high depth were randomly subsampled to a depth within the range of the low-medium depth data using SAMtools view -s (Li et al., 2009) to avoid biases associated with differences in sequence depth. Samples from the Balkan population were used to assemble mitogenomes and determine global autosomal genetic structure (PCA and NGSAdmix), but we did not estimate diversity and neutrality indices from these samples, as we found an excess of changes likely associated with their suboptimal conservation (see supplementary methods).

Due to the focus of our study on demographic reconstruction and neutral evolution, we only considered intergenic regions, representing 61% of the nuclear genome (~1.5 Gb), for most of the analyses (further information on neutral regions definition in supplementary methods). We also identified non-recombining parts of X and Y chromosomes that were excluded from analyses on the autosomes (i.e. PSMC, measures of autosomal diversity) and used to compare patterns of diversity with the autosomes.

Data analysis

Nuclear demographic and divergence reconstruction using PSMC

To infer changes in the effective population size (Ne) through time (3*106-104 ya) on the basis of the nuclear genome, we used a Pairwise Sequentially Markovian Coalescent model (PSMC; (Li & Durbin, 2011)). This model infers population size history from the distribution of the local density of heterozygous sites in a single diploid sequence. Therefore, for this analysis, we used autosomal whole genome data of four *L. lynx* individuals sequenced at higher depth (>19x), two from Asia (Primorksy Krai and Yakutia) and two from Europe (Carpathians and Kirov). For each individual, we generated a diploid consensus file using SAMtools mpileup as suggested by H. Li & Durbin (2011). For this analysis, minimum read depth was set to 7x and maximum read depth to 60x for all individuals.

To infer the divergence time between populations we built pseudo-diploids by randomly combining haplotypes of two individuals sampled in different populations (Cahill et al., 2016; Chikhi et al., 2018). To do so, we used the fasta files generated from the bam files during PSMC pipeline. First, we intersected the two fasta files to get the list of scaffolds represented in both samples, and then we used seqtk mergefa (https://github.com/lh3/seqtk) to randomly sample one allele from each of the two fasta. PSMC analyses were then conducted on the pseudo-diploids as described above for true diploids. We performed 100 bootstraps (Li & Durbin, 2011) for both the original and the pseudo-diploids analysis.

Nuclear demographic reconstruction using Stairway plot

We reconstructed recent demographic trajectories of the populations using Stairway plot (Liu & Fu, 2015), which infers more recent histories than PSMC (5*105-102 ya), although with limited power when sample sizes are below the hundreds (Beichman, Huerta-Sanchez & Lohmueller, 2018). The program is a model-free method that uses the unfolded site frequency spectrum (SFS) to infer population size changes over time. We generated SFS for each population using first ANGSD (Li, 2011; Korneliussen et al., 2013; Korneliussen, Albrechtsen & Nielsen, 2014) to generate the sample allele frequency (SAF) for each population, and then realSFS (Korneliussen et al., 2013) to generate the population SFS.

Nuclear demographic reconstruction using SNeP

Most recent changes in Ne (< 2*103 ya) were inferred using the relationship between linkage disequilibrium and Ne (Hill, 1981) as implemented in the software SNePv1.131 (Barbato et al., 2015). This method uses the genetic distance between markers for estimating Ne in different periods. In order to increase the contiguity between markers we converted our SNP coordinates in the Iberian lynx assembly (41700 scaffolds, N50=1.52Mb) into cat coordinates (chromosomal level assembly) using the synteny previously defined by (Abascal et al., 2016). After splitting our VCF into the different populations, we generated a map and ped file using PLINK 1.9 (www.cog-genomics.org/plink/1.9/ (Chang et al., 2015)). We used a recombination rate of 1.9 cM/Mb (Li et al., 2016) and Sved and Feldman (1973) mutation rate modifier for correcting the recombination rate. We also used the sample size correction for unphased genotypes.

PSMC, Stairway plot and SNeP outputs were plotted, scaled to time and population sizes assuming a mean generation time of five years (Lucena-Perez et al., 2018) and a mutation rate per site per generation of 6 ·10-9 (Abascal et al., 2016).

Nuclear divergence and admixture reconstruction using TreeMix

We used TreeMix v1.12 (Pickrell & Pritchard, 2012) to infer patterns of splits and admixtures between *L. lynx* populations. For this analysis, which requires a set of called variants, we followed the genome VCF (GVCF) workflow in GATK 3.4 (McKenna et al., 2010) (supplementary methods) on all our *L. lynx* populations (except Balkans; excluded from TreeMix analysis), and on the one *L. rufus* sample as outgroup. Allele counts were extracted from each of the two VCF files using a custom script, and both resulting allele count files were merged under the assumption that any SNP absent in one of the species (but present in the other) would be fixed for the reference allele. We ran TreeMix v1.12 (Pickrell & Pritchard, 2012) setting *L. rufus* as outgroup and the block size to 100. We modelled between zero and six migration events ($0 \leq m \leq 6$) and calculated the proportion of variance in relatedness between populations explained by each model. To assess the consistency of migration edges, we performed nine additional runs for $m = 2$ with different random seeds. Both the tree models and the residuals from the fit of the models to the data were visualised using the R script included in TreeMix.

We also ran the three-population test (Reich et al., 2009), as implemented in the threepop programme of TreeMix, to detect past admixture between populations. This test checks whether population X is related to populations A and B through a simple tree (in which case the f statistic, defined as the product of the frequency differences between A and X, and B and X, is expected to be positive), or through an admixture of A and B (where negative f statistic values are expected). To assess the statistical significance of the test, threepop obtains a standard error from blocks (here set to a size of 100 SNPs) and then generates a Z score. Z score values below -2 indicate significant support for admixture. This test was conducted for all possible combinations of i) three representative European populations (Carpathians, Kirov, and the Urals), ii) one Asian test population (Tuva or Yakutia), and iii) the three remaining Asian populations (Vladivostok, Mongolia, and Yakutia or Tuva depending on which was the test population).

Nuclear genomic structure

To assess the genetic relationships among samples we performed a principal component analysis (PCA). We calculated the genotype posterior probabilities using ANGSD (Kim et al., 2011; Li, 2011) and NGSTools/ngsPopGen/ngsCovar (Fumagalli, 2013; Fumagalli et al., 2013). For all the analyses using ANGSD the filters applied were: -uniqueOnly 1 -remove_bads 1 -only_proper_pairs 1 -baq 1 -C 50 -minMapQ 30 -minQ 20 -doCounts 1 –minInd (number of individuals in the population/2) –setMaxDepth (average (AVR) depth for the population + (0.95*stdev depth for the population)) –setMinDepth (AVR depth for the population - (0.95*stdev depth for the population)) -skipTriallelic 1); and we took the base observed in *L. rufus* as the ancestral state (more details on how we reconstructed our ancestral state in supplementary methods). For PCA, and also pairwise genetic distances and admixture analyses we set a SNP_pval of 1e-3. The resulting PCA was plotted using scatterplot3js from threejs library in R (Lewis, 2017). PCA coordinates were scaled to geographical coordinates (Procrustes analysis, following Borg and Groenen 1997) to assess similarities between geographic and genetic distribution using the package MCMCpack (Martin, Quinn & Park, 2011) in R (R Core Team, 2019).

Nuclear genomic structure

We used the genotype likelihoods calculated with ANGSD (Kim et al., 2011; Li, 2011) to perform a structure analysis using NGSadmix (Li, 2011; Skotte, Korneliussen & Albrechtsen, 2013). We ran NGSadmix with a range of a priori populations from K=1 to K=13. The analysis was rerun ten times to evaluate convergence and results were plotted using R. We used CLUMPAK (Kopelman et al., 2015) to evaluate the optimal K following Evanno et al. (2005). We used the identified genetic clusters to confirm that all our a priori defined populations were genetically homogeneous (i.e. all the individuals showed similar ancestry proportions and therefore could be pooled for population-based analysis).

Nuclear genomic differentiation among populations

A 2-dimensional unfolded site frequency spectrum (2d-SFS) was computed using realSFS (Korneliussen et al., 2013) for each population pair. 2d-SFS and SAF files were used as priors to calculate FST using realSFS (Korneliussen et al., 2013). To graphically visualize the genetic relationship among populations, we constructed a neighbour-joining tree based on the pairwise FST matrix using the ape package in R (Paradis, Claude & Strimmer, 2004).

To evaluate the influence of distance on genetic differentiation patterns we calculated the genetic distance among pairs of individuals using ANGSD and ngsDist from NGSTools (Fumagalli et al., 2013; Korneliussen, Albrechtsen & Nielsen, 2014). Geographical distances among sampling points were calculated from their geographic coordinates using the Point Distance tool and Winkel Tripel projection in ArcGIS 10.5 (Esri, 2011). All distances involving samples from Norway were measured via the Scandinavian isthmus to estimate overland distance. Based on our previous results on PCA and admixture proportions, we plotted geographical distance against genetic distance splitting our results depending on whether the comparison was between two samples from Asia, two samples from Europe or one sample from Europe and one from Asia. We tested the significance of these correlations with Mantel tests using the package vegan (Oksanen et al., 2018) in R. We also performed a partial Mantel to test the effect of a possible geographic barrier between Asian and European populations by introducing a binary variable coded as 0 when the two samples were from the same region and as 1 otherwise, and tested the effect of this variable while accounting for the effect of geographic distance.

Nuclear genomic diversity

We calculated the genetic diversity (nucleotide diversity (π) and Watterson estimator (θ)) and Tajima's D neutrality index for each population. For autosomal and X chromosomes, we used ANGSD (Li, 2011; Korneliussen et al., 2013; Korneliussen, Albrechtsen & Nielsen, 2014) and realSFS (Korneliussen et al., 2013) to calculate diversity indices per site for each population (Korneliussen et al., 2013). Using thetaStat (Fumagalli et al., 2014), we performed a sliding-window approach with a window size set to 50,000 bp and a step size of 50,000. We classified our windows as autosomal or X chromosome if all the sites of the given windows belonged to either of these categories (supplementary methods). For comparison among populations, only windows with information for all the populations were used (828 X chromosome, and 24,392 autosomal windows). Sample size of the X chromosome differs from that of autosomal chromosomes (which is 2*number of individuals), as it depends on the number of males and females sampled in the population (2*number of females + number of males). Therefore, we recalculated our Watterson estimator (θ) by adjusting the correction factor (that accounts for sample size) to the actual sample size of the X chromosome in each population. Standard errors were calculated by bootstrapping over windows as implemented in the boot package for R (Canty & Ripley, 2017), to account for the correlation among nearby sites due to linkage disequilibrium (LD).

To infer recent population size changes we compared X chromosome θ vs. autosomal θ, while controlling for divergence (Pool & Nielsen, 2007). Divergence (D) was computed as the number of substitutions between *L. lynx* and *L. rufus* divided by the number of covered sites, based on a genus-wide variant-calling performed following the genome VCF (GVCF) workflow in GATK 3.4 (McKenna et al., 2010) (supplementary methods), using one sample of each of the four species that comprise the Lynx genus (*L. lynx*, *L. pardinus*, *L. canadensis*, and *L. rufus*). The ratio between diversity and divergence (θ/D) was used as a measure of diversity normalized by mutation rate. The average θ/D ratio was calculated for X chromosome (X) and autosomal (A) windows and these were used to obtain an X/A ratio for each population. Standard errors of X/A ratios were calculated by bootstrapping over windows. We used R to plot SFS and diversity values, as well as to plot our diversity estimates along with those reported for other mammals for comparison.

For the Y chromosome, we assigned 71 contigs, adding up to 33,032 bases, by applying our strict sequence depth criteria; i.e. 90% of the total bases in a contig have a female/male ratio depth below 0.3, and an average normalized depth for males between 0.2 and 0.8 (supplementary methods, section Chromosome X and Y regions definition and molecular sexing). We performed a SNP calling using Freebayes, as previously done for the mitogenome, but no SNPs were called under standard quality filters, suggesting overall lack of variation in this chromosome, as previously reported by Hellborg & Ellegren (2004) on the basis of 2040 bp of non-coding genic sequences.

Mitogenomic analyses

Consensus mitogenome sequences were aligned and collapsed into distinct haplotypes using the pegas R package (Paradis, 2010). The number of segregating sites (S), haplotype diversity (Hd), nucleotide diversity (π) and the mean number of pairwise nucleotide differences (k) were calculated using the PopGenome (Pfeifer et al., 2014) and ape (Paradis, Claude & Strimmer, 2004) R packages. Pie chart nodes representing the respective haplotype frequency in each of the populations (Fig. 1) were calculated using R (R Core Team 2019). Phylogenetic relationships among haplotypes were inferred by constructing a median-joining haplotype network (Bandelt, Forster & Röhl, 1999) and represented in PopART (Leigh & Bryant, 2015) coloured by population. R was used to plot k values for different mammal species for comparison.

In order to estimate a substitution rate for the mitochondrial genome in felids, we used a set of 15 felid mitogenomes downloaded from GenBank and BEAST v2.4.8 (Bouckaert et al., 2014 see supplementary methods for details). We incorporated the clock rate estimated as the substitution rate in the BEAST analysis of the 24 *L. lynx* mitogenome haplotypes obtained in our study (Bouckaert et al., 2014). For the intraspecific analyses, we used an ensemble of tRNAs, rRNAs and twelve genes using the HKY+G site model with a single partition. We used a strict clock and a coalescent constant population tree model, adequate when dealing with intraspecific sequences. The MCMC was run for 20 million steps. TreeAnnotator v2.4.8 was used to obtain the Maximum Clade Credibility tree after discarding 10% of initial trees as burn-in. Results were visualised with FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

RESULTS

Demographic and divergence history based on autosomal data

Demographic histories inferred from the WG sequences of four *L. lynx* individuals (from the Carpathian Mountains, Kirov region, Yakutia, and Primorsky Krai populations) using PSMC are concordant throughout most of the reconstructed period, indicating a long span of shared history between all populations (Fig. 2A; Fig. S1). *L. lynx* shows a long period of soft decline going from 3 Mya to 600 kya, and then declining more steeply to around 200 kya. This steep decline is followed by an apparent recovery of the population until 70 kya, although this could instead indicate the emergence of population structure (Mazet et al., 2016; Chikhi et al., 2018). From that point on, the demographic trajectories of European and Asian lynxes start diverging, with the European populations experiencing a sharper decline in Ne than the Asian ones (Fig. 2A; bootstraps presented in Fig. S1). Accordingly, pseudo-diploids constructed by combining Asian-European haplotypes start rising over the trajectory of true diploids around 100 kya, indicating the emergence of population structure, until they sharply increase around 15-20 kya (Fig. 2B; Fig. S1), indicating the time of complete population isolation between European and Asian populations. Unlike Asian-European pseudo-diploids, we did not observe a sudden and sharp increase indicative of complete isolation in pseudo-diploids consisting of two Asian or two European haplotypes. However, the moderate increase in population sizes with respect to real diploids suggests the emergence of some structure in Asia following the split of Asian and European populations, and little or no structure in Europe (at least) before 10 kya.

More recent demographic trajectories inferred for all the populations using the Stairway plot and SNeP are broadly congruent with the pattern inferred by PSMC (Fig. 2C). The Stairway plot reconstructs a steep decline process during the last millennia, with populations from Europe showing consistently smaller population sizes than Asian populations while SNeP reconstructs a smoother population decline spanning the last two millennia again with European populations showing smaller Ne (Fig. 2D). Both reconstructions show a more moderate decline for the European Kirov region population.

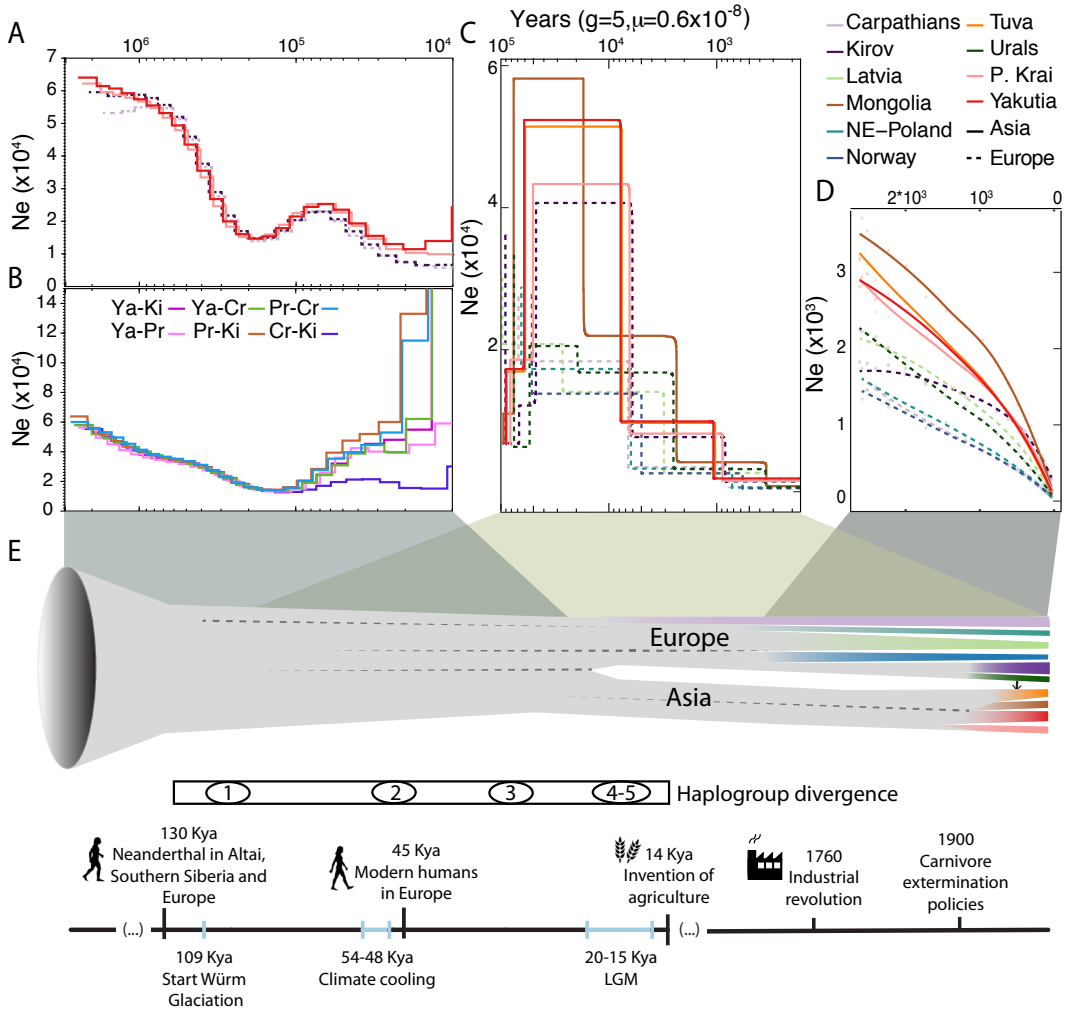


Fig. 2. (A) Demographic reconstructions inferred with PSMC on the basis of autosomal data from one individual from Kirov, Yakutia, Primorsky Krai and Carpathians. (B) Pseudo-diploid trajectories on PSMC. Sequences created by combining haplotypes from two different populations. Sudden population increases inferred for pseudo-diploids are interpreted as the time of complete isolation of the two populations. (C) Demographic reconstruction inferred by Stairway plot for all populations. (D) Recent demographic reconstruction inferred by SNeP for all populations. (E) Timeline of Eurasian lynx main demographic events, including partial (dashed lines), complete isolation (branching) and admixture (arrow) between populations, along with haplogroup divergences, major climatic fluctuations and human milestones.

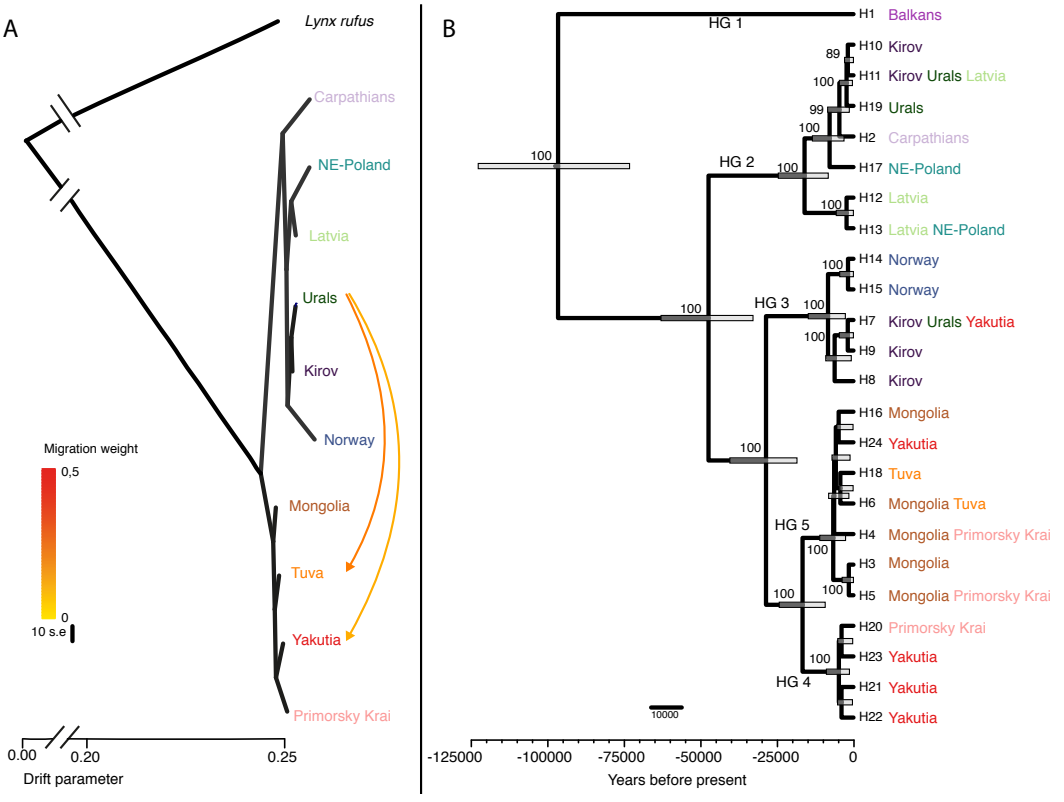


Fig. 3. (A) Population tree inferred with TreeMix from nuclear autosomal data. Each arrow represents a migration event with its weight specified through the color scale. (B) Bayesian maximum clade credibility tree inferred with BEAST depicting the relationship among mitogenome haplotypes. Numbers on nodes represent their posterior probability; only those over 0.85 are depicted.

Considering all populations, the TreeMix analysis based on autosomal intergenic SNPs and using the bobcat (*Lynx rufus*) as outgroup supported a model of population divergence with one or two migration events (Fig. 3A). In line with the PSMC pseudo-diploids results, the most basal split separates a European and an Asian population group, each with some shallower internal structure (Fig. 3A, Fig. S2). In the European group, Norway, NE-Poland, and Carpathians show an increased drift parameter, indicating their larger differentiation compared to the rest of the populations.

Regarding gene-flow, the TreeMix results support significant post-divergence gene flow from the Urals to Tuva population (m=1), or from the Urals to Tuva and to Yakutia population (m=2; Fig. 3A; residuals presented in Fig. S3). Both these migration edges consistently appear in all nine independent runs at m=2 (Fig. S4),

and they are still supported in tree models with three to six migration events (Fig. S2; see fraction of variance explained by each of the models in Fig. S5). Significantly negative Z scores ($Z < -2$) resulting from f statistics using threepop confirmed the admixture in Tuva, but not in Yakutia (Table S3), and identified all western and eastern populations, except Yakutia, as putative sources, with more negative Z scores the closer the western source population is (Ural Mountains < Kirov < Carpathian Mountains).

Mitogenomic divergence

Unlike the pattern inferred for the nuclear genome, the phylogenetic tree reconstructed from whole mitogenome sequences revealed several old mitochondrial lineages in Europe, which may have diverged during periods of isolation in separate glacial refugia (Fig. 3b). The oldest split was dated around 96.5 kya (95% CI: 73-122 kya), and separates the most divergent haplogroup, haplogroup 1 (46 out of 89 segregating sites), currently restricted to the Balkan population. A second mitogenomic split, dated around 47.4 kya (95% CI: 32-62 kya) defines haplogroup 2, whose current distribution includes the Carpathian Mountains and Baltic states populations. A subsequent split around 28.6 kya (95% CI: 18-40 kya) separated haplogroup 3, occurring mostly in northern and eastern Europe, from all the Asian haplogroups (4 and 5). Finally, the split between haplogroup 4 and 5 around 17 kya (95% CI: 9-24 kya) is coincident with an internal diversification within haplogroup 2.

Current nuclear genetic structure

Our a priori populations vary in the extent of the area sampled, what may have affected our estimations of genomic diversity and differentiation. Nevertheless, individual-based clustering and PCA analyses confirmed the relative genetic homogeneity within and differentiation among our a priori defined populations, i.e. ancestry composition is similar among individuals of the same population and different between populations.

In agreement with the historical divergence and demographic processes shown in the previous section, nuclear population data reveals a contemporary major division between individuals in Asia and Europe. Both PCA and NGSAdmix separate these clusters first (PC1, $K=2$, supported as the uppermost level of structure by the Evanno et al. (2005) method (Table S4)) (Fig. 4, Fig. S6; interactive version in original article).

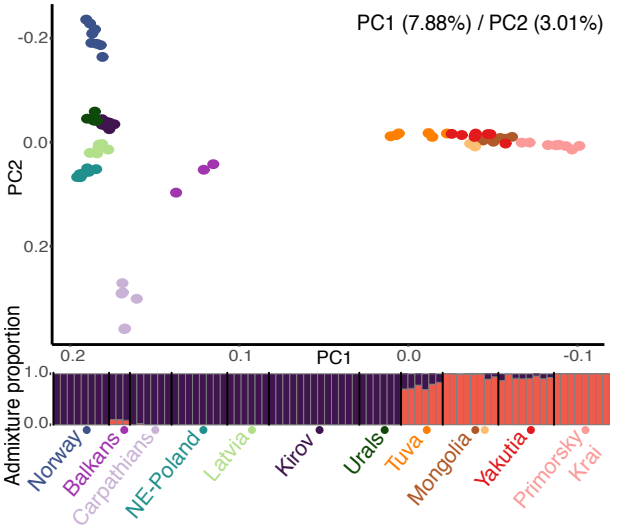


Fig. 4. Relationship among individuals based on nuclear autosomal genotypes. PCA separates eastern and western individuals in the first axis (7.88% of the variance explained), and westernmost populations in the second axis (3.01% of the variance explained). Individual ancestry in each of the two clusters defined in the NGSadmix analysis. Populations are sorted from west to east. Two different colors represent Mongolia population, as this population comprises two different habitats (orange representing Ömnögovi, and brown Central and Khentii Aymag).

Two populations show some admixed ancestry in these analyses: Balkan individuals with some Asiatic ancestry and individuals from Tuva (as well as, to a lesser extent, individuals from Yakutia) with some European ancestry, the latter supporting the historical admixture inferred by TreeMix. The historical isolation of European and Asian populations is also supported by a larger genetic distance between pairs of Asia-Europe individuals than that expected due solely to geographic distance (Fig. S7).

The second and third axis in the PCA, and subsequent partitions in the structure analysis (K=3, K=4 and K=5), separate bottlenecked populations in Europe (Carpathian Mountains, Norway, and NE-Poland) (Fig. 4, Fig. S9-S11; supplementary file 1). The high differentiation exhibited among European populations (average pairwise FST = 0.210, Fig. S8, Table S5) contrasts with the relative homogeneity in Asia (average pairwise FST = 0.098, Fig. S8, Table S5), where only the easternmost population of Primorsky Krai stands out as a separate cluster in some runs at $3 \leq K \leq 6$ (Fig. 1, Fig. S9-S12). Accordingly, when PCA coordinates are projected onto a map, European populations show a major distortion in the projection compared to Asian ones, suggesting an isolation by distance scenario in the case of Asian populations, but a higher differentiation that is not explained solely on the basis of distance for the westernmost populations (Fig. S13, S14). The greater differentiation associated with bottlenecked populations (NE-Poland, Norway and Carpathians) is also observed in isolation by distance plots, where the inclusion of bottlenecked populations increases the slope of the regression line (Fig. S15).

Autosomal genetic diversity

Genetic diversity levels in Eurasian lynx are in the low range of those reported based on genome-wide data for other mammals, including some rare and endangered populations (Fig. S16). This is especially the case for the bottlenecked European populations: NE-Poland, Carpathian Mountains and Norway (Fig. 5, Fig. S16, S17, Table S6). These westernmost populations also show a flatter site frequency spectrum compared to the rest of the populations (Fig. S18), and a significantly positive Tajima's D value (Table S6), both indicative of a recent reduction in population size. NE-Poland, followed by Carpathians, also show the lowest X/A ratio, an additional indication of a recent, severe population size reduction (Pool & Nielsen, 2007). The X/A ratio in Norway is, however, larger and comparable to that found in Ural Mountains, a pattern probably related to its subsequent growth (Pool & Nielsen, 2007). Conversely, the Urals and the Kirov regions show the highest diversity within Europe (θ and π, respectively), but are exceeded by all Asian populations, with the highest genetic diversity present in Tuva despite the relatively small geographic area sampled (Fig. 5, Fig. S17).

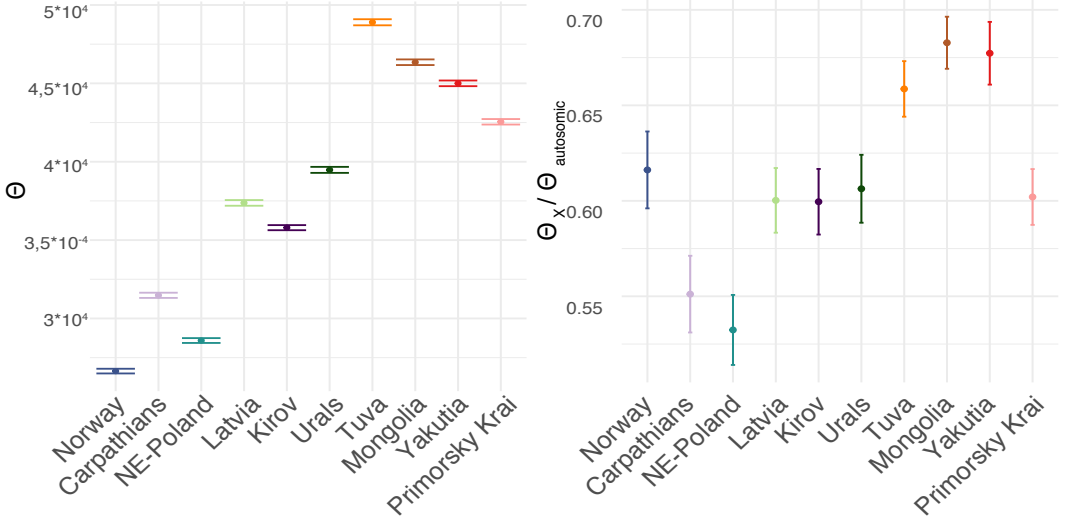


Fig. 5. Waterson's theta (θ) values for autosomal sites and ratio of θ diversity in Xchr vs. autosomes for the different populations. Populations are sorted from west to east.

Mitochondrial structure and diversity

In agreement with the autosomal data, overall mitogenomic diversity at the species level is very low (Fig. S19). Only 89 segregating sites, out of 16,449 sites sequenced, defined the 24 haplotypes found (0.54%, Table S7, Table S8). We observed several populations in which European haplogroups 2 and 3 (e.g. Kirov, the Urals) or Asian haplogroups 4 and 5 co-occur (e.g. Yakutia) (Fig. 1, 3, Fig. S20), but we detected only one instance of the occurrence of a typically European haplotype (haplotype 7, haplogroup 3) in an Asian population (Yakutia), and none in the opposite direction. The westernmost bottlenecked populations also harbour lower diversities in the mitochondrial genome: Carpathians and the Balkans show no mitogenomic diversity at all, while Norway and NE-Poland populations also have low haplotype diversity (Hd = 0.12 and 0.15, respectively; Fig. 1, Fig. S20, Table S7); populations in Asia (Yakutia and Mongolia) show the highest haplotype diversity (Hd = 0.86 in both cases; Fig. 1, Fig. S20, Table S7).

DISCUSSION

Here we report the results of the most comprehensive analyses to date of the Eurasian lynx's evolutionary history and contemporary genetic variation, considering the extent of the sampling both across the species range and across the genome. Our results indicate that the Eurasian lynx was genetically quite homogenous at least until 100 kya, when lynx populations started diverging and the species entered a widespread and continuous demographic decline that affected especially the European populations. Lower populations sizes and increased fragmentation in the westernmost part of the distribution in more recent times likely drove the genetic differentiation between European populations that are otherwise geographically and ecologically close. Conversely, and despite the large range and the wide diversity of habitats, we observed a highly homogenous genetic pattern among Asian population, compatible with an isolation by distance pattern. Climatic oscillations during the Late Pleistocene, together with an increasing human pressure especially after the Last Glacial Maximum (LGM) which ended up with the extirpation of many European populations, likely shaped the current genetic patterns of the species.

Pleistocene

Phylogenetic and demographic analyses support a long common history for Eurasian lynx populations during most of the Pleistocene. The species suffered a massive demographic decline registered by PSMC between 3 Mya and 200 kya (Fig. 2A), which may be reflecting a founder effect associated with the speciation from a common ancestor of Eurasian and Canada lynxes (recently dated around 1-1.2 Mya; (Li et al., 2019). Low species-wide genetic diversity, which is comparable to white African lion, or Greenlandic brown bear (Fig. S16), might be - at least partially - the long-term consequence of this drastic reduction in population size (Frankham et al., 2011; Frankham, 2015). From 200 to 70 kya PSCM shows an apparent increase in population size, after which the species entered in a continuous population decline (Fig. 2A). This transient population size increase is most likely indicating instead the emergence of population structure (Mazet et al., 2016; Chikhi et al., 2018), as suggested by the departure of demographic trajectories of pseudodiploids in PSMC, and by the deepest divergence of mitochondrial haplogroups around 100 kya (Fig. 2B, 3B). The isolation between eastern and western populations became complete around 22-15 kya coinciding with the LGM (20-15 kya), and with the time of the most recent split that is registered in both nuclear and mitogenomic data (Fig. 2, 3). At a nuclear level, clustering and differentiation analyses reflect a basal divergence between Asian and European populations (Fig. 1, 4), and indicate that the genetic differentiation between individuals from the two clusters is larger than expected solely on the basis of geographic distance (Fig. S7). This east-west axis of differentiation was previously reported by Rueness et al. (2014). The lack of clear geographical barriers for lynxes during this period supports the idea that the emergence of the structure was preceded by a range contraction. For instance, barriers previously identified for small mammals between the two continents, including the Yenisei river (Kohli et al., 2015) and the Ural Mountains (Brunhoff et al., 2003), are unlikely to act as strong barriers for such a large and mobile carnivore (Zimmermann & Breitenmoser, 2007).

Within continents, European populations underwent periods of isolation during the Late Pleistocene, as revealed by the divergence of different mitochondrial lineages, whereas populations in Asia remained largely connected as shown by nuclear and mitochondrial data (Fig. 2B, 3B).

Climatic impacts during the Pleistocene

Climatic fluctuations during the Late Pleistocene could have contributed to range contraction and subsequent emergence of populations' structure and divergence revealed by nuclear and mitochondrial data (Fig 2, 3). Accordingly, the date of the complete isolation between Asian and European nuclear genetic groups is coincident with the LGM (20-15 kya) (Svendsen et al., 2004; Schlolaut et al., 2017) (Fig. 2E). Older periods of isolation and divergence due to climatic oscillations during the Pleistocene are only registered in mitogenomic patterns, with the divergence of different haplogroups coinciding with the start of glaciation periods (HG1 divergence, 96.5 kya, start of Würm glaciation period (Ice age), 109 kya), or significant global climate cooling (HG2 divergence, 47.4 kya, 48–54 kya (Kindler et al., 2014); HG4-HG5 divergence and internal diversification of HG2, 17 kya, LGM) (Fig. 2E). At least two of these haplogroups, HG1, and HG2, seem to have diverged in separate glacial refugia, the Balkans and the Carpathians, respectively, as previously suggested for lynx and other temperate mammals (Anijalg et al., 2018; Bilton et al., 1998; Gugolz, Bernasconi, Breitenmoser-Würsten, & Wandeler, 2008; Ratkiewicz et al., 2014; Schmitt & Varga, 2012; Sommer & Nadachowski, 2006; Taberlet, Fumagalli, Wust-Saucy, & Cosson, 1998). HG3, now present mainly in northern and eastern Europe, suggests a possible additional north-eastern refugium for the species in Europe.

Deep mitochondrial divergences contrast with the recent isolation within Europe inferred from nuclear data. This discrepancy could be attributed to male-biased dispersal, which is characteristic of the species (Schmidt, 1998; Holmala et al., 2018). Faster and deeper divergence of mitogenomes during periods of isolation followed by admixture between European populations, driven mostly by males, during the interglacial periods could have led to the pattern that we observe today. For instance, nuclear data suggests the existence of gene flow between the Carpathians and Kirov at least until 10 kya (Fig. 2), while the divergence of the mitogenomic haplogroups typical of these populations was dated to ~50 kya (Fig. 3). The co-occurrence of different haplogroups in different populations suggests that, to a minor extent, females also contributed to the admixture of the populations. For instance, the Carpathians population with a unique haplotype (2), which is basal to HG2, could have acted as a source for postglacial colonization of this HG northward and eastward, as suggested for the fact that the HG2 is also present in populations such as Kirov, Urals

or Latvia. Similar scenarios of isolation in refugia during glacial periods followed by colonization during interglacials have been described for species with similar habitats requirements, such as brown bear (Anijalg et al., 2018), or grey wolf (Vila et al., 1999; Pilot et al., 2010).

Anthropogenic impacts during the Pleistocene

Besides climatic fluctuations, demographic declines during the Late Pleistocene could also be influenced by hominid species that were probably already widespread across Eurasia around 80 to 60 kya (Oppenheimer, 2012; Timmermann & Friedrich, 2016), and more dramatically by modern humans, who arrived in Eurasia around 45 kya and replaced other human populations, exceeding their population size by one order of magnitude (Mellars & French, 2011; Timmermann & Friedrich, 2016; Yang et al., 2017) (Fig. 2E). Direct human impacts on lynx species have been documented during the Late Pleistocene, where a bone remain found in an Upper Palaeolithic site in the Iberian peninsula revealed the use of lynx as meat (Yravedra, 2005). Similarly, bones of leopard (Panthera pardus) have been found associated with hunting by pre-historic humans throughout Europe (reviewed in (Sommer & Benecke, 2006)). Additionally, the decline of ruminants since 100 to 50 kya has been partially attributed to human activities, rather than climatic oscillations (Chen et al., 2019), and hence, the lynx decline during the Late Pleistocene could have been also indirectly attributed to humans through negative effects on prey. Our hypothesis is in line with previous work that support that anthropogenic impacts in combination with climatic oscillations are one of the main drivers of the decline, and in some cases the extinction, of fauna and flora during the Late Pleistocene and early Holocene (Lorenzen et al., 2011; Braje & Erlandson, 2013; Chen et al., 2019; Gretzinger et al., 2019).

Holocene

The population decline of the species continued after the LGM with some differences among populations (Fig. 2). The sustained negative population trends during the Holocene likely contribute to the signals of recent bottlenecks, such as reduced X/A ratios and high Tajima's D values that all populations show. Additionally, even the most diverse population show values of genetic diversity similar to that of the severely bottleneck Apennine brown bear (Benazzo et al., 2017), and only twice that of the extremely eroded sister-species - the Iberian lynx (θ=2.22*10-4, π=2.6*10-4

(Abascal et al., 2016)), whose values are comparable to the least diverse Eurasian lynx populations (Fig. S16). Still, differences in recent demography between populations, with European populations experiencing a severe reduction in population size throughout the Holocene and Asian populations usually maintaining a softer population decline, is reflected in current genetic patterns: European populations, especially westernmost ones, show larger genetic differentiation, increased drift parameters in TreeMix analysis, along with bottleneck signals and lower genome diversity. (Fig. 4-5 S8-S12, S17; Table S5- S6). Patterns of low diversity and high differentiation were previously reported for NE-Poland (Schmidt et al., 2009; Ratkiewicz et al., 2012, 2014), as well as for Scandinavia and the Carpathians, using nuclear microsatellite markers and short mitochondrial sequences (Ratkiewicz et al., 2012, 2014) or a SNP set enriched for coding sequences (Förster et al., 2018). In contrast to westernmost populations, the current structure among Asian populations is shallow, similar to that found for Canada lynx (FST = 0.09-0.10; Meröndun , Murray & Shafer, 2019), the overall pattern is compatible with an isolation by distance scenario (Fig. S14), and there is little support for more than one genetic cluster. Quite homogeneous genetic patterns across Asia are somehow striking given the range of habitats occupied (e.g. from semi-desert in Omnogovi, Mongolia, to boreal forest-tundra in Yakutia), and the several previously defined subspecies in this region.

Anthropogenic impacts during the Holocene

The invention of agriculture produced a rapid demographic expansion that started after LGM in Europe, but only few millennia ago in Asia (Gignoux, Henn & Mountain, 2011; Skoglund et al., 2014; Nielsen et al., 2017), which eventually ended up in the emergence of urbanised and industrialised nation-states. A more extense, contiguous and less anthropogenically altered habitat in Asia during the Holocene might have contributed to the homogeneous genetic pattern in the continent, while in Europe, a higher anthropogenic pressure, intensified on recent times as documented in historical records (Table S2), resulted in genetically structured and eroded populations. Carnivore extermination policies at the turn of 19th and 20th centuries extirpated the species from most of central Europe. Subsequent protection prevented total extirpation in remnant populations and allowed some recovery, but today lynx populations in this region remain highly isolated from each other and from the more contiguous range further north and east (Hellborg et al., 2002; Schmidt et al., 2009; Ratkiewicz et al., 2012). Demographic declines and genetic isolation during the

last century have been particularly intense in Norway (together with neighbouring Sweden) and NE-Poland. In Norway, the population was restricted to a few survivors in the central region from 1926 to 1965, although it has steadily recovered since then, apparently with little contribution of immigrants from outside Scandinavia (Linnell et al., 2010). In NE-Poland the population became restricted to the Białowieża Primeval Forest (BPF), with apparent absence of lynx from 1890 to 1914 (Jędrzejewski et al., 1996; Bieniek, Wolsan & Okarma, 1998), followed by a short bottleneck during the 1960s and 1970s and a modest recovery assisted by immigrants by the end of the 20th century (Jędrzejewski et al., 1996). In contrast, the Carpathians population has been considered relatively large, although largely isolated from other lynx populations, and has been used as the source of animals for reintroductions in central/western Europe (Von Arx et al., 2004). However, this population has not been exempted from intense direct persecution that left around 100 individuals in Romania by 1930 (Kratochvil, 1968), and probably similar numbers in the Slovakian part (Hell & Slamečka, 1996). The protection of the species in both countries in the 1930s allowed a significant recovery of the population with around 500 individuals in Romania by 1950 and 400-500 in Slovakia in the period 1960-1990 (Kubala et al., 2019). Similar scenarios with western fragmented and bottlenecked populations due to continued human pressure during the Holocene versus more contiguous and stable ranges in the east have been also postulated for other carnivore species such as wolf and brown bear in Europe (Adamec et al., 2012; Pilot et al., 2014; Hindrikson et al., 2017).

Implications for conservation, management, and taxonomy

For conservation and taxonomic purposes, it is critical to formally assess possible adaptive intraspecific divergences within the Eurasian lynx, a possibility that is now made feasible by the availability of genomic data. Even though we find little to no support for most of the subspecies discussed in the literature in recent years, the finding of shallow differentiation at neutral regions of the genome does not exclude the occurrence of locally divergent selection at particular genes, i.e. local adaptation. We can also not discard that the morphological variation that sustained previous subspecific delimitations are plastic responses to local environments mediated by epigenetic changes, as suggested by a recent study of *Lynx canadensis* (Meröndun, Murray & Shafer, 2019). Nevertheless, given our results, we argue that management plans should focus on reversing the demographic trends to prevent further genetic erosion in the most affected populations, and allowing natural evolutionary processes,

including the facilitation of population connectivity through migratory corridors in human-altered habitats, as already postulated for lynx and other large carnivores in Europe (Boitani et al., 2015). Management based on maintaining the current distinctiveness of endangered European populations may not be warranted given the shared history and sustained historical gene flow inferred in this study, and should be balanced against risks of inbreeding depression, which in the absence of further conclusive evidence are likely to exceed those of outbreeding depression in these populations (Frankham et al., 2011; Frankham, 2015).

AUTHOR CONTRIBUTION

K.S., J.A.G. and M.R. conceived the project and designed the study; A.P.S., G.N., I.O., I.V.S., J.A.G., K.S., M.G.D., and M.P. provided samples; E.M., K.W., and M.L.P. performed the lab work; B.M.C., D.K.R., E.M., K.W. and M.L.P. analyzed data; B.M.C., D.K.R., E.M., J.A.G., K.S., M.L.P., and M.R. interpreted the results; M.L.P. drafted the manuscript with support from B.M.C., D.K.R., E.M., and J.A.G. and critical input from A.P.S., K.S., and M.R; J.A.G. supervised the project. All authors approved the final version of the manuscript.

ACKNOWLEDGEMENTS

This project was funded by the National Science Center, Poland through project no 2014/15/B/ NZ8/00212, project "Biodiversity of East-European and Siberian large mammals on the level of genetic variation of populations" (BIOGEAST), contract No. PIRSESGA-2009-247652, and by the Dirección General de Investigación Científica y Técnica of the Spanish Ministerio de Ciencia e Innovacion (MICINN) through projects CGL2013-47755-P and CGL2017-84641-P, and by the Russian Science Foundation (project No. 18-14-00093). Samples from Kirov, the Urals and most Siberian populations were obtained as part of the Russian Science Foundation project No. 18-14-00093. The project leading to these results has received funding from "La Caixa" Foundation (ID 100010434), under agreement LCF/BQ/SO14/52250035 and LCF/BQ/SO15/52260006. MLP and DKR were supported by PhD contracts from the Programa Internacional de Becas "La Caixa-Severo Ochoa". Logistical support was provided by the Laboratorio de Ecología Molecular (LEM-EBD) certified

to ISO9001:2015 and ISO14001:2015 quality and environmental management systems, and Laboratorio de SIG y Teledetección (LAST-EBD) in the Estación Biológica de Doñana. Calculations were carried out in the Genomics servers of Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD). Authors also acknowledge the computing and storage resources provided by Fundación Pública Galega Centro Tecnolóxico de Supercomputación de Galicia (CESGA). EBD-CSIC received support from the Spanish Ministry of Economy and Competitiveness under the 'Centro de Excelencia Severo Ochoa 2013-2017' program, grants SEV-2012-0262. We also acknowledge Ostrava Zoo (Check Republic) and Jerez Zoo (Spain) for providing samples. We are very thankful to Vanina Tonzo for helping with the Stairway plot analysis. The license for Eurasian lynx live-trapping and blood sampling in Poland was obtained from the National Ethics Committee for Animal Experiments (no. DB/KKE/PL—110/2001) and the Local Ethics Committee for Animal Experiments at the Medical University of Białystok, Poland (no. 52/2007). Import of samples was licensed by CITES permissions no 12RU000512 and 15NO-046-EX, respectively. We are grateful to Ingrid Reinkind for kind providing the Eurasian lynx samples from Norway. Besides the free software mentioned in the text, this work was possible thanks to the free, open source packages: dplyr (Wickham, Francois, Henry, & Müller, 2017), ggplot2 (Wickham, 2009), tidyr (Wickham & Henry, 2018), broom (Robinson, 2018), RColorBrewer (Neuwirth, 2014), stringr (Wickham & Henry, 2018), gridExtra (Auguie, 2017), knitr (Xie, 2014, 2015, 2017), reshape2 (Wickham, 2007), lemon (Edwards, 2017), data.table (Dowle & Srinivasan, 2017), rlang (Henry & Wickham, 2018), png (Urbanek, 2013), and rgdal (Bivand, Keitt, & Rowlingson, 2017).

10 APÉNDICE

GENOMIC PATTERNS IN THE WIDESPREAD EURASIAN LYNX SHAPED BY LATE QUATERNARY CLIMATIC FLUCTUATIONS AND ANTHROPOGENIC IMPACTS

SUPPLEMENTARY METHODS

History and status of the sampled populations

Information on the recent and current demographic status of Eurasian lynx populations is large lacking and mostly unreliable, being mostly based on unvalidated expert opinion and hunting statistics. Nevertheless, the populations sampled do cover a range of recent demographic histories and exposure to anthropogenic impacts (Table S2). Westernmost populations in Europe are the remnants of an ancient range that started its contraction as early as the 7th century when the species disappeared from Great Britain and followed by a progressive eastward withdrawal in continental Europe starting probably in the 16th century and peaking in the 19th and 20th century with the implementation of extermination policies in many countries. Such intense persecution led to the extermination of the species in most of western Europe, with the few westernmost populations surviving going through intense and extended bottlenecks which were followed by a slow recovery once the species became protected in the 1930s-1940s. Bottlenecks and genetic isolation have been particularly intense in the Balkans, Norway (together with neighbouring Sweden) and NE-Poland, which were at the brink of extirpation during the first half of 20th century and with little or no gene flow from other lynx populations. The Carpathians Mountains (mainly Romanian and Slovakian parts) went also through important contractions leaving 100-200 individuals by the 1930s, but recovered to more than 1000 individuals by the second half of the 20th century. The Carpathian population has been since considered relatively large, but largely isolated from other lynx populations, and has been used as the source of animals for reintroductions in western Europe (Von Arx et al., 2004). However, the first model-based estimates obtained recently suggest lynx abundance in the region has been overestimated in the past (Krojerová-Prokešová et al., 2019; Kubala et al., 2019).

Latvia is part of the northwestern population, together with Finland, western Russia, and Estonia. While it also went through declines starting in the 17th century and intensifying in the early 20th century, the population recovered and now occupies its full historical range with 600-800 individuals estimated at present. Importantly, the population remained well connected to the large and continuous core population farther east in European Russia, which is represented in our sampling by Kirov and the Urals.

Information on the lynx populations in Asia is even more scarce, scattered and mostly consists of fur harvest data, which not only reflect population numbers but also socio-economic circumstances. The Asian distribution is generally considered continuous and with high densities, although these typically fluctuate due to natural (prey availability) and anthropic causes (harvest pressure), with amplitudes ranging from 7 to 13 years, being larger at northern latitudes (Matyushkin & Vaisfeld, 2003). Asian lynxes underwent also a heavy exploitation during last centuries, dating back to the XVII century, with 10 000 lynx pelts were recorder yearly in Siberia only (Brass, 1911). This, however, did not lead to the extirpation of entire populations typical for the bottlenecked populations in the western peripheries of the species range.

Samples

We sampled 80 *L. lynx* across the distribution range of the species, including individuals for five out of the six subspecies proposed by the IUCN Cat Specialist Group (*L. l. lynx*, *L. l. balcanicus*, *L. l. carpathicus*, *L. l. isabellinus*, and *L. l. wrangeli*) (Kitchener et al., 2018)). Majority of samples were tissues collected from legally hunted individuals (Norway, Latvia, Russia, and Romania; n = 68) or animals found dead (Poland, Carpathian Mountains, and Mongolia; n = 19). Three museum specimens (Balkans) and six blood samples of live-trapped lynx (Białowieża Primeval Forest in NE Poland and Carpathians Mountains, in Poland) were also obtained. Sampling of harvested individuals was opportunistic – the lynx harvesting was not intended for obtaining genetic samples. The license for lynx live-trapping and blood sampling in Poland was obtained from the National Ethics Committee for Animal Experiments (no: DB/KKE/PL – 110/2001) and the Local Ethics Committee for Animal Experiments at the Medical University of Białystok (no: 52/2007). These licenses adhere to the Directive 2010/63/EU on the protection of animals used for scientific purpose. No animals were harmed during live-trapping and handling.

Genome re-sequencing

Samples were sequenced at Centro Nacional de Análisis Genómico (CNAG-CRG) or Macrogen using different sequencing strategies (Table S1). These strategies depend on the quality of the gDNA extracted or the goal depth.

Library preparation

Library preparation 1. Good quality samples aimed to be sequenced at low to medium depth. gDNA samples were used for preparing Illumina sequencing compatible paired-end libraries using NO-PCR protocol, TruSeq™DNA Sample Preparation Kit v2 (Illumina Inc.) and the KAPA Library Preparation kit (Kapa Biosystems). In short, 2.0 micrograms of genomic DNA sheared on a Covaris™ E210 was end-repaired, adenylated and ligated to Illumina specific indexed paired-end adaptors. The DNA was size selected with AMPure XP beads (Agencourt, Beckman Coulter) in order to reach the fragment insert size of 220-550bp. The final libraries were quantified using Library Quantification Kit (Kapa Biosystems).

Library preparation 2. Low input gDNA samples aim to be sequenced at low to medium depth. Low input gDNA was used for short-insert paired-end libraries for whole genome sequencing were prepared with KAPA HyperPrep kit (Roche-Kapa Biosystems) with some modifications. In short, 0.8-1.0 microgram of genomic DNA was sheared on a Covaris™ LE220 (Covaris) in order to reach the fragment size of ~500bp. The fragmented DNA was further size-selected for the fragment size of 220-550bp with AMPure XP beads (Agencourt, Beckman Coulter). The size selected genomic DNA fragments were end-repaired, adenylated and ligated to Illumina sequencing compatible indexed paired-end adaptors (NEXTflex® DNA Barcodes). The adaptor-modified end library was size selected and purified with AMPure XP beads to eliminate any not ligated adapters. The final libraries were quantified using Library Quantification Kit (Kapa Biosystems).

Library preparation 3. Good quality samples aim to be sequenced at medium to high depth. Library preparations were performed according to TruSeq DNA PCR-Free library prep guide (Illumina). Briefly, one microgram of quantified genomic DNA was sheared using an LE220Focused-ultrasonicator (Covaris, Inc.) with a Duty factor of 15%, Peak incident power 450 W, 200 cycles per burst for 50 seconds. Sheared DNA fragments were end-repaired, size-selected to obtain DNA fragments around 350 bp, and adenylated according to the manufacturer's instructions. After ligating indexing adapters to the ends of the DNA fragments, quality and band size of libraries were assessed using D1000 Screen Tapes (Agilent) on a TapeStation 2200 (Agilent). Library concentration was measured by qPCR using KAPA library Quantification Kit (KAPA Biosystems).

Library preparation 4. Low input gDNA samples aim to be sequenced at medium to high depth. Library preparations were performed according to TruSeq Nano DNA library prep guide (Illumina). Briefly, 100 nanogram of quantified genomic DNA was sheared using an LE220Focused-ultrasonicator (Covaris, Inc.) with a Duty factor of 15%, Peak incident power 450 W, 200 cycles per burst for 50 seconds. Sheared DNA fragments were end-repaired, size-selected to obtain DNA fragments around 350 bp, and adenylated according to the manufacturer's instructions. After ligating indexing adapters to the ends of the DNA fragments, DNA libraries were enriched using eight cycles of PCR according to the manufacturer's instructions. Quality and band size of libraries were assessed using D1000 Screen Tapes (Agilent) on a TapeStation 2200 (Agilent) after size selection and after PCR amplification. Libraries were quantified by qPCR using KAPA library Quantification Kit (KAPA Biosystems).

Clustering and Sequencing

HiSeq2000 v3, 2x101bp. Some libraries were sequenced using TruSeq SBS Kit v3-HS (Illumina Inc.), in paired end mode, 2x101bp, in a fraction of a sequencing lane of HiSeq2000 flowcell v3 (Illumina Inc.) according to standard Illumina operation procedures. Primary data analysis, the image analysis, base calling and quality scoring of the run, were processed following standard Illumina procedures.

HiSeq2000 v4, 2x126bp. Other libraries were sequenced on HiSeq2000 (Illumina, Inc) in paired-end mode with a read length of 2x126bp using TruSeq SBS Kit v4 following the manufacturer's protocol. Image analysis, base calling and quality scoring of the run were processed following standard Illumina procedures.

HiSeq X-10, 2x150bp. Libraries aim to be sequenced at medium to high depth were sequenced on HiSeq X-10 sequencer (Illumina, Inc) in paired-end mode with a read length of 2x150bp following the manufacturer's protocol. Image analysis, base calling and quality scoring of the run were processed following standard Illumina procedures.

*Library preparation 1 and 2 and Sequencing using HiSeq2000 was carried out in CNAG facilities. Library preparation 3 and 4 and Sequencing using HiSeqX-10 was carried out in MACROGEN facilities. In all cases, samples were sequenced using Illumina protocols, and primary data analysis was carried out with the standard Illumina pipeline.

Quality controls, trimming & mapping to nuclear genome

We evaluated the quality of the data using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and removed adaptors using SeqPrep (https://github.com/jstjohn/SeqPrep). Trimmed reads were mapped to a 2.4Gb *Lynx pardinus* (Iberian lynx) reference genome (Abascal et al., 2016) using BWA-MEM (Li, 2013) with default parameters. After adding read groups to each sample using picard-tools (https://broadinstitute.github.io/picard/), we merged the bam files from the same individuals, which are also coming from the same library, with SAMtools merge (Li et al., 2009). We marked the duplicates using picard-tools (https://broadinstitute.github.io/picard/) and performed a local realignment using GATK 3.4 (McKenna et al., 2010). Then, we performed a Base Quality Score Recalibration (BQSC), a pre-processing step to detect systematic errors made by the sequencer. As BQSC needs a set of known variants, which we did not have, we followed GATK recommendations (https://gatkforums.broadinstitute.org/gatk/discussion/44/base-quality-score-recalibration-bqsr), and did an initial round of SNP calling on our data unrecalibrated, and used it afterwards to do BQSR. Variant calling was done using HaplotypeCaller (McKenna et al., 2010) on a priori defined populations (considered as sampling locations, see Table S1, Table S2). We filtered the resulting VCF files (https://github.com/vcflib/vcflib, "QD > 2 & MQ > 40 & FS < 100"), and used those VCF file only to perform BQSC of the data using GATK 3.4 (McKenna et al., 2010). We calculated overall stats using SAMtools flagstat (Li et al., 2009) and average depth using SAMtools depth (Li et al., 2009). The depth of our samples was on average 7.4 +/- 2.3 sd for those sequenced using HiSeq 2000 and 23.7+/- 3.7 sd for those sequenced with HiSeq X-10.

Neutral regions definition and exclusion

To determine neutral regions of the nuclear genome we excluded all the regions annotated as functional categories in the Iberian lynx reference genome, i.e. genes, lncRNAs, mRNAs, and ncRNAs (Abascal et al., 2016). Besides, we annotated and excluded promoters of protein-coding genes and lncRNA defined as 1000-bp upstream of the gene or lncRNA. We annotated and excluded ultra conserved non-coding elements (UCNEs) using human coordinates (https://ccg.vital-it.ch/UCNEbase/). This information was translated into cat coordinates using LiftOver (http://rohsdb.cmb.usc.edu/GBshape/cgi-bin/hgLiftOver), and latter, we translated these cat

coordinates into lynx coordinates by using lynx to cat sinteny (Abascal, 2016). We also added a security buffer of +/-1000bp to any functional region to avoid regions under the influence of functional areas. To exclude these functional regions and our security buffer from the bam files we used bedtools subtract and bedtools intersect (Quinlan & Hall, 2010).

Mitogenome reconstruction

To reconstruct the mitogenome, we mapped only 10M random reads in the case of WG samples while all reads coming from the capture experiment were used. The average depth was calculated using Samtools depth piped to a custom bash script.

We called SNPs using the following settings on Freebayes (Garrison & Marth, 2012): haploid genome, minimum read mapping quality 30, minimum base quality 20, a minimum of two bases supporting an alternative allele call, and a maximum read mismatch fraction of 0.2. Repetitive regions RS2 and RS3 (positions 16096-16382 and 16908-174, respectively) (Sindičić et al., 2012), where no SNPs could be called reliably, were excluded from the analysis. The number of reference positions not covered by any read was calculated for each individual using the command genomecov in BEDTools (Quinlan & Hall, 2010). A consensus for each mitochondrial genome was constructed using the FastaAlternateReferenceMaker command in GATK (McKenna et al., 2010). Consensus sequences were rearranged to present them with standard coordinates but lacking the excluded repetitive regions.

Treatment of suboptimal samples

Balkans samples were museum material with unknown collection date. Some evidence indicates that at least two of them were more than 70 years old. All three samples were maintained in suboptimal conditions for DNA studies. Accordingly, we processed the samples in a sterile lab especially used for museum or suboptimal material. DNA extraction yielded relatively low amount of DNA and the DNA was highly degraded. Still, given the interest of the samples, we proceeded with library preparation. Libraries failed standard quality control because of degradation and RNA contamination. One of them also failed in DNA quantity. Standard procedures were implemented to prevent contamination in the lab throughout the whole process. As they were suboptimal material they underwent two sequencing runs to reach our

goal depth, in contrast with the rest of the samples that got this depth with only one run. FASTQC analysis indicates that the reads show a distorted GC, which could be an indication of bacterial DNA. We mapped the reads to the nuclear genome, and calculated the endogenous content as: reads mapped vs. total reads. Balkans samples showed 67%, 73%, and 85%, way below the average for the rest of the samples (~97%). We also calculate Ts/Tv ratio. High Ts/Tv ratios are indication of damaged DNA, and are typical of museum and ancient materials. While one of the samples (labeled as historical) showed signs of damage when compared to the rest of contemporary samples, the other two showed a very low value. We performed diversity measures using ANGSD and SNP calling using GATK and we found unexpectedly high diversity values (above twice the diversity of Tuva (the population with the highest diversity). Therefore, we decide to exclude these samples from this analysis. Still, we think that analysis such as structure or PCA, although taken with caution, could give some useful information on the relationship of Balkan samples to the rest of the samples. Contamination, errors, and damage are unlikely to be shared among samples, so the variants shared by all three samples and separating Balkan samples from the rest are likely to be authentic.

For the mitogenome analysis, Balkans samples underwent a BLAST+ (Camacho et al., 2009) based filtering step between mapping and SNP calling. BLAST+ searches were performed in our own server against NCBI database using as queries all mapping reads showing one or more mismatches to the reference genome. Reads were excluded if the list of hits with the lowest E-values did not include at least one felid species. Besides, we manually revised and confirmed all SNPs called from these samples. Therefore, the extra divergence of the Balkan mitogenomes, obtained from low quality samples, is not likely caused by contamination or post-mortem damage, as many divergent reads potentially coming from exogenous DNA were efficiently removed through this BLAST filtering step (sample:percentage of removed reads, mapping reads retained post-filtering: h_ll_ba_0214: 34%, 18370; h_ll_ba_0215: 17%, 26622; c_ll_ba_0216: 8%, 30749). Moreover, the exact same haplotype was independently reconstructed from the three samples; suggesting that they are not affected by contamination or damage, which are expected to be random and sample-specific.

Identification of the ancestral state

We inferred the ancestral state using *L. rufus* for our analysis with ANGSD. Raw paired-end Illumina reads were mapped to the Iberian lynx reference genome and the resulting bam file was used to call variants using SAMtools mpileup (–q30) (Li et al., 2009). The output was piped into pu2fa (-C45) program from Chrom-Compare-master (https://github.com/Paleogenomics/Chrom-Compare) to do a pseudo-haploidisation. We inferred the ancestral state for 97% of bases in our reference genome.

Chromosome X and Y regions definition and molecular sexing

The annotation of the reference genome do not include sexual chromosomes so we characterized such regions as they should be treated apart of the autosomes because of their different ploidy. For each sexual chromosome we defined two different sets of regions depending on their posterior use:

i) A lax definition was used to define any region having signals of belonging to a sex chromosome. We used these regions to sex molecularly our individuals and to exclude them from any analysis intended on the autosomes (i.e. PSMC, measures of autosomal diversity).

ii) An strict definition was used to define regions which showed clear signs of belonging to the sexual chromosomes and therefore were used in the analysis aiming to study haploid and sex-linked genetic signals.

i) For our lax definition, we defined non-recombining parts of sexual chromosomes regions using differential depth expected in autosomal vs Y and non-recombinant X chromosome regions. Expecting that female/male ratio would be: ~1 in autosomal regions ~0 in Y non-recombinant regions ~0.5 in X non-recombinant regions. Similar approaches were previously used by Hall et al. (2013), and later modified by Smeds et al. (2015). We used *Lynx pardinus* individuals (eleven males and 15 females) sequenced to 5x depth to do this analysis. First because the sex of those individuals was known, and also, because our reference genome belongs to *Lynx pardinus* and only perfectly matching reads are taken into account to calculate read depth -species

divergence could cause lack of depth in some areas of the genome. Then, we extrapolated our results to our *L. lynx* individuals, assuming that the divergence of the two species is so recent that chromosomal reorganizations, if any, should be a minority. We computed the per base average normalized depth for both females and males and then calculated the female/male ratio.

First, we used samtools view to select only primary alignments with a minimum mapping quality of 30 and no mismatches followed by samtools depth to get each base depth. Then, for each individual we used its mode to normalize each base depth to later compute the average depth per base and sex which was used to get the per base female normalized depth / male normalized depth ratio. The distribution of such ratios was explored in R using a random subset of 5M positions.

We decided to make the definition (X, Y or autosomic) of each region using contigs as the meaningful units, as single bases belonging to a chromosome when the adjacent ones do not does not make biological sense and scaffolds might be chimerical due to errors in the scaffolding process. For each contig we obtained its length as well as the number of bases with no data available.

To further inform the definition of the X chromosome, we took advantage of the definition of the regions annotated as syntenic to cat's X or autosomal chromosomes in Abascal (2016). First, we compared the distributions of the ratios found in those syntenic regions in contrast with regions syntenic to autosomes to help us to define the thresholds used. A base was assigned to the X chromosome if it showed a female depth/male depth ratio above 1.5 and/or if it had been described as syntenic to the cat X chromosome. We defined a contig as belonging to the X chromosome in our lax definition if more than 30% of its bases met our requirements.

Chromosome\Definition	Lax (scaffolds/bases)	Strict (scaffolds/bases)
Y Chromosome	401 / 335,035	77 / 33,032
X Chromosome	14,467 / 72,426,678	1,891 / 45,245,057

The lax definition of the Y chromosome contigs required them to have more than 30% of their bases with a female/male depth ratio below 0.3 and at least a normalized depth of 0.2 in males (to avoid non-covered regions by sequencing or mapping biases).

ii) As the strict X chromosome definition we simply used the bases that had already been described as X chromosome using the synteny to the cat genome.

As the strict definition, we required at least 90% of the total bases in a contig to have a ratio below 0.3 and an average normalized depth for males between 0.2 and 0.8 (to also exclude multicopy Y chromosome regions) for the Y chromosome.

To sex molecularly our individuals, we called SNPs on each sample's original BAM file using the lax definition of the Y chromosome as target region using GATK 3.4 (McKenna et al., 2010) and following GATK best practice recommendations (DePristo et al., 2011). Then we visually inspected the VCF files to define as females those samples showing virtually no depth for any reported SNP compared to males, which showed consistent depth in all regions.

Y chromosome variation

We called SNPs on males exclusively using their pooled BAM files, selecting our strict definition of Y chromosome bases as the target region, and the same setting used for the mitogenome calling.

SNP calling and genotyping

When required for the analysis we called variants following the genome VCF (GVCF) workflow in GATK 3.4 (McKenna et al., 2010)), i.e. we ran the HaplotypeCaller tool individually on each *L. lynx* sample to produce per-sample intermediate GVCFs, which were parsed to the GenotypeGVCFs tool for a joint variant calling and genotyping. Next, INDELs and multiallelic SNPs were filtered out from the VCF, as were all other variants that did not meet the recommended GATK hard filters (variants were excluded if "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0 || SOR > 3.0").

Phylogenetic analyses of mitogenomes

In order to estimate a substitution rate for the mitochondrial genome in felids, we downloaded from GenBank the sequences and the respective annotations of the following felid mitogenomes: *Panthera leo* (GenBank: KP202262), *P. pardus* (GenBank: KP202265), *P. onca* (GenBank: KP202264), *P. uncia* (GenBank: KP202269), *P. tigris* (GenBank: KP202268), *Neofelis nebulosa* (GenBank: KP202291), *Acinonyx jubatus* (GenBank: KP202271), *Puma concolor* (GenBank: KP202261), *Catopuma temminckii* (GenBank: KP202267), *L. lynx* (GenBank: KP202283), *Pardofelis marmorata* (GenBank: KP202263), *Felis catus* (GenBank: FCU20753), *Prionailurus bengalensis* (GenBank: KP246843), *Caracal caracal* (GenBank: KP202272), and *Leopardus guigna* (GenBank: KP202293), and *Prionodon pardicolor* (GenBank: NC_024569) as a non-felid outgroup. Sequences were aligned using ClustalW v2.1 in Geneious v11.1.2 (Kearse et al., 2012). We used the Greedy algorithm with linked branch lengths implemented in PartitionFinder v2.1.1 (Lanfear et al., 2017) to investigate the set of partitions and substitution models adequate to all possible combinations of tRNAs (pulled together), the two rRNAs and twelve genes in the dataset of the 16 mitochondrial sequences, encompassing 14,994 bp. The ND6 gene and the hypervariable Control Region were excluded from the analyses. We inferred a substitution rate of 1,54*10-8 mutations/site/year for the mitogenome.

We performed a Bayesian phylogenetic analysis with the 15 feline species using BEAST v2.4.8 (Bouckaert et al., 2014) and the optimal partitions selected with PartitionFinder among the models implemented in BEAST. The tree was calibrated using a set of fossil constraints (Table below). We assumed a strict clock as justified before (Paijmans et al., 2017) and a calibrated Yule model as the tree prior, more appropriate when considering sequences from different species. The Markov chain Monte Carlo (MCMC) chain was run for 50 million steps to ensure convergence of the chain and ESS values well above 200 for all parameters estimated. Results were visualised with Tracer v1.7 (Beta version) (Rambaut et al., 2018) discarding the initial 10% as burn-in.

			Log Normal Parameters	
Fossil (estimated age)	Minimum hard bound	Maximum soft bound	M	S
Lynx (5.3 Ma)	5.3	10	1.0	1.25
Acinonyx (3.8 Ma)	3.8	10	1.2	1.3
Caracal (3.8 Ma)	3.8	16	2.0	1.5
Oldest Panthera tigris (1.5 Ma)	1.5	10	1.5	1.35
Oldest Panthera (3.8 Ma)	3.8	16	2.0	1.5

SUPPLEMENTARY FIGURES

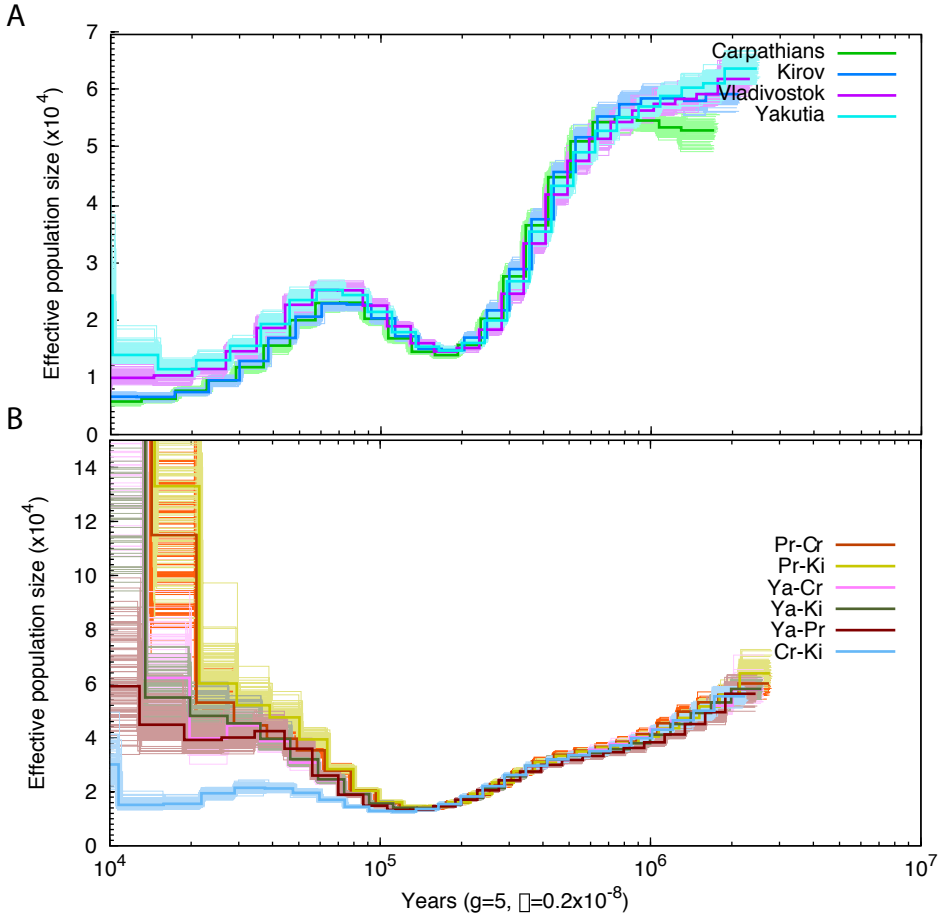


Fig. S1. PSMC results with bootstrapping for individuals from Kirov, Yakutia, Primorsky Krai and Carpathians, and for pseudo-diploid sequences created by combining haplotypes from pairs of populations.

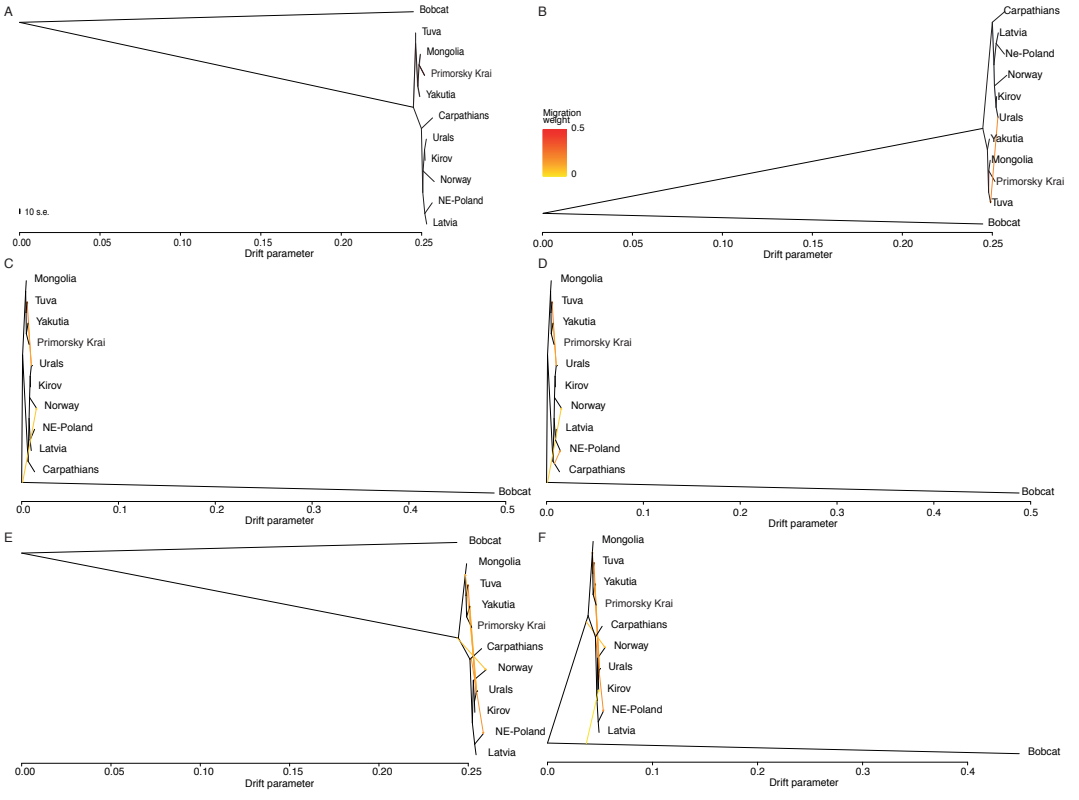


Fig. S2. Population trees inferred by Treemix when allowing zero, one, three, four, five and six migration events.

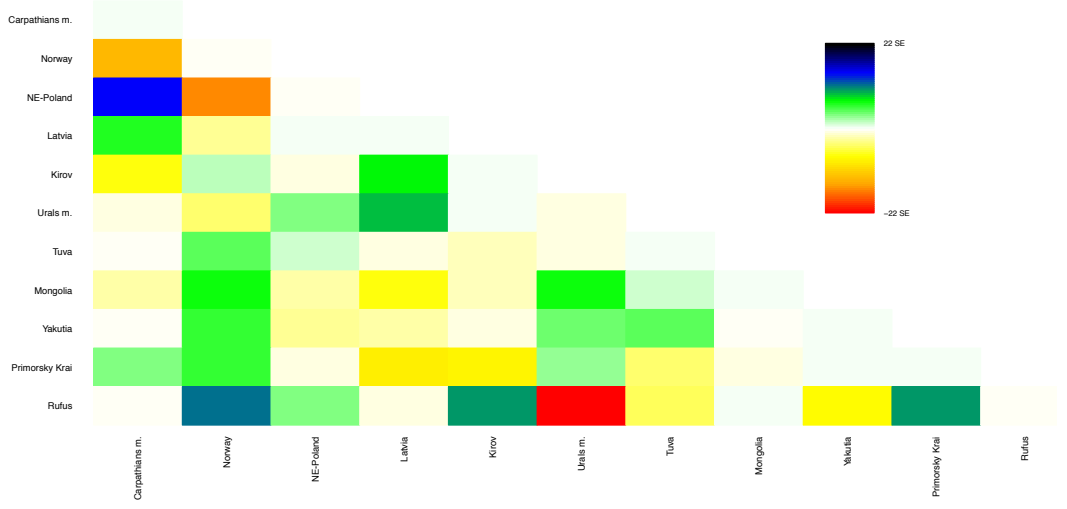


Fig. S3. Visualization of the residuals from the fit of the model to the data for m=2 in TreeMix.

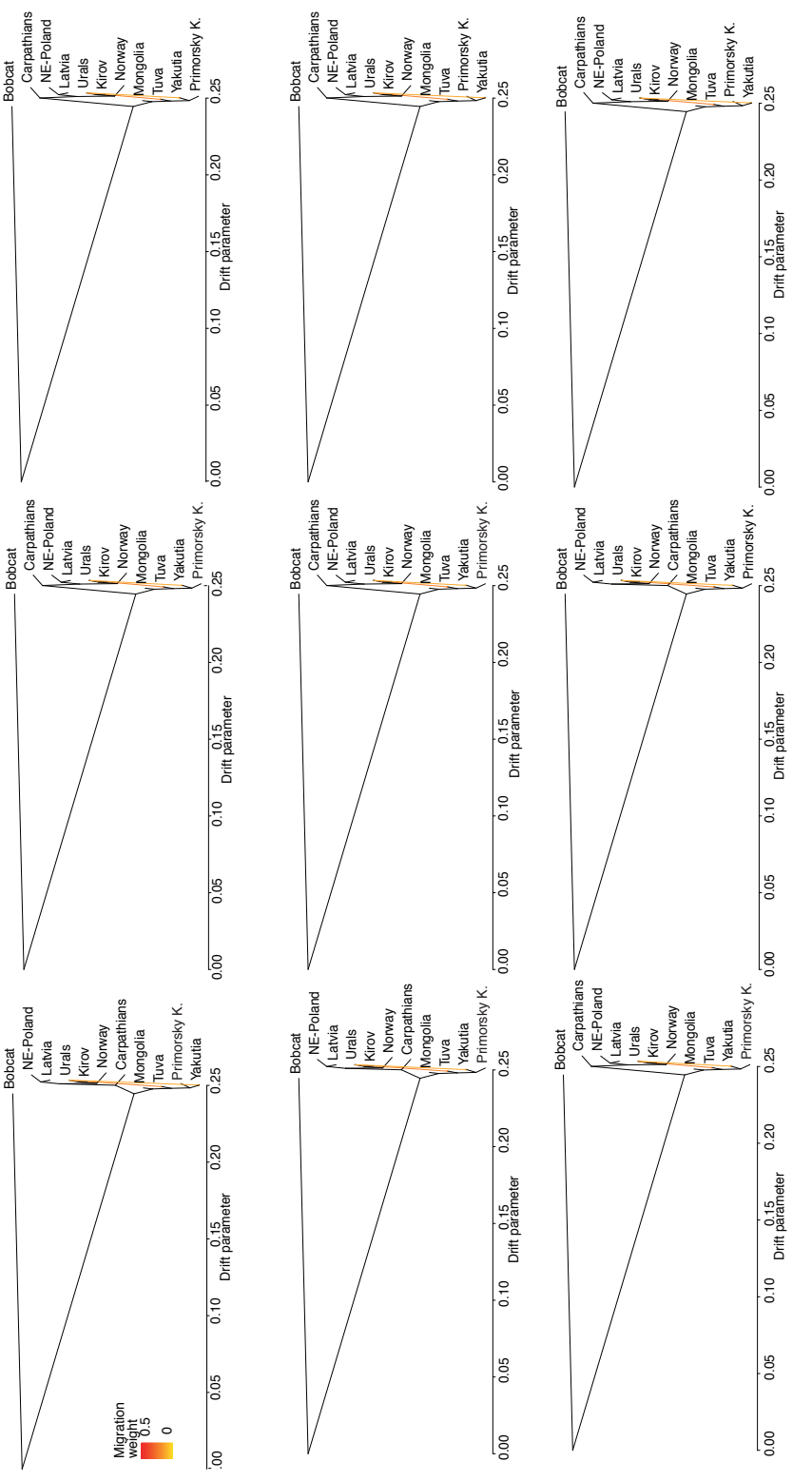


Fig. S4. Treemix results for 10 runs allowing two migration events.

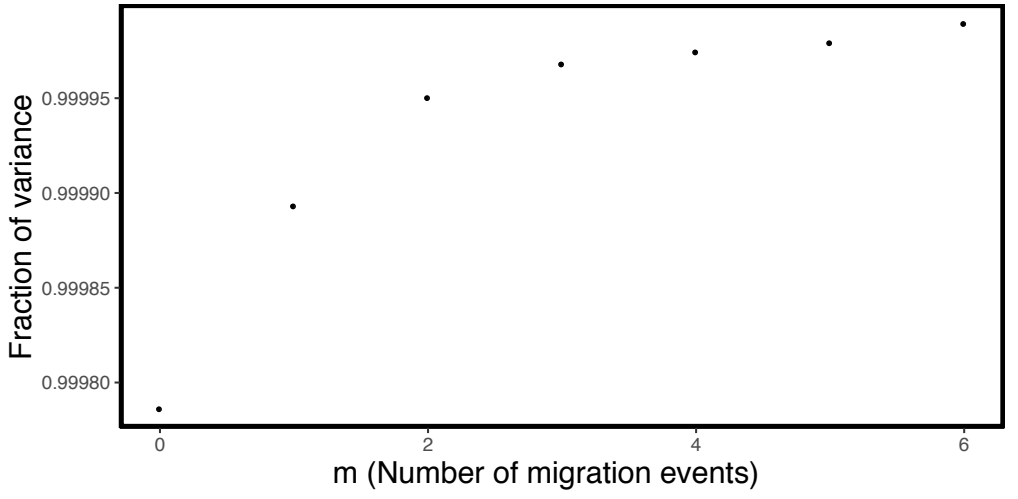


Fig. S5. Fraction of variance explained by the model allowing 0 to 6 migration events in Treemix.

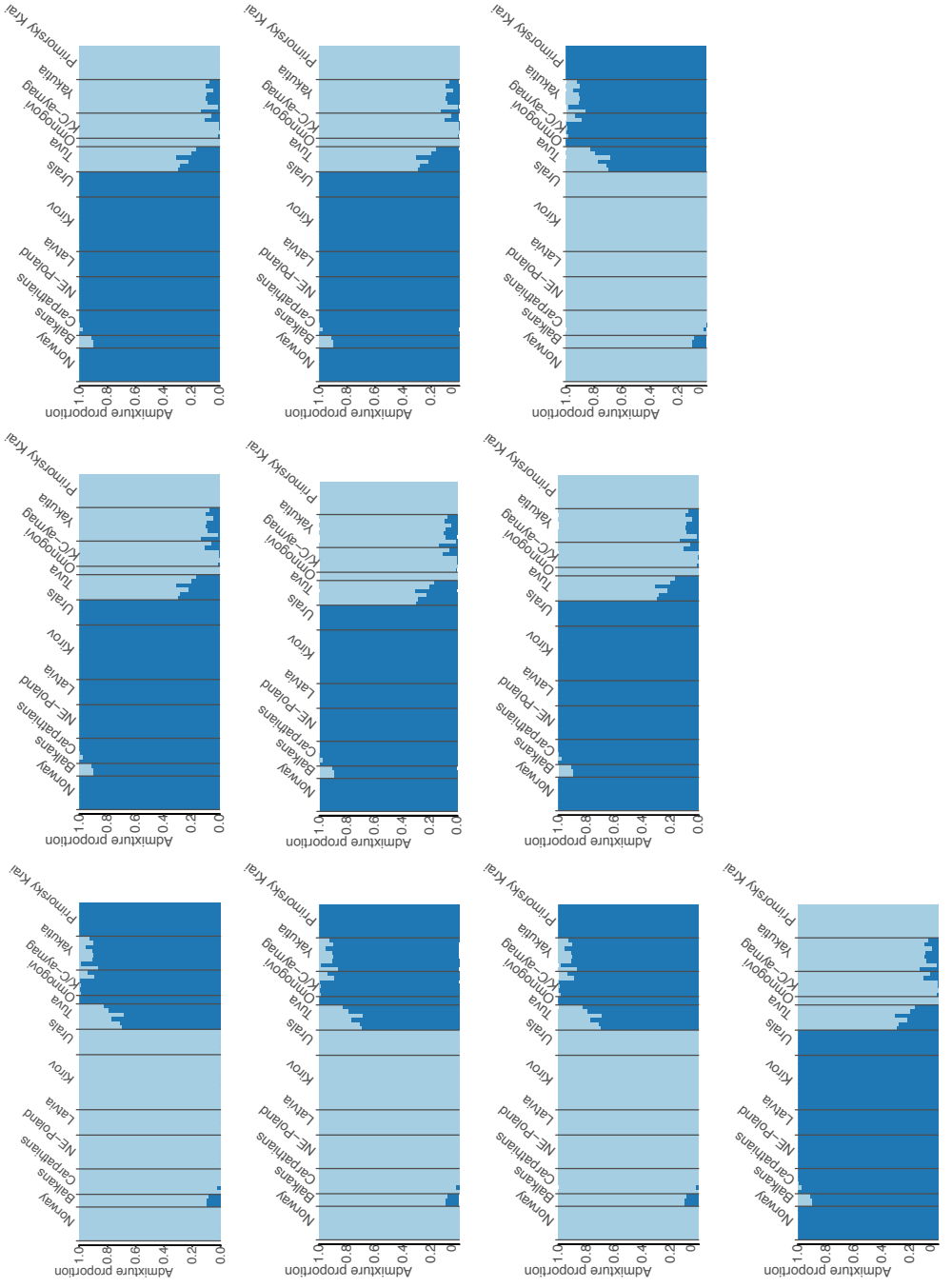


Fig. S6. Alternative clustering patterns obtained in 10 runs of NGSadmix for K=2, Mongolia population comprise K/C-aymag (i.e. Central and Khentii aymag) and Ömnögovi.

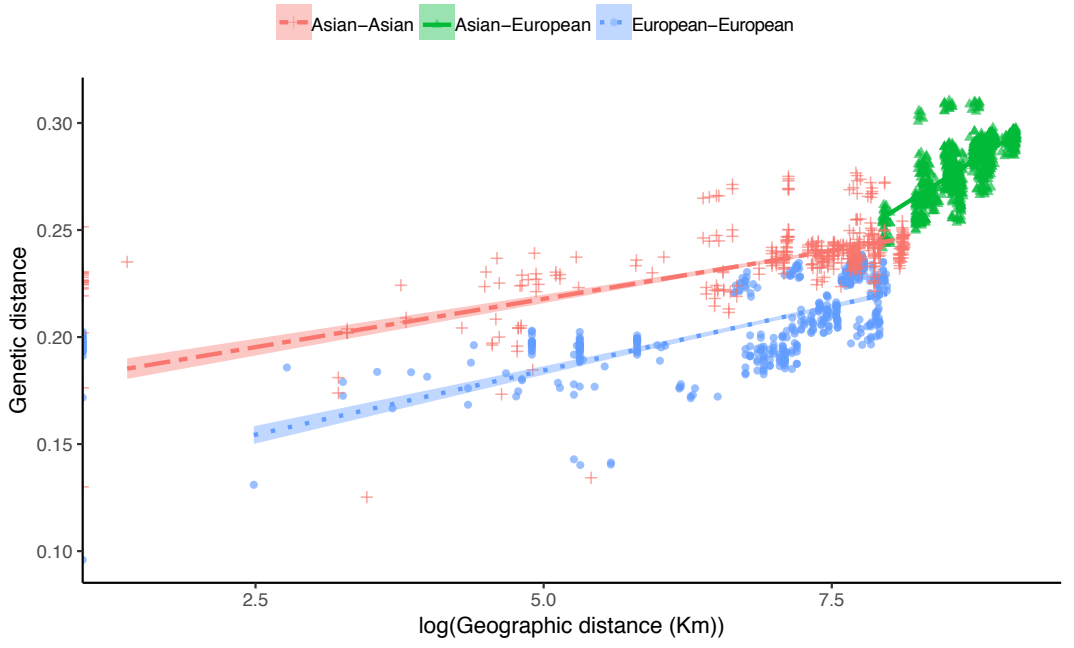


Fig. S7. Correlation of genetic vs. geographic distances between pairs of individuals. Correlations were assessed separately for Asian-Asian, European-European and Asian-European pairs. European bottlenecked populations (Carpathians, Norway and NE-Poland) were not included in this analysis. We tested an effect additional to geographic distance contributing to the genetic distance separating Asian and European individuals through a partial Mantel test. This test was significant (r=0.5459; p=0.001).

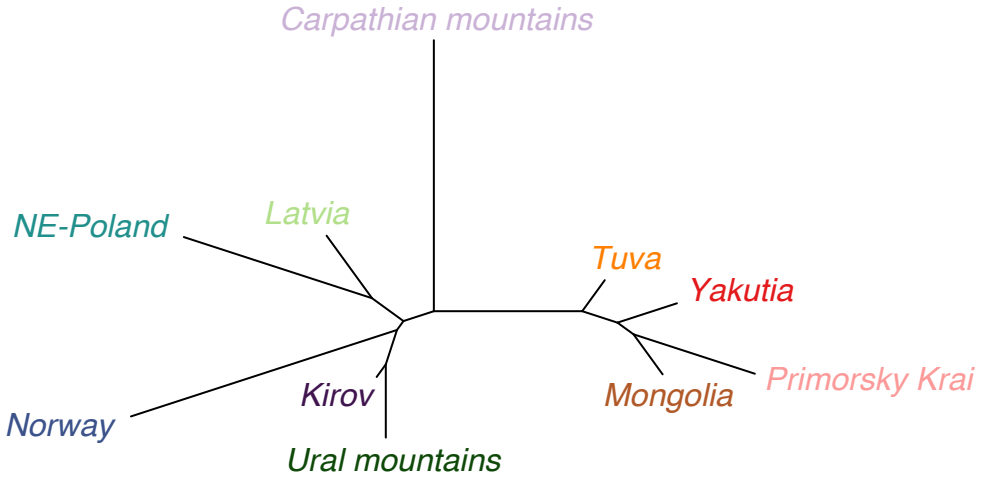


Fig. S8. NJ-tree of populations based on pairwise FST.

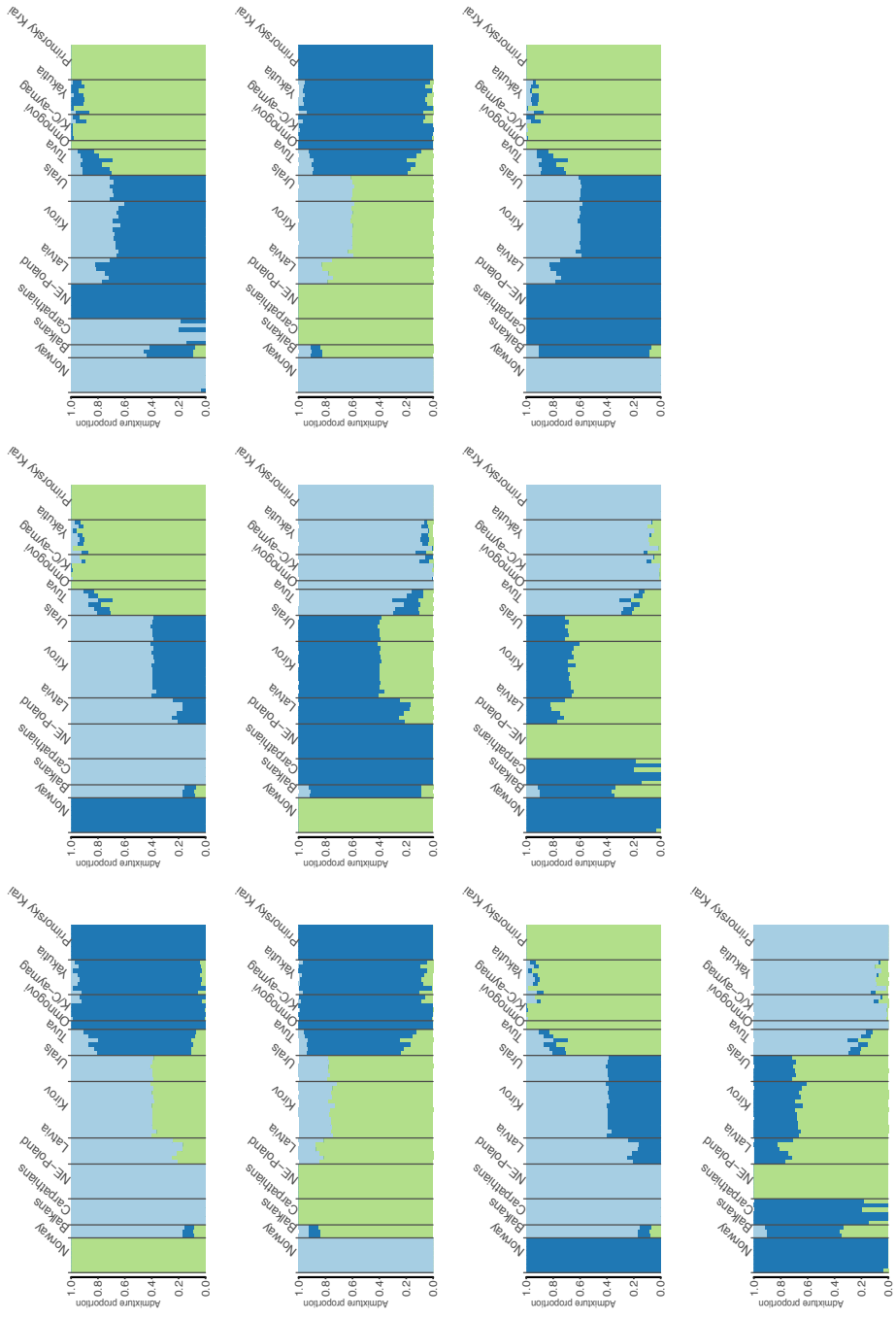


Fig. S9. Alternative clustering patterns obtained in 10 runs of NGSadmix for K=3. Mongolia population comprise K/C-aymag (i.e. Central and Khentii aymag) and Ömnögovi.

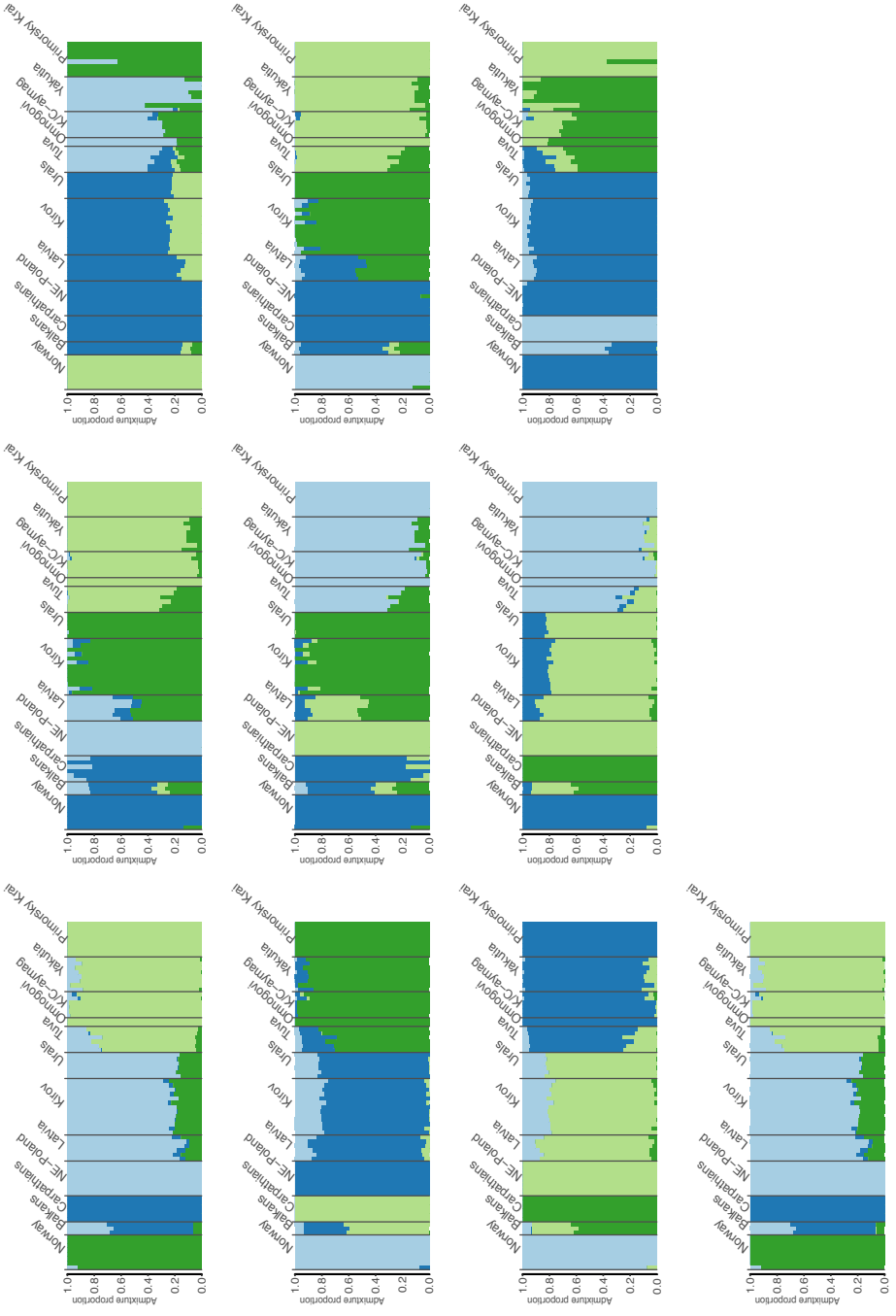


Fig. S10. Alternative clustering patterns obtained in 10 runs of NGSadmix for K=4. Mongolia population comprise K/C-aymag (i.e. Central and Khentii aymag) and Ömnögovi.

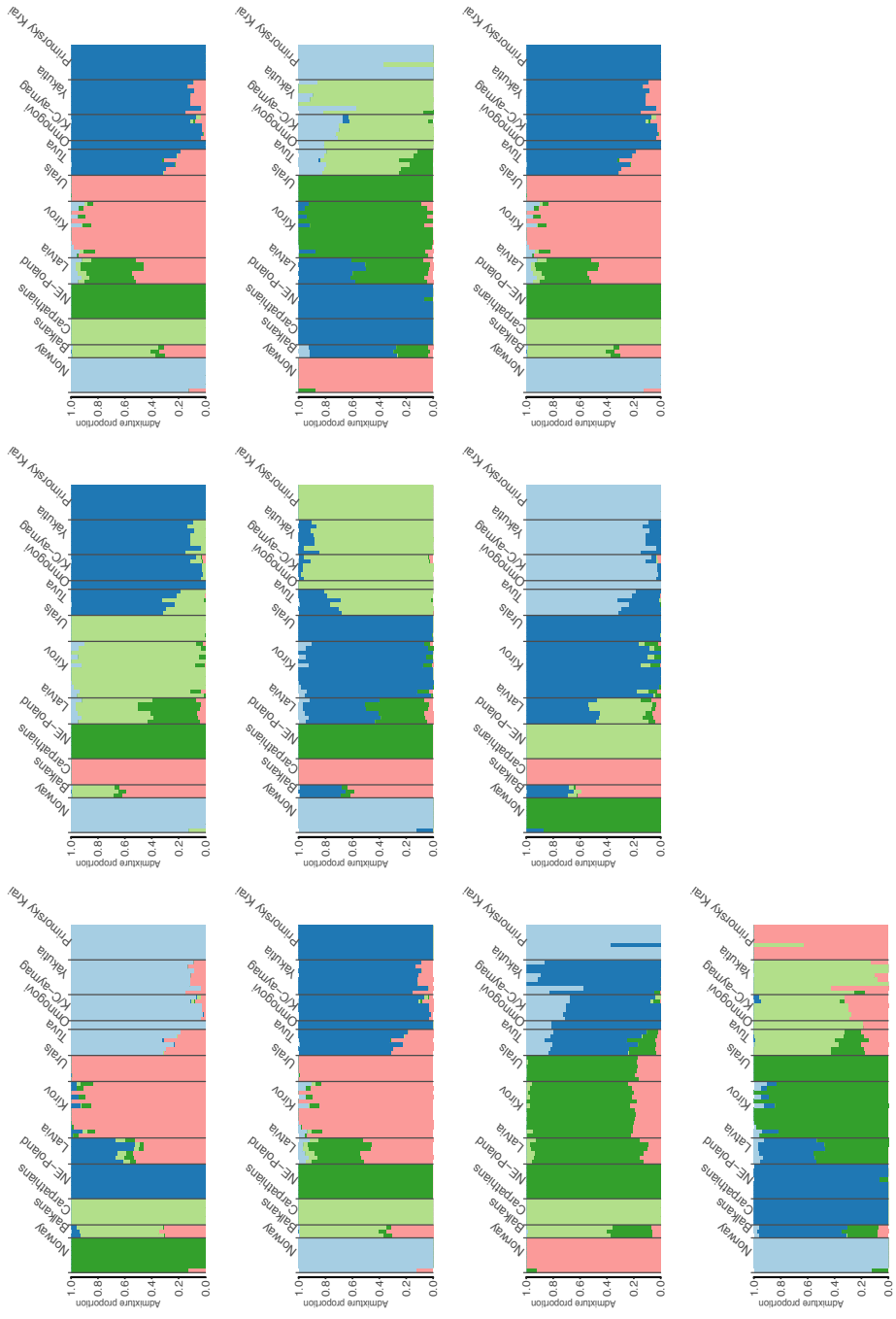


Fig. S11. Alternative clustering patterns obtained in 10 runs of NGSadmix for K=5. Mongolia population comprise K/C-aymag (i.e. Central and Khentii aymag) and Ömnögovi.

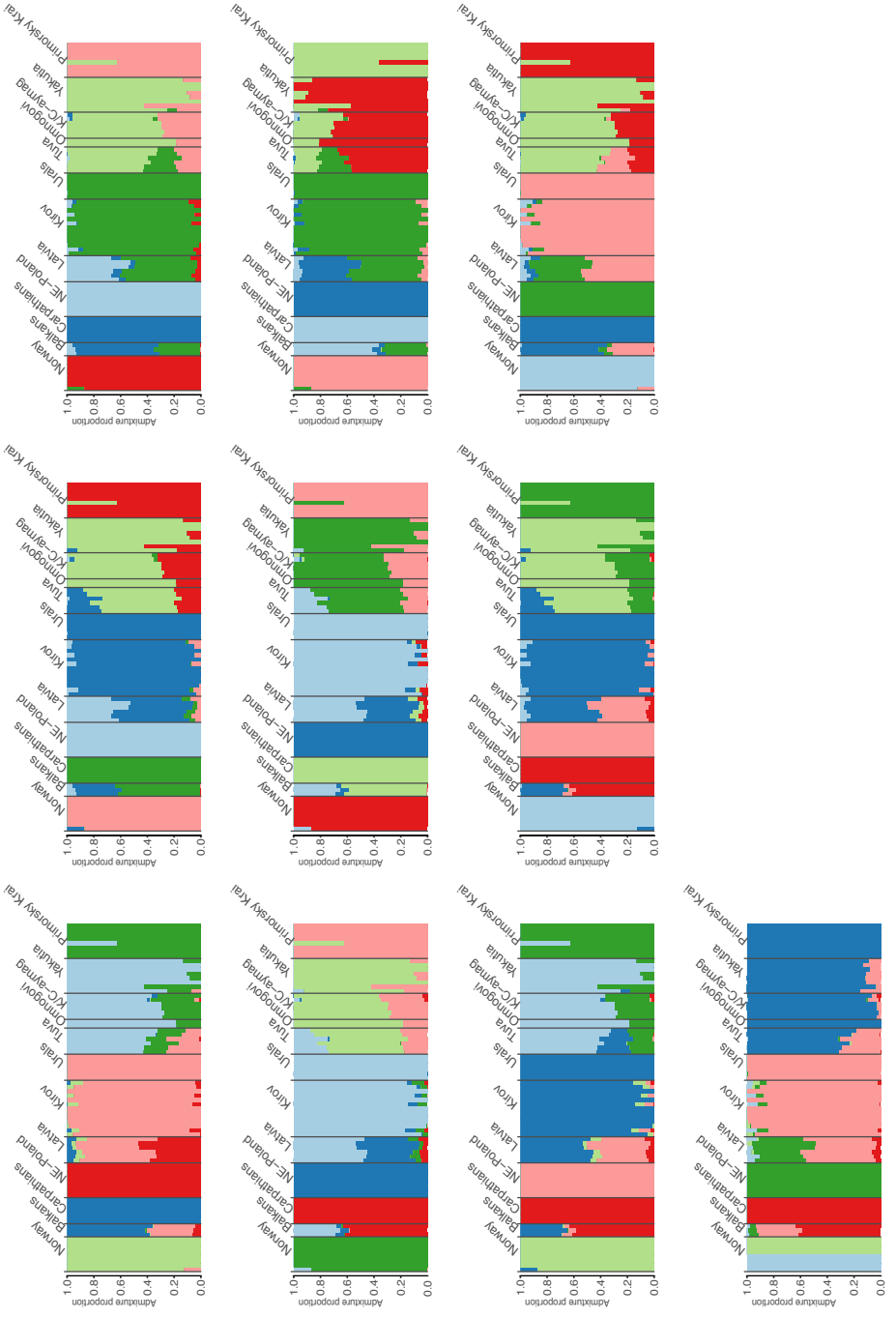


Fig. S12. Alternative clustering patterns obtained in 10 runs of NGSadmix for K=6. Mongolia population comprise K/C-aymag (i.e. Central and Khentii aymag) and Ömnögovi.

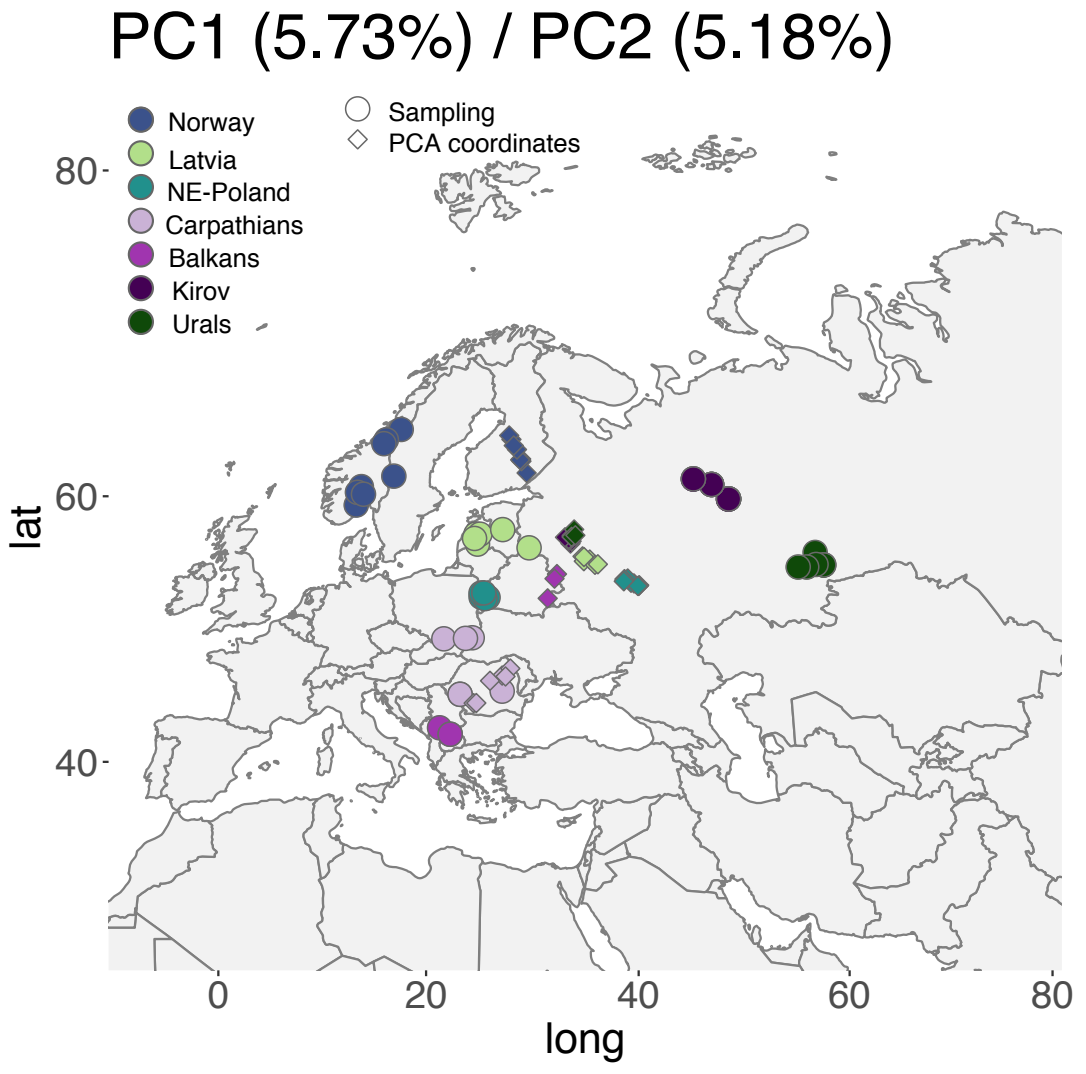


Fig. S13. Procrustes-transformed PCA coordinates to geographical coordinates in European populations, and sampling sites.

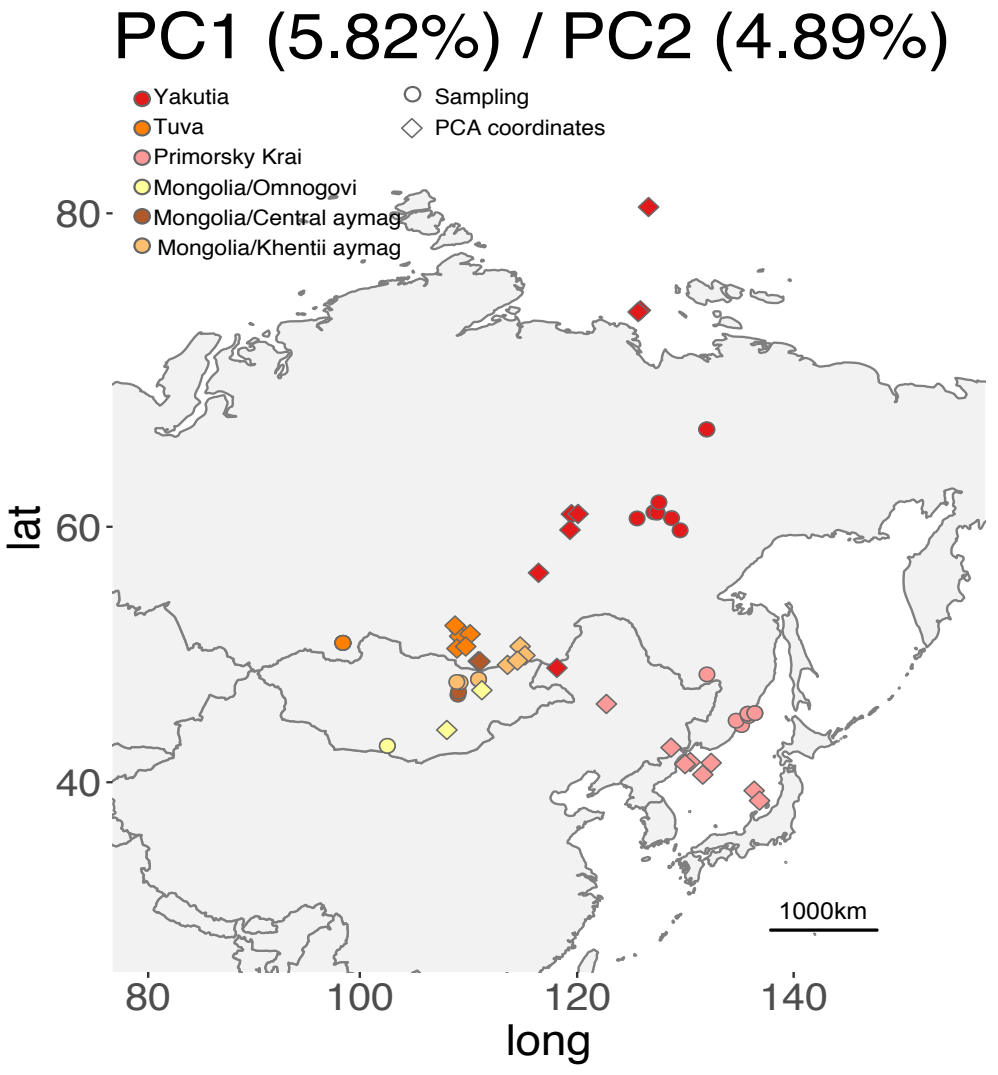


Fig. S14. Procrustes-transformed PCA coordinates to geographical coordinates in Asian populations, and sampling sites.

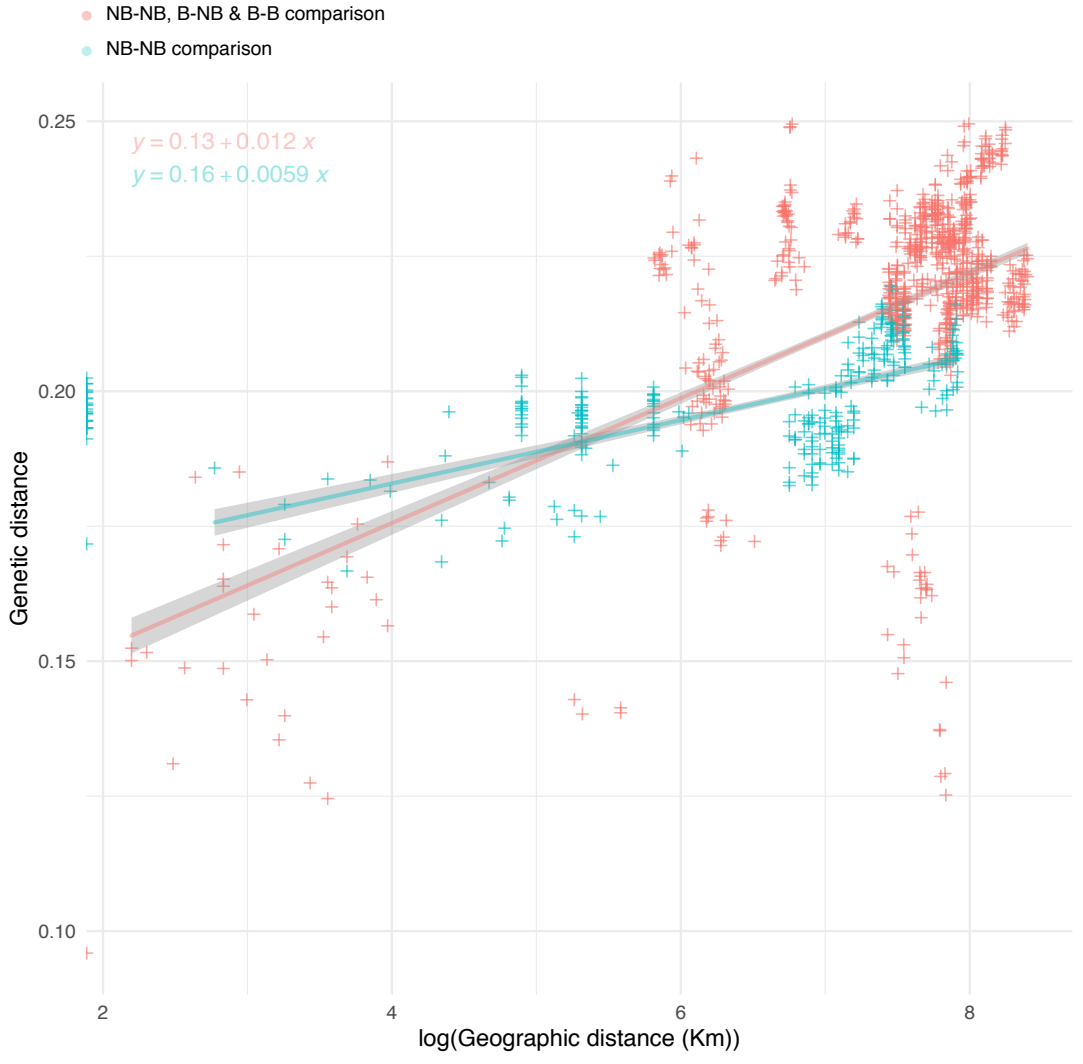


Fig. S15. Correlation of genetic vs. geographic distances between pairs of individuals from European populations, including and excluding individuals from European bottlenecked populations (Carpathians, Norway and NE-Poland). NB-NB: comparisons between two individuals from non-bottlenecked populations (NB); B-NB: comparisons between one individual from bottlenecked (B) and one individual from non-bottlenecked population (NB); B-B: comparison between two individuals from bottlenecked populations (B).

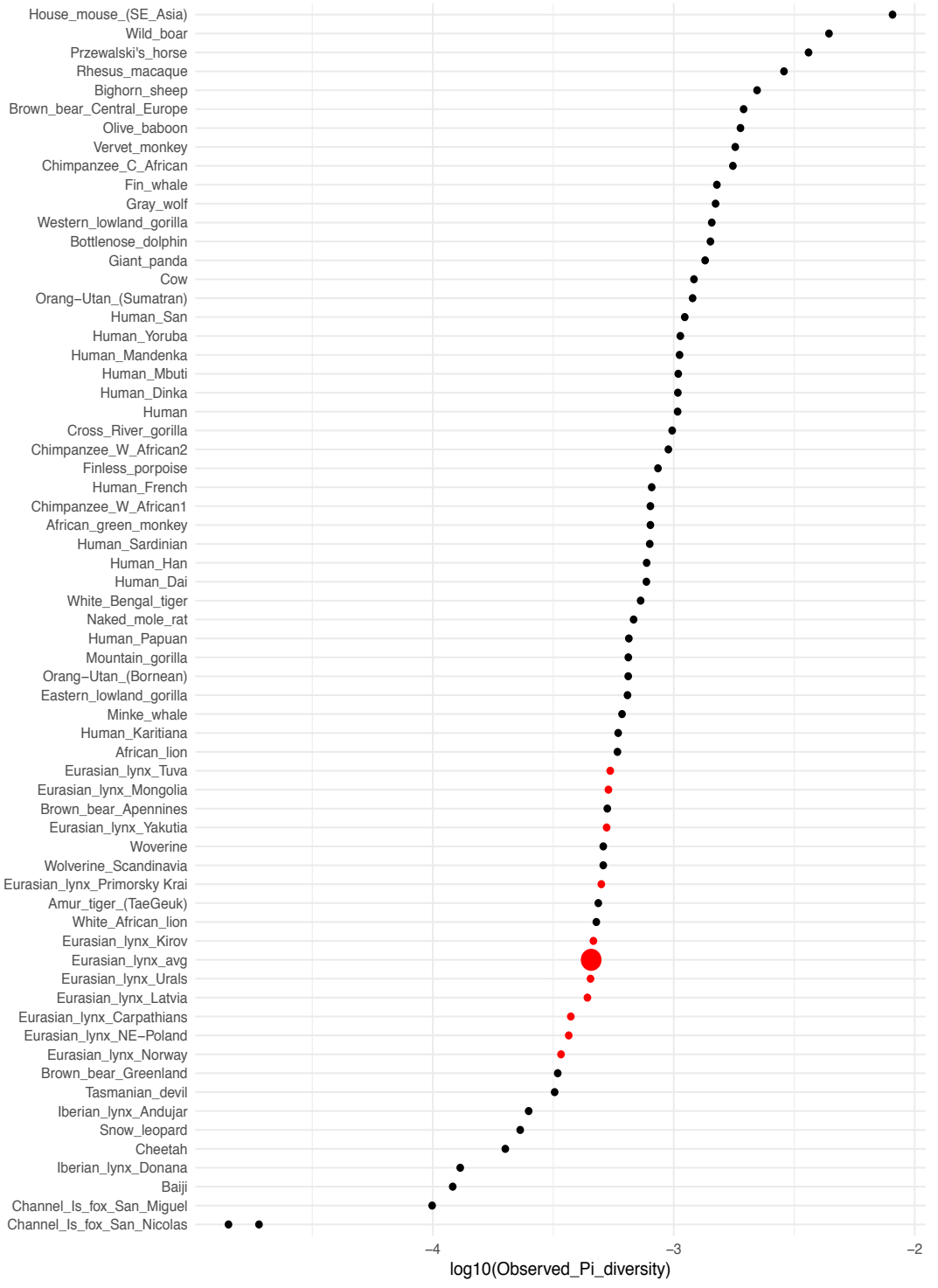


Fig. S16. Comparison of genomic nucleotide diversity among mammals. Modified from Robinson et al. (2016).

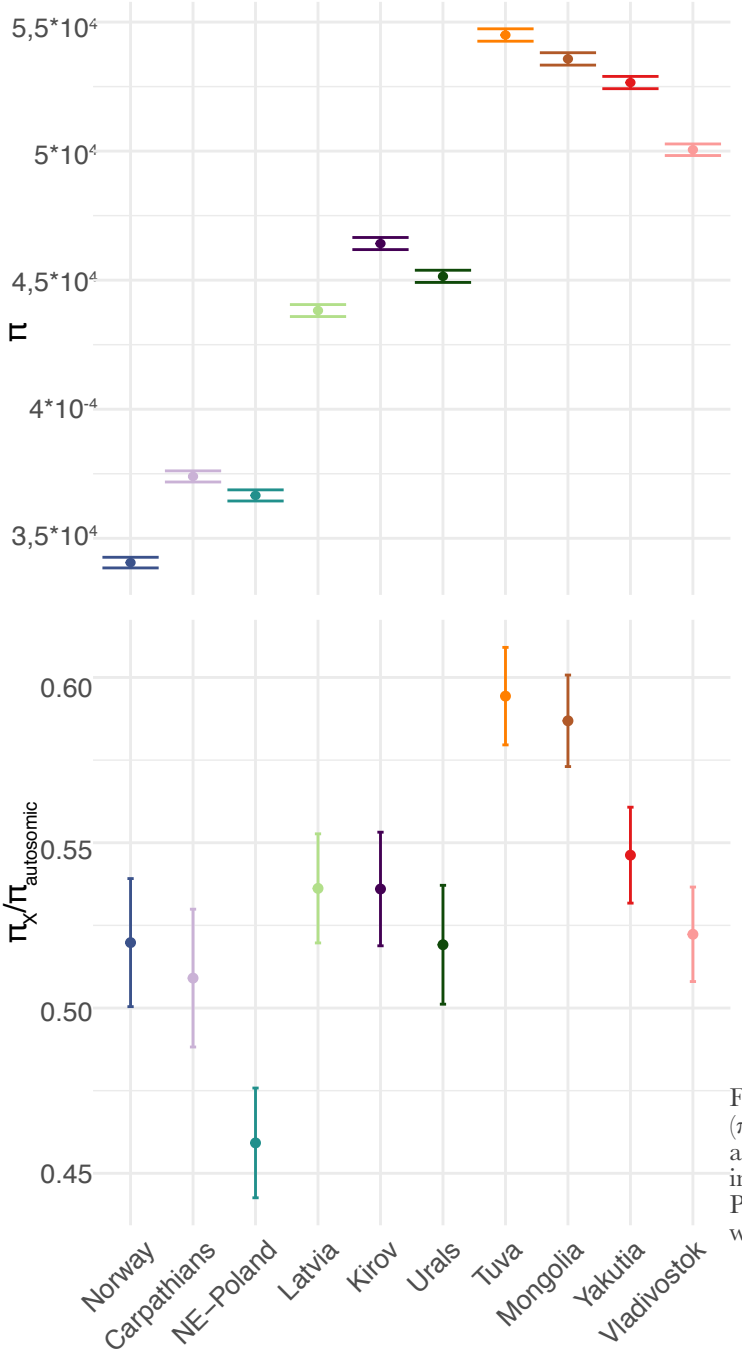


Fig. S17. Nucleotide diversity (π) values for autosomal sites and ratios of π diversity in Xchr vs. autosomes. Populations are sorted from west to east.

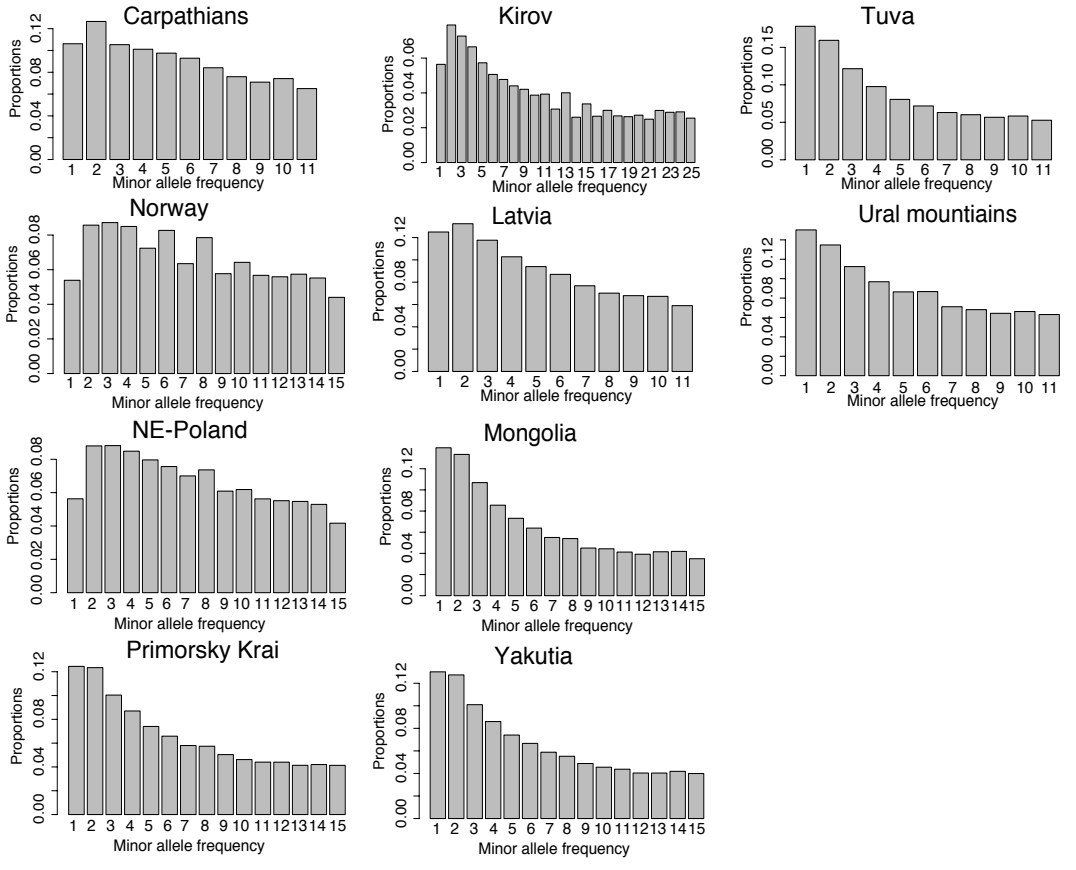


Fig. S18. Site frequency spectrum (SFS) of the different Eurasian lynx populations.

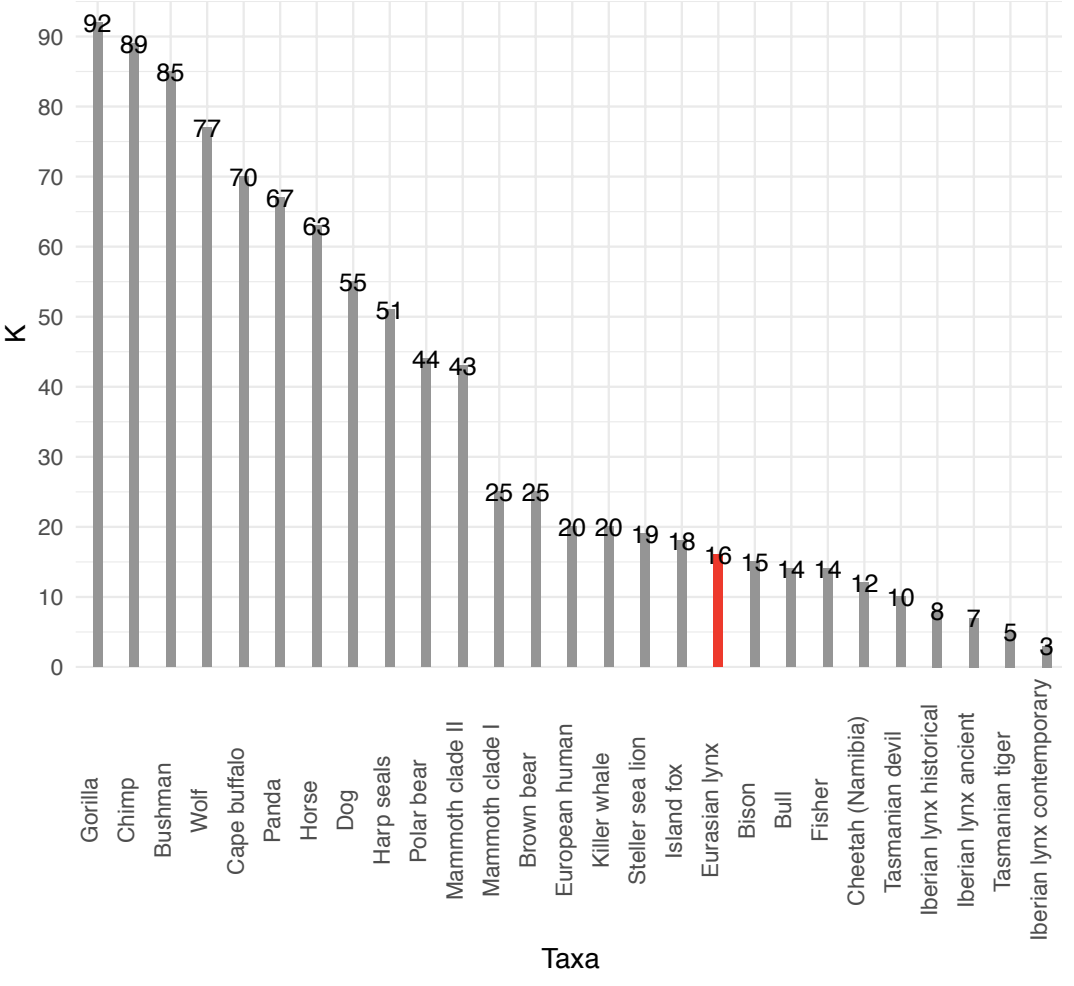


Fig. S19. Comparison of mitogenomic pairwise differences (k) among different species. Modified from (Casas-Marce et al., 2017).

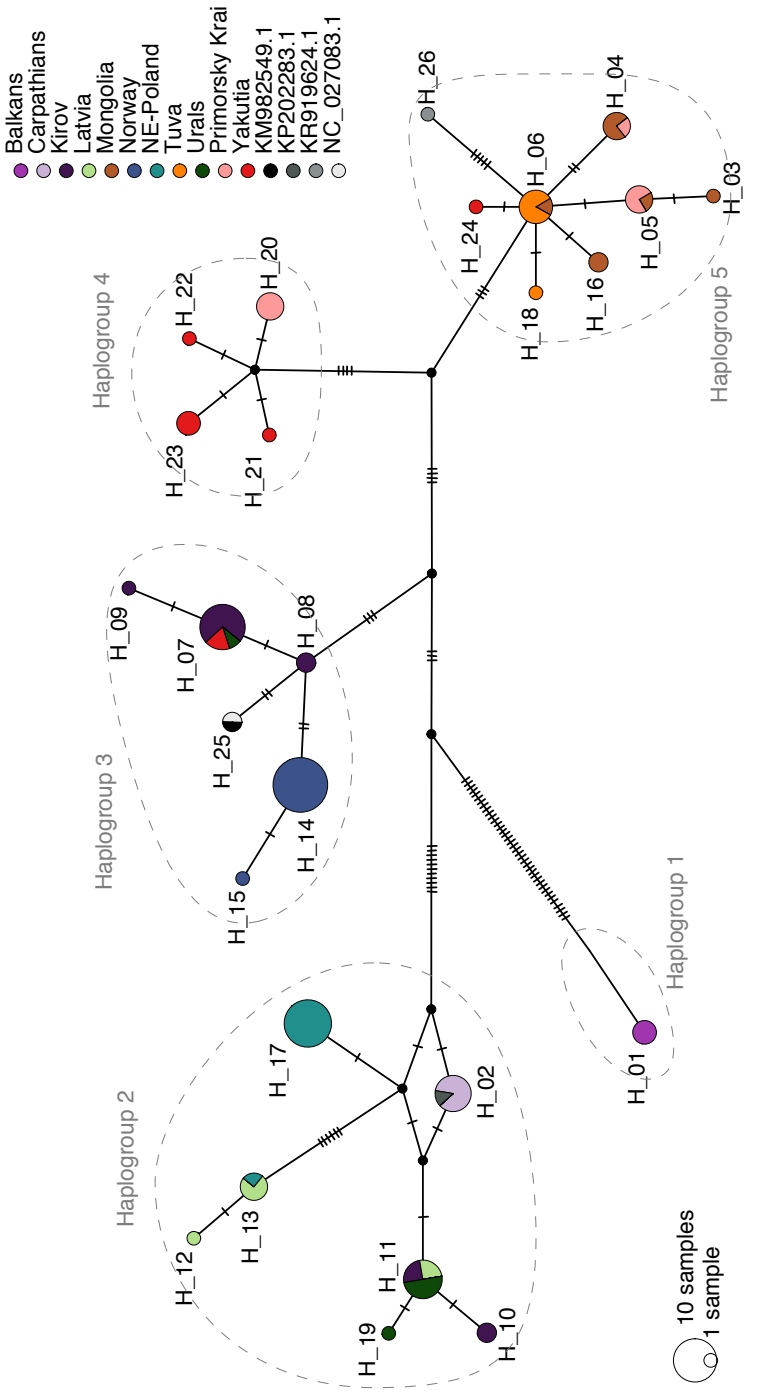


Fig. S20. Mitogenome haplotype network.

SUPPLEMENTARY TABLES

Table S1. Detailed information about the *L. lynx* samples analysed in this study.

Sample	Subspecies last IUCN report	Subspecies Wozencraft, 2005
h_ll_ba_0214	Lynx lynx balcanicus	Lynx lynx balcanicus
h_ll_ba_0215	Lynx lynx balcanicus	Lynx lynx balcanicus
c_ll_ba_0216	Lynx lynx balcanicus	Lynx lynx balcanicus
c_ll_cr_0205	Lynx lynx carpathicus	Lynx lynx carpathicus
c_ll_cr_0206	Lynx lynx carpathicus	Lynx lynx carpathicus
c_ll_cr_0207	Lynx lynx carpathicus	Lynx lynx carpathicus
c_ll_cr_0208	Lynx lynx carpathicus	Lynx lynx carpathicus
c_ll_cr_0209	Lynx lynx carpathicus	Lynx lynx carpathicus
c_ll_cr_0212	Lynx lynx carpathicus	Lynx lynx carpathicus
c_ll_ka_0184	Lynx lynx wrangeli	Lynx lynx kozlovi
c_ll_ka_0186	Lynx lynx wrangeli	Lynx lynx kozlovi
c_ll_ka_0188	Lynx lynx wrangeli	Lynx lynx kozlovi
c_ll_ka_0189	Lynx lynx wrangeli	Lynx lynx kozlovi
c_ll_ki_0086	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0088	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0090	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0091	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0092	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0093	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0094	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0095	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0096	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0097	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0098	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0099	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0100	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0101	Lynx lynx lynx	Lynx lynx lynx
c_ll_ki_0102	Lynx lynx lynx	Lynx lynx lynx
c_ll_la_0044	Lynx lynx lynx	Lynx lynx lynx
c_ll_la_0047	Lynx lynx lynx	Lynx lynx lynx
c_ll_la_0048	Lynx lynx lynx	Lynx lynx lynx
c_ll_la_0052	Lynx lynx lynx	Lynx lynx lynx
c_ll_la_0053	Lynx lynx lynx	Lynx lynx lynx
c_ll_la_0054	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0063	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0064	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0065	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0068	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0069	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0070	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0071	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0072	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0074	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0075	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0076	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0077	Lynx lynx lynx	Lynx lynx lynx

Population	Nuclear WG	WG depth	Library preparation
Balkans	Yes	Low-medium	Library_preparation_2
Balkans	Yes	Low-medium	Library_preparation_2
Balkans	Yes	Low-medium	Library_preparation_2
Carpathian montains	Yes	Low-medium	Library_preparation_1
Carpathian montains	Yes	Low-medium	Library_preparation_1
Carpathian montains	Yes	Low-medium	Library_preparation_1
Carpathian montains	Yes	Low-medium	Library_preparation_1
Carpathian montains	Yes	Low-medium	Library_preparation_1
Carpathian montains	Yes	Medium-high	Library_preparation_4
Mongolia (Central/Khentii aymag)	Yes	Low-medium	Library_preparation_2
Mongolia (Central/Khentii aymag)	Yes	Low-medium	Library_preparation_2
Mongolia (Central/Khentii aymag)	Yes	Low-medium	Library_preparation_2
Mongolia (Central/Khentii aymag)	Yes	Low-medium	Library_preparation_2
Kirov region	No	NA	Library preparation 5
Kirov region	No	NA	Library preparation 5
Kirov region	Yes	Medium-high	Library_preparation_4
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Kirov region	Yes	Low-medium	Library_preparation_1
Latvia	Yes	Low-medium	Library_preparation_2
Latvia	Yes	Low-medium	Library_preparation_2
Latvia	Yes	Low-medium	Library_preparation_2
Latvia	Yes	Low-medium	Library_preparation_2
Latvia	Yes	Low-medium	Library_preparation_2
Latvia	Yes	Low-medium	Library_preparation_2
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	No	NA	Library preparation 5
Norway	Yes	Low-medium	Library_preparation_1
Norway	Yes	Low-medium	Library_preparation_1
Norway	Yes	Low-medium	Library_preparation_1

Sequencing platform	Depth of coverage WG	Mitogenome	Latitude	Longitude	Sex
HiSeq2000,v4	4,77714	Yes	20°48'40.59"E	42°13'52.22"N	Male
HiSeq2000,v4	4,83197	Yes	20°48'40.59"E	42°13'52.22"N	Female
HiSeq2000,v4	4,3064	Yes	19°54'28.59"E	42°44'41.39"N	Male
HiSeq2000,v4	7,97891	Yes	19°45'30.73"E	49°40'17.04"N	Female
HiSeq2000,v4	7,54133	Yes	25°35'57.94"E	45°38'26.45"N	Male
HiSeq2000,v4	8,68415	Yes	22°12'4.41"E	45°23'50.94"N	Female
HiSeq2000,v4	7,61293	Yes	22°17'56.19"E	49°45'2.97"N	Female
HiSeq2000,v4	8,62264	Yes	22°12'4.41"E	45°23'50.94"N	Female
HiSeq X-10	19,2437	Yes	22°26'5.33"E	49°42'28.08"N	Male
HiSeq2000,v4	6,96814	Yes	108°38'11.00"E	48°22'43.00"N	Male
HiSeq2000,v4	6,86957	Yes	108°4'25.00"E	48°24'35.00"N	Female
HiSeq2000,v4	6,00429	Yes	110°29'11.00"E	48°37'21.00"N	Male
HiSeq2000,v4	5,92381	Yes	110°29'11.00"E	48°37'21.00"N	Male
HiSeq2000,v3	NA	Yes	52°48'51.71"E	59°11'53.21"N	NA
HiSeq2000,v3	NA	Yes	50°23'47"E	59°48'53"N	NA
HiSeq X-10	21,8923	Yes	50°23'47"E	59°48'53"N	Female
HiSeq2000,v3	6,10748	Yes	50°23'47"E	59°48'53"N	Male
HiSeq2000,v3	5,31978	Yes	50°23'47"E	59°48'53"N	Female
HiSeq2000,v3	5,29063	Yes	50°23'47"E	59°48'53"N	Male
HiSeq2000,v3	5,84507	Yes	50°23'47"E	59°48'53"N	Male
HiSeq2000,v3	5,29078	Yes	48°24'59"E	60°48'44"N	Male
HiSeq2000,v3	6,12156	Yes	48°24'59"E	60°48'44"N	Male
HiSeq2000,v3	6,02764	Yes	48°24'59"E	60°48'44"N	Female
HiSeq2000,v3	5,77744	Yes	48°24'59"E	60°48'44"N	Female
HiSeq2000,v3	6,00544	Yes	48°24'59"E	60°48'44"N	Female
HiSeq2000,v3	6,17273	Yes	46°43'38"E	61°11'38"N	Male
HiSeq2000,v3	6,29829	Yes	46°43'38"E	61°11'38"N	Female
HiSeq2000,v3	6,07494	Yes	46°43'38"E	61°11'38"N	Male
HiSeq2000,v4	7,73762	Yes	22°31'40.58"E	56°38'33.42"N	Male
HiSeq2000,v4	6,9258	Yes	28°8'43.59"E	56°22'36.28"N	Female
HiSeq2000,v4	5,26228	Yes	22°49'7.98"E	57°21'30.79"N	Female
HiSeq2000,v4	6,83197	Yes	22°35'48.23"E	57°21'2.27"N	Male
HiSeq2000,v4	7,77824	Yes	22°49'0.39"E	57°2'45.09"N	Female
HiSeq2000,v4	8,24717	Yes	25°34'18.55"E	57°40'21.74"N	Male
HiSeq2000,v3	NA	Yes	27°28'49.97"E	70° 4'33.69"N	Female*
HiSeq2000,v3	NA	Yes	23° 1'49.14"E	69°56'22.70"N	NA
HiSeq2000,v3	NA	Yes	21°20'24.98"E	69°29'14.21"N	Female*
HiSeq2000,v3	NA	Yes	23° 1'21.28"E	69° 0'19.20"N	Male*
HiSeq2000,v3	NA	Yes	21°21'54.37"E	69°33'20.90"N	Male*
HiSeq2000,v3	NA	Yes	10°10'28.73"E	62°57'0.70"N	NA
HiSeq2000,v3	NA	Yes	11°37'29.30"E	63°47'59.08"N	Female*
HiSeq2000,v3	NA	Yes	11°37'29.30"E	63°47'59.08"N	Female*
HiSeq2000,v3	NA	Yes	10°48'51.78"E	63°52'41.00"N	Male*
HiSeq2000,v3	5,42769	Yes	12°21'40.04"E	64°30'26.98"N	Male
HiSeq2000,v3	5,79824	Yes	10°40'45.23"E	63°49'28.47"N	Male
HiSeq2000,v3	5,13626	Yes	10°57'15.35"E	63°34'2.38"N	Male

c_ll_no_0078	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0079	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0080	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0081	Lynx lynx lynx	Lynx lynx lynx
c_ll_no_0082	Lynx lynx lynx	Lynx lynx lynx
c_ll_og_0181	Lynx lynx isabellinus	Lynx lynx isabellinus
c_ll_og_0187	Lynx lynx isabellinus	Lynx lynx isabellinus
c_ll_po_0001	Lynx lynx lynx	Lynx lynx lynx
c_ll_po_0002	Lynx lynx lynx	Lynx lynx lynx
c_ll_po_0003	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0004	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0005	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0007	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0010	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0011	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0014	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0017	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0019	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0105	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_po_0106	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_to_0190	*Lynx lynx wrangeli*	*Lynx lynx kozlovi*
c_ll_to_0191	*Lynx lynx wrangeli*	*Lynx lynx kozlovi*
c_ll_tu_0153	*Lynx lynx wrangeli*	*Lynx lynx kozlovi*
c_ll_tu_0157	*Lynx lynx wrangeli*	*Lynx lynx kozlovi*
c_ll_tu_0158	*Lynx lynx wrangeli*	*Lynx lynx kozlovi*
c_ll_tu_0159	*Lynx lynx wrangeli*	*Lynx lynx kozlovi*
c_ll_tu_0165	*Lynx lynx wrangeli*	*Lynx lynx kozlovi*
c_ll_tu_0166	*Lynx lynx wrangeli*	*Lynx lynx kozlovi*
c_ll_ur_0194	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_ur_0195	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_ur_0196	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_ur_0199	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_ur_0200	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_ur_0203	*Lynx lynx lynx*	*Lynx lynx lynx*
c_ll_vl_0107	*Lynx lynx wrangeli*	*Lynx lynx neglectus*
c_ll_vl_0108	*Lynx lynx wrangeli*	*Lynx lynx neglectus*
c_ll_vl_0109	*Lynx lynx wrangeli*	*Lynx lynx neglectus*
c_ll_vl_0110	*Lynx lynx wrangeli*	*Lynx lynx neglectus*
c_ll_vl_0112	*Lynx lynx wrangeli*	*Lynx lynx neglectus*
c_ll_vl_0113	*Lynx lynx wrangeli*	*Lynx lynx neglectus*
c_ll_vl_0128	*Lynx lynx wrangeli*	*Lynx lynx neglectus*
c_ll_vl_0132	*Lynx lynx wrangeli*	*Lynx lynx neglectus*
c_ll_ya_0138	*Lynx lynx wrangeli*	*Lynx lynx wrangeli*
c_ll_ya_0139	*Lynx lynx wrangeli*	*Lynx lynx wrangeli*
c_ll_ya_0140	*Lynx lynx wrangeli*	*Lynx lynx wrangeli*
c_ll_ya_0142	*Lynx lynx wrangeli*	*Lynx lynx wrangeli*
c_ll_ya_0143	*Lynx lynx wrangeli*	*Lynx lynx wrangeli*
c_ll_ya_0145	*Lynx lynx wrangeli*	*Lynx lynx wrangeli*
c_ll_ya_0146	*Lynx lynx wrangeli*	*Lynx lynx wrangeli*
c_ll_ya_0147	*Lynx lynx wrangeli*	*Lynx lynx wrangeli*
c_lr_zz_0001	*Lynx rufus*	*Lynx rufus*

Norway	Yes	Low-medium	Library_preparation_1
Norway	Yes	Low-medium	Library_preparation_1
Norway	Yes	Low-medium	Library_preparation_1
Norway	Yes	Low-medium	Library_preparation_1
Norway	Yes	Low-medium	Library_preparation_1
Mongolia (Omnogovi)	Yes	Low-medium	Library_preparation_2
Mongolia (Omnogovi)	Yes	Low-medium	Library_preparation_2
Białowieża Primeval Forest	Yes	Low-medium	Library_preparation_1
Białowieża Primeval Forest	Yes	Low-medium	Library_preparation_1
Białowieża Primeval Forest	Yes	Low-medium	Library_preparation_1
Białowieża Primeval Forest	No	NA	Library preparation 5
Białowieża Primeval Forest	No	NA	Library preparation 5
Białowieża Primeval Forest	No	NA	Library preparation 5
Białowieża Primeval Forest	No	NA	Library preparation 5
Białowieża Primeval Forest	Yes	Low-medium	Library_preparation_1
Białowieża Primeval Forest	Yes	Low-medium	Library_preparation_1
Białowieża Primeval Forest	No	NA	Library preparation 5
Białowieża Primeval Forest	Yes	Low-medium	Library_preparation_1
Knyszyn Primeval Forests	Yes	Low-medium	Library_preparation_1
Knyszyn Primeval Forests	Yes	Low-medium	Library_preparation_1
Mongolia (Central/Khentii aymag)	Yes	Low-medium	Library_preparation_2
Mongolia (Central/Khentii aymag)	Yes	Low-medium	Library_preparation_2
Tuva Republic	Yes	Low-medium	Library_preparation_2
Tuva Republic	Yes	Low-medium	Library_preparation_2
Tuva Republic	Yes	Low-medium	Library_preparation_2
Tuva Republic	Yes	Low-medium	Library_preparation_2
Tuva Republic	Yes	Low-medium	Library_preparation_2
Tuva Republic	Yes	Low-medium	Library_preparation_2
Ural Mountains	Yes	Low-medium	Library_preparation_1
Ural Mountains	Yes	Low-medium	Library_preparation_1
Ural Mountains	Yes	Low-medium	Library_preparation_1
Ural Mountains	Yes	Low-medium	Library_preparation_1
Ural Mountains	Yes	Low-medium	Library_preparation_1
Ural Mountains	Yes	Low-medium	Library_preparation_1
Primorsky Krai	Yes	Low-medium	Library_preparation_1
Primorsky Krai	Yes	Low-medium	Library_preparation_1
Primorsky Krai	Yes	Low-medium	Library_preparation_1
Primorsky Krai	Yes	Low-medium	Library_preparation_1
Primorsky Krai	Yes	Medium-high	Library_preparation_3
Primorsky Krai	Yes	Low-medium	Library_preparation_1
Primorsky Krai	Yes	Low-medium	Library_preparation_1
Primorsky Krai	Yes	Low-medium	Library_preparation_1
Yakutia Republic	Yes	Low-medium	Library_preparation_1
Yakutia Republic	Yes	Low-medium	Library_preparation_1
Yakutia Republic	Yes	Low-medium	Library_preparation_1
Yakutia Republic	Yes	Low-medium	Library_preparation_1
Yakutia Republic	Yes	Low-medium	Library_preparation_1
Yakutia Republic	Yes	Low-medium	Library_preparation_1
Yakutia Republic	Yes	Medium-high	Library_preparation_3
Yakutia Republic	Yes	Low-medium	Library_preparation_1
Na	Yes	Medium-high	Library_preparation_4

HiSeq2000,v3	5,25445	Yes	12°8'49.62"E	61°23'2.66"N	Male
HiSeq2000,v3	5,09435	Yes	8°24'42.53"E	59°23'55.42"N	Male
HiSeq2000,v3	5,21795	Yes	8°43'38.54"E	60°39'53.79"N	Female
HiSeq2000,v3	5,53956	Yes	8°58'2.81"E	60°16'28.74"N	Female
HiSeq2000,v3	5,36419	Yes	9°1'43.70"E	60°8'40.04"N	Male
HiSeq2000,v4	6,79024	Yes	101°2'52.00"E	43°13'30.00"N	Male
HiSeq2000,v4	5,63946	Yes	101°2'52.00"E	43°13'30.00"N	Female
HiSeq2000,v4	6,51928	Yes	23°35'43.68"E	52°57'15.72"N	Male
HiSeq2000,v4	6,61355	Yes	23°50'35.00"E	52°49'22.81"N	Female
HiSeq2000,v4	6,43528	Yes	23°43'12.86"E	52°47'48.32"N	Female
HiSeq2000,v3	NA	Yes	23°23'39.63"E	52°50'33.07"N	Female*
HiSeq2000,v3	NA	Yes	23°50'35.00"E	52°49'22.81"N	NA
HiSeq2000,v3	NA	Yes	23°50'35.00"E	52°49'22.81"N	Female*
HiSeq2000,v3	NA	Yes	23°45'52.35"E	52°39'54.96"N	Male*
HiSeq2000,v4	6,44556	Yes	23°30'51.37"E	52°53'30.49"N	Male
HiSeq2000,v4	6,38699	Yes	23°56'43.94"E	52°46'6.68"N	Male
HiSeq2000,v3	NA	Yes	23°36'41.72"E	52°39'40.52"N	Female*
HiSeq2000,v4	6,376	Yes	23°44'36.44"E	52°40'50.61"N	Male
HiSeq2000,v4	6,05977	Yes	23°28'25.76"E	53°5'26.62"N	Male
HiSeq2000,v4	6,10897	Yes	23°39'32.61"E	53°6'16.21"N	Female
HiSeq2000,v4	7,33784	Yes	108°16'23.00"E	47°23'55.00"N	Female
HiSeq2000,v4	7,71857	Yes	108°41'1.00"E	47°37'48.00"N	Male
HiSeq2000,v4	8,10689	Yes	96°32'00.0"E	51°29'00.0"N	Female
HiSeq2000,v4	7,80348	Yes	96°32'00.0"E	51°29'00.0"N	Male
HiSeq2000,v4	8,0083	Yes	96°32'00.0"E	51°29'00.0"N	Female
HiSeq2000,v4	7,73842	Yes	96°32'00.0"E	51°29'00.0"N	Female
HiSeq2000,v4	7,76115	Yes	96°32'00.0"E	51°29'00.0"N	Female
HiSeq2000,v4	8,05401	Yes	96°32'00.0"E	51°29'00.0"N	Male
HiSeq2000,v4	11,0239	Yes	59°38'33.00"E	56°2'53.00"N	Female
HiSeq2000,v4	11,7718	Yes	60°7'49.94"E	55°8'58.01"N	Male
HiSeq2000,v4	12,3771	Yes	60°7'49.94"E	55°8'58.01"N	Male
HiSeq2000,v4	13,1596	Yes	59°47'19.00"E	55°10'54.00"N	Male
HiSeq2000,v4	12,6612	Yes	59°0'1.97"E	55°0'2.00"N	Male
HiSeq2000,v4	13,3973	Yes	57°21'45.00"E	54°59'58.00"N	Female
HiSeq2000,v4	6,50257	Yes	136°28'21.71"E	44°56'8.45"N	Female
HiSeq2000,v4	11,8787	Yes	135°34'45.07"E	45°17'37.84"N	Female
HiSeq2000,v4	6,04097	Yes	136°43'52.76"E	45°39'59.98"N	Male
HiSeq2000,v4	7,07492	Yes	132°54'50.74"E	49°0'53.30"N	Male
HiSeq X-10	28,9235	Yes	137°0'33.44"E	45°51'25.99"N	Female
HiSeq2000,v4	8,36169	Yes	137°0'33.44"E	45°51'25.99"N	Male
HiSeq2000,v4	7,96396	Yes	137°25'50.84"E	45°54'27.11"N	Male
HiSeq2000,v4	8,31332	Yes	137°25'50.84"E	45°54'27.11"N	Male
HiSeq2000,v4	7,94849	Yes	132°4'9.98"E	59°54'3.71"N	Male
HiSeq2000,v4	7,97872	Yes	129°12'17.68"E	61°11'21.23"N	Male
HiSeq2000,v4	8,15324	Yes	129°9'42.35"E	61°9'33.70"N	Male
HiSeq2000,v4	8,29523	Yes	136°10'36.67"E	66°53'17.09"N	Male
HiSeq2000,v4	7,8525	Yes	136°10'36.67"E	66°53'17.09"N	Male
HiSeq2000,v4	8,26097	Yes	129°26'43.23"E	61°53'35.84"N	Male
HiSeq X-10	22,9182	Yes	130°40'44.70"E	60°46'34.41"N	Male
HiSeq2000,v4	8,07335	Yes	127°18'16.14"E	60°45'5.54"N	Female
HiSeq X-10	25,6619	No	NA	NA	NA

Table S2. Climatic, habitat, historical records, taxonomic description and estimated current Ne of the studied *Lynx lynx* populations.

	Population	**Climate**[1]	**Land cover**[2]	**Demography**[3]
1	North-Eastern Poland (Białowieża and Knyszyn Primeval Forests (BPF & KPF))	Cold, without dry season, warm summer	Temperate Broadleaf & Mixed Forests	Bottlenecked
2	Balkans	Temperate, without dry season, warm summer	Temperate Broadleaf & Mixed Forests	Bottlenecked
3	Carpathian Mountains	Cold, without dry season, warm summer	Temperate Conifer Forests	Bottlenecked
4	Latvia	Cold, without dry season, warm summer	Temperate Broadleaf & Mixed Forests	Healthy
5	Norway	Cold, without dry season, cold summer	Boreal Forests/Taiga	Bottlenecked
6	Kirov region, Russia	Cold, without dry season, cold summer	Boreal Forests/Taiga	Healthy
7	Ural Mountains, Russia	Cold, without dry season, warm summer	Boreal Forests/Taiga	Healthy
8	Tuva (the Republic of Tyva), Russia	Cold, dry winter, cold summer	Temperate Conifer Forests	Healthy
9	Yakutia (Republic of Sakha), Russia	Cold, dry winter, very cold winter	Boreal Forests/Taiga	Healthy
10	Primorsky Krai, Russia	Cold, dry winter, warm summer	Temperate Broadleaf & Mixed Forests	Healthy
11	Mongolia (Omnogovi)	Arid, desert, cold	Deserts & Xeric Shrublands	Healthy
12	Mongolia (Central/Khentii aymag)	Cold, dry winter, cold summer	Temperate Grasslands, Savannas & Shrublands	Healthy

Human impact data[3]
• Lynx population reached minimum at the turn of 19th and 20th century, restricted to BPF (Bieniek et al. 1998) • Heavy extirpation at BPF; 20yrs of practical lynx absence 1890-1914 (Jędrzejewski et al. 1996) • Short bottleneck 1960-1970, followed by recovery of the population helped by immigrating
• Considered as the most endangered population of the Eurasian lynx due to its isolation, past bottleneck and currently decreasing numbers (Von Arx et al. 2004) • At the turn of 19th and 20th centuries the population was extirpated from most of its range, with only 15 – 20 individuals surviving • Recovered to 50 – 70 individuals in 1950 – 1960s and reached up to 280 lynx in 1974
• Considered large but isolated • Intense persecution at the turn of 19th and 20th centuries • Romania: Extermination policiy between 1891 – 1921 left around 100 indivdiuals in 1930. Recovery after protection in 1933, with 10-fold increase in the following 25 years and around 500 lynxes estimated in 1950 (Kratochvil 1968) • Slovakia: Extermination policiy in the period 1883-1936 (Hell and Slamečka 1996). Partially protected since then resulting in 400 to 550 individuals in 1960s -1990s and 200 individuals currently
• Considered part of the core of the Northwestern lynx population • Started to decline at the end of 17th century • Significant decline at the beginning of the 20th century; with 30 lynx counted by rangers in 1933 • Population recovery starting 1945, and increasing steadily after 1960 (Andersone et al. 2003) • Recent status assessed as favourable, range fully recovered and 600 – 800 individuals estimaated at present (Bagrade et al. 2016)
• 1846-1981 large carnivore extermination policy led to dramatic reduction of lynx numbers with few survivors left in central Norway for 40 yrs (1926-1965) (Linnell et al. 2010) • In 1960s the recovery of the population started from few core remnant areas, and little contribution of immigrants from Russia (Rueness et al. 2002) • Similar dinamycs in neighbouring Sweden: only 30 to 100 individuals left in the country between
• 50 to 400 lynx harvested yearly between 1930 and 1980 (Kozlovskiy, 2003) • Population dynamics fluctuate by 7-8 folds with peaks every 16-17 years (Kozlovskiy, 2003)
• Recorded population fluctuations since 1934 with fur harvest peaking during 1940s (300 skins yearly) and a decline towards 1970s (Malafeev et al. 1986)
• 100 individuals harvested anually ~1845 (Kogan 1933) • High numbers suggested by harvest records between 1927-1988 • Population decline between 1969-1983 (Only 2 skins harvested) • Maximum harvest during 1941-1949 (526 skins) (Smirnov, 1997)
• Cyclic population fluctuations with peaks every 10-13 years • 20-730 fur harvested during 1930-1990 with highest numbes reported in 1930 (Mordosov 2003) • On the border of the 19th and 20th centuries Siberia gave 10,000 pelts annually (Brass 1911)
• Described as abundant and widely distributed • Peak yearly harvest during 1930-1940s with decline towards 1980 (Matyushkin et al. 2003) • 400-1200 individuals harvested yearly during 1920s (Heptner and Sludskiy, 1972)
• 1500 individuals harvested anually ~1845 (Kogan 1933) • Very heavy exploitation of lynx at the beginning of 20th century with up to 4000 skins exported annually (Matyushkin et al. 2003)

Subspecies[4]	Estimated Ne	Ne_0.5(Jack-Knife on samples)	Ne_0.95(Jack-Knife on samples)
L.l.lynx	46,8	24,4	271,8
L.l.balcanicus	NA	NA	NA
L.l.carpathicus	11	3,1	87,7
L.l.lynx	INF	64	INF
L.l.lynx	29	15	133,1
L.l.lynx	1093,9	229,2	INF
L.l.lynx	INF	INF	INF
L.l.wrangeli / L.l.kozlovi	INF	INF	INF
L.l.wrangeli	32,7	11	INF
L.l.wrangeli/L.l. neglectus	46,8	14,9	INF
L.l. isabellinus	45,3	12,3	INF
L.l. wrangeli			

1 Peel, M. C., Finlayson, B. L., and McMahon, T. A.: Updated world map of the Köppen-Geiger climate classification, Hydrol. Earth Syst. Sci., 11, 1633-1644

2 http://ecoregions2017.appspot.com/

3 Andersone Z., Ozolins Ya., Pupila A., Bagrade G. 2003. The East European and Caucasian parts of lynx range (the western group of regions): Latvia. – In: Matyushkin Ye.N., Vaisfeld M.A. (eds.) The Lynx: Regional Features of Ecology, Use and Protection. Moscow: Nauka, pp. 92 – 104. (in Russian/ English)

3 Bagrade G., Ruņģis D.E., Ornicāns A., Šuba J., Žunna A., Howlett S.J., Lūkins M., Gailīte A., Stepanova A., Done G., Gaile A., Bitenieks K., Mihailova L., Baumanis J., Ozoliņš J. 2016. Status assessment of Eurasian lynx in Latvia linking genetics and demography—a growing population or a source–sink process? Mammal Research, 61: 337–352

3 Bieniek, M., Wolsan, M., & Okarma, H. (1998). Historical biogeography of the lynx in Poland. Acta Zoologica Cracoviensia, 41,143-167.

3 Brass, E. (1911). Aus dem Reiche der Pelze. - Imberg & Lefson GmbH, Berlin. 709 s. pp 353

3 Hell P., Slamečka, J. 1996: Current status of the lynx *(Lynx lynx)* in Slovakia. Acta Sc. Nat. Brno 30 (3): 64-78.

3 Hellborg, L., Walker, C. W., Rueness, E. K., Stacy, J. E., Kojola, I. et al. (2002). Differentiation and levels of genetic variation in northern European lynx (*Lynx lynx*) populations revealed by microsatellites and mitochondrial DNA analysis. Conservation Genetics 3, 97-111

3 Heptner, V. G., & Sludskiy, A. A. (1972). Carnivores. Hyaenas and Felids. In: Mammals of the Soviet Union. Vol. 2. Part 2. Heptner, V. G. & Naumov, N. P. (Eds) Vysshaya Shkola Publishers, Moscow. [in Russian]

3 Jędrzejewski, W., Jędrzejewska, B., Okarma, H., Schmidt, K., Bunevich, A. N., & Miłkowski, L. (1996). Population dynamics (1869-1994), demography and home ranges of the lynx in Bialowieza Primeval Forest (Poland and Belaruss). Ecography, 19, 122-138.

3 Kogan M.I. 1933. Foreign fur markets. Soviet Asia.Publ., Moscow. 243 s.[in Russian]

3 Kozlovskiy, I. S. (2003). The Vyatka-Kama Region. In: Matyushkin EN, Vaisfeld MA, (Eds) The Lynx. Regional features of ecology, use and protection. Moscow: Nauka. pp. 168-217

3 Kratochvil J. 1968. The lynx population in Rumania. Acta Sc. Nat. Brno 2 (5/6): 65-70.

3 Kubala J., Smolko P., Zimmermann F., Rigg R., Tám B., Ilko T., Foresti D., Breitenmoser-Würsten C., Kropil R., Breitenmoser U. 2019. Robust monitoring of the Eurasian *lynx Lynx* lynx in the Slovak Carpathians reveals lower numbers than officially reported. Oryx 53: 548-556

3 Linnell, J. D. C., Broseth, H., Odden, J., & Nilsen, E. B. (2010). Sustainably harvesting a large carnivore? Development of Eurasian lynx populations in Norway during 160 years of shifting policy. Environmental Management 45, 1142–1154

3 Malafeev, Y. M., Kryazhimskii, F. W., & Dobrinskii, L. N. (1986). Analysis of lynx population in the Middle Ural. Sverdlovsk: Akademya Nauk SSSR [in Russian].

3 Matyushkin, E. N., Podolskiy, S. A., & Tkachenko K. N. (2003). The south of Far East. In: Matyushkin, E. N., & Vaisfeld, M. A. (Eds) The Lynx. Regional features of ecology, use and protection. Moscow: Nauka. pp 367-387, 423-472. [in Russian with English summary]

3 Kogan M.I. 1933. Foreign fur markets. Soviet Asia.Publ., Moscow. 243 s. pp 214, 219 [in Russian]

3 Mordosov, I. I. (2003). Yakutia. In: Matyushkin, E.N., & Vaisfeld, M.A. (Eds) The Lynx. Regional features of ecology, use and protection. Moscow: Nauka. pp 290-303. [in Russian with English summary]

3 Rueness, E. K., Jorde, P. E., Hellborg, L., Stenseth, N. C., Ellegren, H., & Jakobsen, K. S. (2003). Cryptic population structure in a large, mobile mammalian predator: the Scandinavian lynx. Molecular Ecology 12, 2623-2633.

3 Smirnov, M. N. (1997). Peculiarities of distribution and ecology of the lynx (Felis lynx L. 1756) in the upper reaches of Yenissei. In: Fauna and ecology of terrestrial vertebrates of Siberia. Ministry of General and Professional Education of the Russian Federation, Krasnoyarsk.[in Russian]

3 Von Arx M, Breitenmoser-Würsten C, Zimmermann F, Breitenmoser U (2004) Status and conservation of the Eurasian lynx (*Lynx lynx*) in 2001. Muri: KORA Bericht no 19. pp 330.

4 Kitchener A. C., Breitenmoser-Würsten Ch., Eizirik E., Gentry A., Werdelin L., Wilting A., Yamaguchi N., Abramov A. V., Christiansen P., Driscoll C., Duckworth J. W., Johnson W., Luo S.-J., Meijaard E., O'Donoghue P., Sanderson J., Seymour K., Bruford M., Groves C., Hoffmann M., Nowell K., Timmons Z. & Tobe S. 2017. A revised taxonomy of the Felidae. The final report of the Cat Classification Task Force of the IUCN/SSC Cat Specialist Group. Cat News Special Issue 11, 80 pp.

Table S3. Z score values from ThreePop analyses. The test population is indicated in the upper-left cell, with western populations displayed in rows and eastern populations in columns. A) Tuva as the test population. B) Yakutia as the test population.

A)

Tuva	Mongolia	Yakutia	Primorsky Krai
Ural mountains	-39,5	3,49	-49,07
Kirov	-34,5	5,93	-47,4
Carpathian Mountains	-5,56	14,74	-7,8

B)

Yakutia	Mongolia	Tuva	Primorsky Krai
Ural mountains	56,91	133,06	20,43
Kirov	62,24	135,24	22,88
Carpathian Mountains	61,72	114,2	33,94

Table S4. Delta values for different K values calculated following Evanno (2005).

Delta(K=2)	177988180,8
Delta(K=3)	4,373180862
Delta(K=4)	0,135164532
Delta(K=5)	1,405316641
Delta(K=6)	1,919383265
Delta(K=7)	1,732917785
Delta(K=8)	1,558812815
Delta(K=9)	0,1360717
Delta(K=10)	0,369484886
Delta(K=11)	0,594390785
Delta(K=12)	0,263857417

Table S5. Pairwise weighted FST values.

	Kirov	Latvia	Norway	NE-Pol	Mongolia	Tuva	Urals	P. Krai	Yakutia
Carpathians	0.213065	0.241779	0.347542	0.30614	0.299333	0.271568	0.249896	0.342122	0.298078
Kirov	NA	0.094969	0.187046	0.171447	0.227579	0.175641	0.051613	0.274871	0.216736
Latvia	NA	NA	0.237903	0.159947	0.244883	0.2017	0.130275	0.292338	0.238369
Norway	NA	NA	NA	0.313762	0.32172	0.28839	0.231778	0.365709	0.317561
NE-Poland	NA	NA	NA	NA	0.303215	0.267466	0.209452	0.346982	0.299569
Mongolia	NA	NA	NA	NA	NA	0.073926	0.239251	0.102906	0.081348
Tuva	NA	NA	NA	NA	NA	NA	0.188265	0.125433	0.087282
Urals	NA	NA	NA	NA	NA	NA	NA	0.289338	0.22914
Primorsky Krai	NA	NA	NA	NA	NA	NA	NA	NA	0.117371

Table S6. Diversity values (π, and Θ), Tajima's D, and Ne (and confidence interval) for *L. lynx* populations.

Population	Watterson estimator	Pi	Tajima's D	Ne	Ne_0.5 (Jack-Knife on samples)	Ne_0.95 (Jack-Knife on samples)
Yakutia	4,50E-04	5,27E-04	2,50E-05	32,7	11	INF
Urals	3,95E-04	4,51E-04	2,32E-05	INF	INF	INF
Tuva	4,89E-04	5,45E-04	1,74E-05	INF	INF	INF
Primorsky Krai	4,25E-04	5,01E-04	2,60E-05	46,8	14,9	INF
Norway	2,66E-04	3,41E-04	3,67E-05	29	15	133,1
NE-Poland	2,86E-04	3,67E-04	3,93E-05	46,8	24,4	271,8
Mongolia	4,63E-04	5,36E-04	2,15E-05	45,3	12,3	INF
Latvia	3,74E-04	4,38E-04	2,54E-05	INF	64	INF
Kirov	3,58E-04	4,64E-04	4,12E-05	1093,9	229,2	INF
Carpathians	3,15E-04	3,74E-04	2,75E-05	11	3,1	87,7

Table S7. Mitogenomic diversity in Eurasian Lynx populations.

Population	N	# Hap	S	Hd (sd)	π (sd) (‰)	K	Tajima's D	Fu&Li F
Yakutia	8	5	19	0,86 (0,11)	0,46 (0,28)	7.61	0.1979	0.55
Kirov	15	5	22	0,7 (0,11)	0,54 (0,3)	8.86	12.752	1.55
Primorsky Krai	8	3	11	0,68 (0,12)	0,34 (0,21)	5.61	16.065	1.24
Tuva	6	2	1	0,33 (0,22)	0,02 (0,03)	0.33	-0.933	-0.7
Mongolia	8	5	5	0,86 (0,11)	0,13 (0,09)	2.18	0.5884	0.85
Carpathian m.	6	1	0	0 (0)	0 (0)	0	NA	NA
Latvia	6	3	9	0,73 (0,16)	0,28 (0,18)	4.6	0.9949	1.36
NE-Poland	13	2	7	0,15 (0,13)	0,07 (0,05)	1.08	-1,982*	-2.44
Norway	17	2	1	0,12 (0,1)	0,01 (0,01)	0.12	-11.639	-1.48
Balcans	3	1	0	0 (0)	0 (0)	0	NA	NA
Ural m.	6	3	22	0,6 (0,22)	0,45 (0,28)	7.33	-14.981	-1.2
Total	96	24	89	0,93 (0,01)	1 (0,5)	16.31	-0.1929	0.71

Table S8. List of mitogenome haplotypes.

	Population															
Haplotype	Balkans	Carpathian m.	Kirov	Latvia	Mongolia	Norway	NE-Poland	Tuva	Ural m.	Primorsky Krai	Yakutia	KM982549.1	KP202283.1	KR919624.1	NC_027083.1	TOTAL
H_01	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
H_02	0	6	0	0	0	0	0	0	0	0	0	0	1	0	0	7
H_03	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
H_04	0	0	0	0	3	0	0	0	0	1	0	0	0	0	0	4
H_05	0	0	0	0	1	0	0	0	0	3	0	0	0	0	0	4
H_06	0	0	0	0	1	0	0	5	0	0	0	0	0	0	0	6
H_07	0	0	8	0	0	0	0	0	1	0	2	0	0	0	0	11
H_08	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2
H_09	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
H_10	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2
H_11	0	0	2	2	0	0	0	0	4	0	0	0	0	0	0	8
H_12	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
H_13	0	0	0	3	0	0	1	0	0	0	0	0	0	0	0	4
H_14	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	16
H_15	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
H_16	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2
H_17	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	12
H_18	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
H_19	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
H_20	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	4
H_21	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
H_22	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
H_23	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	3
H_24	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
H_25	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	2
H_26	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
TOTAL	3	6	15	6	8	17	13	6	6	8	8	1	1	1	1	100

Haplotype	SNPs Position																		
	4	18	160	611	859	924	1096	1186	2076	2801	3341	3359	3451	3526	3665	3762	4645	4670	5047
H_01	C	G	C	C	G	C	C	C	T	T	T	T	G	T	T	G	C	T	T
H_02	.	.	T	.	A	.	.	T	C	C	C	.	.	.	C	.	.	C	.
H_03	.	.	T	.	A	T	.	T	.	C	.	C	A	.	C	.	.	.	.
H_04	.	.	T	.	A	T	.	T	.	C	.	C	A	.	C	.	T	.	.
H_05	.	.	T	.	A	T	.	T	.	C	.	C	A	.	C	.	.	.	.
H_06	.	.	T	.	A	T	.	T	.	C	.	C	A	.	C	.	.	.	.
H_07	.	.	T	.	A	.	.	T	.	C	.	.	.	.	C	.	.	.	.
H_08	.	.	T	.	A	.	.	T	.	C	.	.	.	.	C	.	.	.	.
H_09	.	.	T	.	A	.	.	T	.	C	.	.	.	.	C	.	.	.	.
H_10	.	.	.	.	A	.	.	T	C	C	C	.	.	.	C	.	.	C	.
H_11	.	.	.	.	A	.	.	T	C	C	C	.	.	.	C	.	.	C	.
H_12	.	.	.	.	A	.	.	T	C	C	.	.	.	.	C	.	.	C	C
H_13	.	.	.	.	A	.	.	T	C	C	.	.	.	.	C	.	.	C	C
H_14	.	.	T	.	A	.	.	T	.	C	.	.	.	.	C	.	.	.	.
H_15	.	.	T	.	A	.	.	T	.	C	.	.	.	.	C	.	.	.	.
H_16	.	.	T	.	A	T	.	T	.	C	.	C	A	.	C	.	.	.	.
H_17	.	.	.	.	A	.	.	T	C	C	C	.	.	.	C	.	.	C	.
H_18	.	.	T	.	A	T	.	T	.	C	.	C	A	.	C	.	.	.	.
H_19	.	.	.	.	A	.	.	T	C	C	C	.	.	.	C	.	.	C	.
H_20	.	.	T	T	A	.	.	T	.	C	.	.	A	.	C	A	.	.	.
H_21	.	.	T	T	A	.	.	T	.	C	.	.	A	C	C	.	.	.	.
H_22	.	.	T	T	A	.	.	T	.	C	.	.	A	.	C	.	.	.	.
H_23	.	.	T	T	A	.	.	T	.	C	.	.	A	.	C	.	.	.	.
H_24	.	.	T	.	A	T	.	T	.	C	.	C	A	.	C	.	.	.	.
H_25	.	.	T	.	A	.	T	T	.	C	.	.	.	.	C	.	.	.	.
H_26	G	A	T	.	A	T	.	T	.	C	.	C	A	.	C	.	.	.	.

5151	5465	6046	6205	6340	6512	6787	6973	7132	7189	7448	7584	7591	7948	8066	8230	8380	8466	8511	8751	9080	9263
A	T	G	C	A	C	A	A	T	T	G	C	C	T	C	A	T	C	A	T	G	A
T	C	A	T	.	T	G	.	.	C	A	.	T	.	.	G	.	.	.	C	.	.
T	.	A	T	.	T	G	.	.	C	A	T	T	.	T	G	.	.	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	C	.	G	.	T	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	C	.	G	.	T	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	C	.	G	.	T	G	C	.	G
T	C	A	T	.	T	G	.	.	C	A	.	T	.	.	G	.	.	.	C	.	.
T	C	A	T	.	T	G	.	.	C	A	.	T	.	.	G	.	.	.	C	.	.
T	C	A	T	.	T	G	G	.	C	A	.	T	.	.	G	C	.	.	C	.	.
T	C	A	T	.	T	G	G	.	C	A	.	T	.	.	G	C	.	.	C	.	.
T	.	A	T	G	T	G	.	.	C	A	.	T	C	.	G	.	T	G	C	.	G
T	.	A	T	G	T	G	.	.	C	A	.	T	C	.	G	.	T	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	.	G
T	C	A	T	.	T	G	.	.	C	A	.	T	.	.	G	.	.	.	C	.	.
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	.	G
T	C	A	T	.	T	G	.	C	C	A	.	T	.	.	G	.	.	.	C	.	.
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	A	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	A	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	A	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	A	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	C	.	G	.	T	G	C	.	G
T	.	A	T	.	T	G	.	.	C	A	.	T	.	T	G	.	.	G	C	.	G

9689	9773	9866	9895	10395	10586	10608	10826	10879	10925	10966	10988	11071	11318	11340	11411	11507	11710	12237	12362	12372	12437
G	A	A	A	G	A	A	G	C	T	G	T	A	C	G	T	G	G	T	C	A	C
A	G	.	G	A	G	.	A	.	C	A	C	G	T	.	.	A	A	.	.	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	C	A	.	.	T	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	C	A	.	.	T	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	C	A	.	.	T	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	C	A	.	.	T	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	.	A	.	.	.	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	.	A	.	.	.	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	.	A	.	.	.	G	T
A	G	.	G	A	G	.	A	.	C	A	C	G	.	.	.	A	A	.	.	.	T
A	G	.	G	A	G	.	A	.	C	A	C	G	T	.	.	A	A	.	.	.	T
A	G	.	G	A	G	.	A	.	C	A	C	G	T	.	.	A	A	C	.	G	T
A	G	.	G	A	G	.	A	.	C	A	C	G	T	.	.	A	A	C	.	G	T
A	G	.	.	A	.	G	A	.	C	.	C	G	.	.	.	A	.	.	.	G	T
A	G	.	.	A	.	G	A	.	C	.	C	G	.	.	.	A	.	.	.	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	C	A	.	.	T	G	T
A	G	G	G	A	G	.	A	.	C	A	C	G	T	.	.	A	A	.	.	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	C	A	.	.	T	G	T
A	G	.	G	A	G	.	A	.	C	A	C	G	T	.	.	A	A	.	.	.	T
A	G	.	.	A	.	.	A	T	C	.	C	G	.	.	.	A	.	.	T	G	T
A	G	.	.	A	.	.	A	T	C	.	C	G	.	.	.	A	.	.	T	G	T
A	G	.	.	A	.	.	A	T	C	.	C	G	.	.	.	A	.	.	T	G	T
A	G	.	.	A	.	.	A	T	C	.	C	G	.	.	.	A	.	.	T	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	A	C	A	.	.	T	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	.	A	.	.	.	G	T
A	G	.	.	A	.	.	A	.	C	.	C	G	.	.	C	A	.	.	T	G	T

12815	12908	12950	12973	13245	13355	13533	13637	13712	14104	14163	14352	14404	14419	14440	14570	14964	15109	15114	15362	15858	16108
A	C	G	T	G	A	A	G	T	T	C	A	T	A	A	A	T	G	T	A	T	C
G	.	A	.	.	.	.	A	.	.	.	G	C	.	G	.	.	A	.	.	C	A
G	.	A	.	.	G	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	G	.	A	C	.	.	.	C	.	.	.	.	A	C	.	C	A
G	.	A	.	.	G	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	G	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	C	.	.	.	C	.	.	.	C	A	.	.	C	A
G	.	A	.	.	.	.	A	.	.	.	G	C	.	G	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	.	.	.	G	C	.	G	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	.	C	.	G	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	.	.	.	G	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	T	A	.	.	.	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	C	.	G	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	.	.	.	G	C	.	G	.	.	A	.	.	C	A
G	.	A	.	A	G	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	.	.	.	G	C	.	G	.	.	A	.	.	C	A
G	.	A	.	.	G	.	A	C	.	.	.	C	G	.	.	.	A	.	.	C	A
G	.	A	.	.	G	.	A	C	.	.	.	C	G	.	.	.	A	.	.	C	A
G	.	A	.	.	G	.	A	C	.	T	.	C	G	.	.	.	A	.	.	C	A
G	.	A	.	.	G	G	A	C	.	.	.	C	G	.	.	.	A	.	.	C	A
G	.	A	.	.	G	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	.	.	A	C	.	.	.	C	.	.	.	.	A	.	.	C	A
G	.	A	.	.	G	.	A	C	.	.	.	C	.	.	G	.	A	.	G	C	A

16154	16160	16161	16197	16208	16226	16306	16389	16437	**Haplogroup**
T	A	C	G	A	T	T	T	A	Hg_1
C	G	.	A	.	C	.	C	.	Hg_2
C	G	.	A	G	C	C	C	.	Hg_5
C	G	.	A	.	C	C	C	.	Hg_5
C	G	.	A	G	C	C	C	.	Hg_5
C	G	.	A	.	C	C	C	.	Hg_5
C	G	T	A	.	C	C	C	G	Hg_3
C	G	T	A	.	C	C	C	.	Hg_3
C	G	T	A	.	C	C	C	G	Hg_3
C	G	.	A	.	C	.	C	.	Hg_2
C	G	.	A	.	C	.	C	.	Hg_2
C	G	.	A	.	C	C	C	.	Hg_2
C	G	.	A	.	C	C	C	.	Hg_2
C	G	T	A	.	C	C	C	.	Hg_3
C	G	T	A	.	C	C	C	.	Hg_3
C	G	.	A	.	C	C	C	.	Hg_5
C	G	.	A	.	C	C	C	.	Hg_2
C	G	.	A	.	C	C	C	.	Hg_5
C	G	.	A	.	C	.	C	.	Hg_2
C	G	.	A	.	C	C	C	.	Hg_4
C	G	.	A	.	C	C	C	.	Hg_4
C	G	.	A	.	C	C	C	.	Hg_4
C	G	.	A	.	C	C	C	.	Hg_4
C	G	.	A	.	C	C	C	.	Hg_5
.	G	T	A	.	C	C	C	.	Hg_3
.	G	.	A	.	C	C	C	.	Hg_5

02 CAPÍTULO

ANCIENT GENOME PROVIDES INSIGHTS INTO THE HISTORY OF EURASIAN LYNX IN IBERIA AND WESTERN EUROPE

María Lucena-Pérez

Enrico Bazzicalupo, Johanna Paijmans, Daniel Kleinman, Love Dalen, Michael Hofreiter, Miguel Delibes, Miguel Clavero, José A. Godoy.

ABSTRACT

The Eurasian lynx (Lynx lynx) is one of the most widely distributed felids in the world. However, most of its populations are nowadays declining following a trend that started a few millennia ago. Historical declines have been especially severe in Europe, and particularly in Western Europe, from where the species disappeared during the last few centuries. Here, we analyze the whole genome sequence of an Eurasian lynx inhabiting the Iberian Peninsula 2,500 ya, to gain insights into the phylogeographic position and genetic status of this extinct population. Also, we contextualize previous ancient data in the light of the new phylogeographic studies for the specie. Our results suggest that the Iberian population is part of an extinct European lineage closely related and probably ancestral to current Carpathian-Baltic lineages. Also, this sample holds the lowest Eurasian lynx diversity reported so far, similar to the highly endangered Iberian lynx. A combination of historical factors, such as a founder effect while colonizing the peninsula, together with intensified human impacts during the Holocene in the Cantabrian strip, could have led to a genetic impoverishment of the population and precipitated its extinction. Mitogenomic lineages distribution in space and time supports the long-term coexistence of several lineages of Eurasian lynx in Western Europe with fluctuating ranges. While the mitogenomic lineage related to the one currently found in Balkans and Caucasus was predominant during the Pleistocene, another lineage related to the one currently distributed in Central Europe prevailed during the beginning of the Holocene.The use of ancient genomics has proven to be a useful tool to understand the biogeographic pattern of the Eurasian lynx in the past and to provide background knowledge and baselines to assist conservation actions.

INTRODUCTION

The genus Lynx diverged from the lineage of Asian cats in the genera Catopuma and Pardofelis between 8 and 10 Mya (Li et al., 2016). There are 4 extant Lynx species, the bobcat (L. rufus) and Canadian lynx (L. canadensis) in North America and the Eurasian lynx (L. lynx) and the Iberian lynx (L. pardinus) in Eurasia (Sunquist & Sunquist, 2002). The Eurasian lynx has a vast distribution area, covering from Korea to France (where it has been reintroduced) (Breitenmoser et al., 2015). A first range-wide analysis of mitochondrial and microsatellite diversity suggested the existence of three main lineages, Asian, European, and Southern (Rueness et al., 2014), but more recent studies based on whole mitochondrial genomes and genome-wide nuclear data revealed a more complex history (Lucena-Perez et al., 2020; Lucena-Perez et al. *in prep.*). The Eurasian lynx has, at least, six mitochondrial haplogroups diverging between 97-86 and 17-15 kya, with the three most divergent lineages occurring in Balkans, Caucasus, and the rest of the distribution (Lucena-Perez et al., 2020; Lucena-Perez et al. *in prep.*). However, genome-wide nuclear variation suggest that the populations have homogenized to some extent by gene flow in Europe, resulting in two main nuclear genetic clusters: East (Asia), and West, which would include southern populations (Balkans and Caucasus) (Lucena-Perez et al., 2020; Lucena-Perez et al. *in prep.*). Genetic patterns differed among lineages: while Asian populations are quite homogeneous across a wide geographic range and show an isolation by distance pattern, western and central European ones usually show higher genetic differentiation and lower levels of genetic diversity than those in Asia (Lucena-Perez et al., 2020). Southern populations show completely opposite patterns. While Caucasus is highly diverse, Balkans show the lowest diversity reported so far for the species (Lucena-Perez, et al. *in prep.*). Differences in genetic patterns among extant Eurasian lynx populations are a consequence of their contrasting demographic trajectories.

The Eurasian lynx originated in Asia and was once distributed along Europe, where it arrived around the early Late Pleistocene (Werdelin, 1981; Kurten & Werdelin, 1984; Kahlke, 1999). Western and central European populations progressively disappeared during the last few centuries, with most populations being already extirpated by the 1950s (Kratochvil, 1968). While several authors considered that the pre-collapse, historical range of the Eurasian lynx never encompassed the Iberian Peninsula (Kratochvil, 1968), Clavero & Delibes (2013) reviewed historical records to conclude that the species had been present in the northern fringe of the Iberian Peninsula up

to, at least, the early 19th century. This interpretation of written records was later confirmed by the genetic identification in the Iberian Peninsula of ancient Eurasian lynx remains (Rodríguez-Varela et al., 2016).

Here, we analyse the whole genome sequence of an Eurasian lynx that inhabited the northern Iberian Peninsula in ancient times, ca. 2,500 ya, to gain insights into the Eurasian lynx population inhabiting the Iberian Peninsula, which was the westernmost within the historical distribution range of the species. In particular, we analyse the relationship of this sample with contemporary populations and its genomic diversity, based both on the nuclear genome and the mitogenome, with the aim of reconstructing its phylogeographic history, and assess whether genomic factors could have contributed to the extinction of this population.

MATERIALS AND METHODS

Samples

The ancient Eurasian lynx sample was collected at Sima de Pagolusieta, Gorbea, Vizcaya, Spain, and was stored as part of the collection of the ARANZADI Natural History collection. This sample was previously described in Altuna (1980), and molecularly identified as Lynx lynx by Rodriguez-Varela (2016). Accordingly, the sample was processed following current recommendations for damaged and contaminated material in a dedicated ancient DNA laboratory (Evolutionary Adaptive Genomics Group, Potsdam University, Germany). We used ~10 mg of bone powder for digestion at 37ºC during ~18h in 1 ml extraction buffer (0.25mg/ml Proteinase k in 0.45M EDTA). After digestion, DNA was extracted according to the protocol of Dabney et al. (2013), maximizing DNA recovery thanks to an increased ratio of sample to binding buffer (1:13). 20µl of DNA extract was used to construct a single - stranded indexed Illumina library following (Gansauge & Meyer, 2013). Briefly, DNA extract was treated with uracil-DNA glycosylase (UDG) and endonuclease VII, removing postmortem damage. Then, DNA was denaturalized and ligated to a biotinylated oligo at the 3'end. The immobilization using Streptavidin-coated beads allows repairing the blunt ends of the molecules and the subsequent ligation of the P5 adapter. To avoid clonality due to PCR amplification, the resulting

library was amplified doing a qPCR to evaluate the concentration and set up the number of amplification cycles accordingly. Using AccuPrime Pfx DNA polymerase (ThermoFisher) we amplified and indexed the library incorporating an index in both P5 and P7 adapters. Due to low concentration we had to reamplify the ancient library using IS5 and IS6 primers, which led to an excess of heteroduplicates. Heteroduplexes were removed in a single-cycle-PCR using 0,5ul of the amplified library in a total of 20ul (Meyer & Kircher). The indexed library was quantified using Qubit 2.0 fluorometer (ThermoFisher Scientific) and 2200 TapeStation Instrument, and pooled in an equimolar ratio for sequencing. The library was sequenced in the Swedish Museum of Natural History in an Illumina HiSeq X instrument, after doing some test runs in Illumina MiSeq sequencing platform, and checking library complexity using Preseq (http://smithlabresearch.org/software/preseq/). Contemporary data was used from previous studies (Lucena-Perez et al., 2020; Lucena-Perez et al. *in prep.*).

Quality of all libraries was analysed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). The ancient sample was mapped to 2.8 Gb Iberian lynx (L. pardinus) reference genome (Abascal et al., 2016) exactly as in Lucena-Perez et al. (2020) and Lucena-Perez et al. (*in prep.*) with the aim of assessing its structure patterns together with contemporary samples with similar data (PCA, individual-based ancestry). Also, the ancient Iberian sample together with 12 Eurasian lynx from six different populations covering most of the Eurasian lynx range were mapped to the 2.5 Gb domestic cat (*Felis catus*) reference genome v. 9.0 (GCF_000181335.3) (Buckley et al., 2020) using bwa aln 7.17 (Li & Durbin, 2009). We use this strategy to calculate heterozygosity as ancient DNA has certain features that can inflate diversity estimates (i.e. DNA damage, contamination, short reads) that are handled better using bwa-aln due to its stringency with mismatches during mapping and suitability for short reads. Before this, overlapping reads were merged using SeqPrep (https://github.com/jstjohn/SeqPrep), with default parameters. Bam files were filtered for mapping quality above 30 (Li & Durbin, 2009), and read groups were added (https://broadinstitute.github.io/picard). Then, we merged runs coming from the same individual using samtools (Li et al., 2009), marked duplicates using picard-tools (https://broadinstitute.github.io/picard), and realigned using GATK (McKenna et al., 2010).

Coverage, mapping statistics, and read length distribution was calculated using samtools (Li et al., 2009). Endogenous content was measured as the number of reads mapped to the reference excluding duplicates divided by the total number of reads sequenced. We used mapDamage (Jónsson et al., 2013) to confirm the authenticity of the ancient sample based on the excess of C to T substitutions at the end of the reads.

We also sexed the Iberian sample by calculating the ratio of depth of coverage in X chromosome to A1 autosome. This method was validated using contemporary samples of known sex.

Radiocarbon dating

2 gr of sample were sent to Centro Nacional de Aceleradores for radiocarbon dating through accelerator mass spectrometry.

Mitogenome

The mitogenome sequence of the sample was reconstructed as previously reported for contemporary samples (Lucena-Perez et al., 2020). Briefly, reads were mapped to the *L. lynx* mitochondrial reference genome, SNPs were called using Freebayes in haploid mode and a consensus sequence was generated, with sites not covered by any read or with low-quality genotypes coded as N. The calling procedure was done conjunctly with the previously analysed samples. Repetitive regions RS2 and RS3 (positions 16096-16382 and 16908-174, respectively; Sindičić et al., 2012), were excluded from the analysis.

Consensus mitogenome sequence was aligned and collapsed into distinct haplotypes using the pegas R package (Paradis, 2010). The number of segregating sites (S), haplotype diversity (Hd), nucleotide diversity (π) and the mean number of pairwise nucleotide differences (k) were calculated using the PopGenome (Pfeifer et al., 2014) and ape (Paradis, Claude & Strimmer, 2004) R packages. Phylogenetic relationships among haplotypes were inferred by constructing a median-joining haplotype network (Bandelt, Forster & Röhl, 1999) in PopART (Leigh & Bryant, 2015), and Maximum-likelihood phylogenetic trees with RAxML 8.2.11 (Stamatakis, 2014) within the Geneious Prime, under a GTR model with a gamma distribution of rates across sites and a proportion of invariant sites (GTRGAMMAI), with node support estimated with 100 non-parametric bootstrap replicates. The inferred tree was visualised and formatted with FigTree v1.4.3 (https://github.com/rambaut/figtree).

We also reanalysed the mtDNA haplotypes reported by Rodríguez-Varela (Rodríguez-Varela et al., 2015, 2016) based on two fragments of the cytochrome b gene (79+74=153 bp) and two fragments of the control region (93+91= 184 bp), along with overlapping bases in our mitogenome haplotypes. Sequence alignments were constructed and manually edited to remove gaps, and ML trees were constructed as described above for mitogenome haplotypes.

PCA & individual ancestry

Principal component analysis (PCA) and individual base admixture proportions were calculated using the data from Lucena-Perez et al. (2020) and Lucena-Perez et al. (*in prep.*), mapped using bwa-mem to Iberian lynx genome. We only considered intergenic regions, representing 61% of the nuclear genome (~1.5 Gb) to focus on neutral evolution. We also filtered out low complexity and low mappability regions previously identified in Abascal et al. (2016).

We used ANGSD (Kim et al., 2011; Li, 2011) and NGSTools/ngsPopGen/ngsCovar (Fumagalli, 2013; Fumagalli et al., 2013) to calculate principal components (PCA) directly from genotype posterior probabilities. Two different PCA were calculated with and without the Iberian sample. The PCA represented here is the contemporary PCA with a projection of the Iberian coordinates to avoid distortion in the contemporary PCA driven by ancient data. The filters used were: -uniqueOnly 1 –remove_bads 1 –only_proper_pairs 1 –baq 1 –C 50 –minMapQ 30 –minQ 20 –doCounts 1 –minInd (number of individuals in the population/2) –setMaxDepth (AVR depth for the population + (0.95*AVR depth for the population)) –setMinDepth (AVR depth for the population - (0.95*AVR depth for the population)) -skipTriallelic 1) –SNP_pval 1e-3. PCA was plotted using scatterplot3js from threejs library in R (http://bwlewis.github.io/rthreejs/).

Also, Individual-based admixture proportions based on genotype likelihoods were calculated using ANGSD (Kim et al., 2011; Li, 2011), and NGSadmix (Li, 2011; Skotte, Korneliussen & Albrechtsen, 2013) with the same filters mentioned before. The analysis was run from K=2 to K=10 to assess the sequential clustering of the populations with an increasing number of genetic groups, and replicated 10 times each run to evaluate convergence. The results were plotted using R (R Core Team, 2019).

Diversity

Diversity was measured as individual observed heterozygosity calculated using data mapped to the domestic cat genome with bwa-aln. To avoid depth bias, which is likely to affect diversity estimates, we subsampled contemporary data to ~4.6x, i.e. depth of the Iberian sample, using samtools (Li et al., 2009). We estimated heterozygosity by running ANGSD (Li, 2011; Korneliussen et al., 2013), and then realSFS for each individual (Korneliussen et al., 2013). The site frequency spectrum (SFS) was used as input for obtaining heterozygosity values in non-overlapping windows, using first ANGSD to get the post-probability estimates and then thetaStat (Fumagalli et al., 2014) with a window size of 500 kb. The filters used were: -uniqueOnly 1 -remove_bads 1 -only_proper_pairs 1 -baq 1 -C 50 -minMapQ 30 -minQ 20 -doCounts. As post-mortem damage is usually represented in the form of transitions CG to TA, and can inflate diversity estimates artificially (Hansen et al., 2001) we calculated heterozygosity relying only on transversions.

RESULTS

Ancient sample details, and sequencing

Shotgun sequencing of the ancient Iberian sample resulted in a total of 711 950 164 reads. Duplication level was 27%, and fragment size mode was 40 bp, after excluding adapters and merging the reads. Percentage of endogenous content was 28 and 32% when mapping to domestic cat and Iberian lynx reference genomes, respectively. After trimming, mapping and removing duplicates, depth of coverage was 5,6 when mapped to Iberian lynx and 4,6x when mapped to domestic cat.

The processed ancient Iberian sample was a male, as inferred from a ratio of depth of coverage between X to A1 chromosome of 0.458. Radiocarbon dating estimated the sample to be from 2570+/-30 BP. Accordingly the sample shows typical damage pattern C to T at the ends of the reads (Fig. 1).

A total of 3461 reads mapped to the reference Eurasian lynx mitogenome sequence, yielding an average depth of 24.5x (SD = 7.8x). The original sequence was 16987 bp long and included 2 ambiguous bases (Ns), whereas the sequence finally aligned after deletion of control region RS1 and RS2 sequences was 16,448 bp long and contained no Ns.

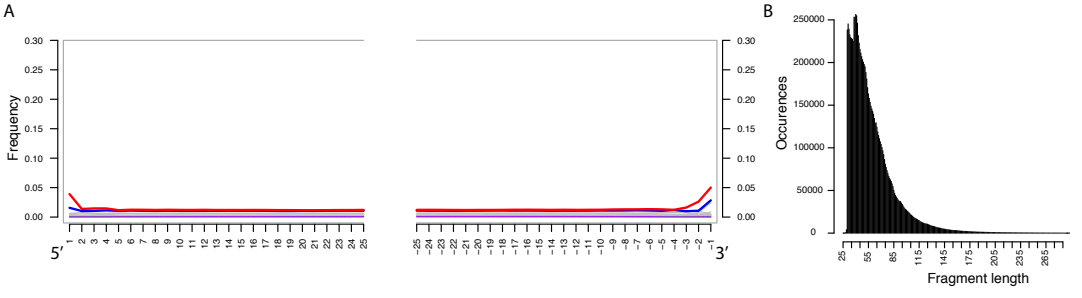


Fig. 1. Authentication of ancient DNA data using MapDamage. A) Cytosine deamination patterns. X-axis indicates the individual nucleotide position of the reads. Plots show the proportion of T where the reference genome possesses a C (red) and proportion of A where the reference possesses a G (blue). C to T changes are typical of ancient damage patterns. Plot on the right represents the fragment length distribution determined by the merging of the two reads. Minimum red length for mapping used was 30bp, resulting in a truncation of the plot.

Phylogeography and genetic status of the Iberian Eurasian lynx

Most of the ancestry of the Iberian sample belongs to the European lineage, as revealed by PCA and individual-based clustering analysis (Fig. 2). Still the uppermost supported level of structure (K=2, Evanno, Regnaut, & Goudet, 2005) supports ancestry proportions similar to those from Caucasus, with partial Asian ancestry for the Iberian sample (Fig. 2). This is confirmed by subsequent partitions in the structure analysis (K=3, K=4, K=5), in which the Iberian sample exhibits a larger proportion of the ancestry shared with individuals from Caucasus, followed by other European populations such as Kirov and Carpathians. Also, for these partitions the Iberian sample is the only one showing Asian ancestry. 3D PCA including PC1, PC2, and PC3 support an intermediate position of the Iberian sample among the European samples, slightly closer to Caucasian samples but well separated from all other extant populations.

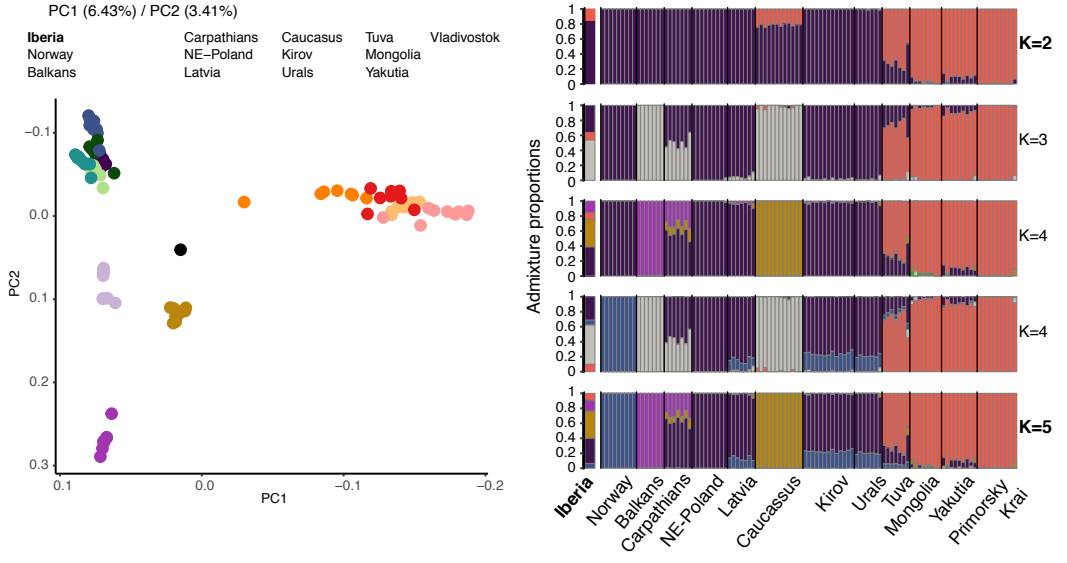


Fig. 2. Relationship among individuals based on nuclear autosomal intergenic genotypes. The ancient Iberian sample occupies a central position within European populations closer to Asian lineage. First axis (with a 6.43 % of the variance explained) separates eastern and western populations, while the second axis (with 3.41%) separates populations within Europe. Below individual ancestry proportions are determined for K=2, K=3, K=4, and K=5. Populations are sorted from west to east. When the clustering separates a population, colour code is maintained between PCA and individual ancestry proportion plot. The uppermost supported level of admixture is K=2, followed by K=5. K=4 yield two different results separating either Caucasian and Balkans populations, or Norway and Caucasian-Balkans. The Iberian individual is represented on the left and it has been widened allowing visualization.

Observed heterozygosity of the Iberian sample is the lowest among analysed samples, including those from the highly eroded Balkans population (Fig. 3). Precisely, considering only transversions, the Iberian sample observed heterozygosity is 1.1e-05, about 50% (40-60%) of that of Balkans individuals.

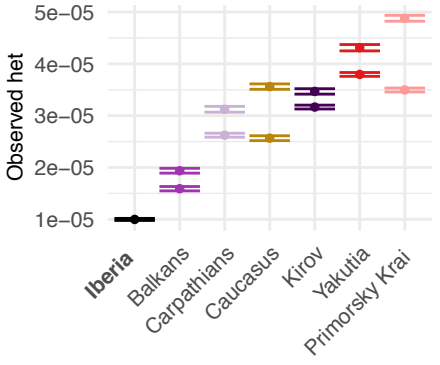


Fig. 3. Observed heterozygosity for the individuals analysed. Standard errors were calculated by bootstrapping over 500 kb windows. Populations are sorted from west to east.

Mitogenome

The alignment of the mitogenome sequence of the Iberian sample from northern Spain to previously reported sequences reveal a novel haplotype (H32), which clustered (100% bootstrap support) with haplogroup HG2, currently distributed in the Carpathians and the Baltic region (Fig. 4). H32 haplotype appears as an external short branch within the HG2, consistent with its ancient dating and likely reflecting an ancestral position with respect to the rest of the haplotypes in HG2. Such a basal and ancestral position is confirmed by the network analyses, where H32 appears in the branch leading to HG2, separated by only one mutation of the closest haplotype within this haplogroup and by 15 mutations from the closest haplotype in a different haplogroup (H9 in Northern HG3) (Fig. 4).

The analysis of cytochrome b and control region including data from Rodríguez-Varela et al. (2016; 2015) results in lower resolution due to a small number of variable sites (11 variable sites, vs. 104 in the mitogenome alignment). Major mitogenomic haplogroups receive low support or are collapsed with this reduced dataset (Fig. 5). The sequence obtained by Rodríguez-Varela for the same sample analysed here is identical to ours (H32 identical to their haplotype 1), and collapsed with our haplotypes H4-H6, H16-18, H20-24 (HG4 and HG5, only observed in Asian samples) and with H12-13 (part of our HG3, currently recorded in the Latvia and NE-Poland).

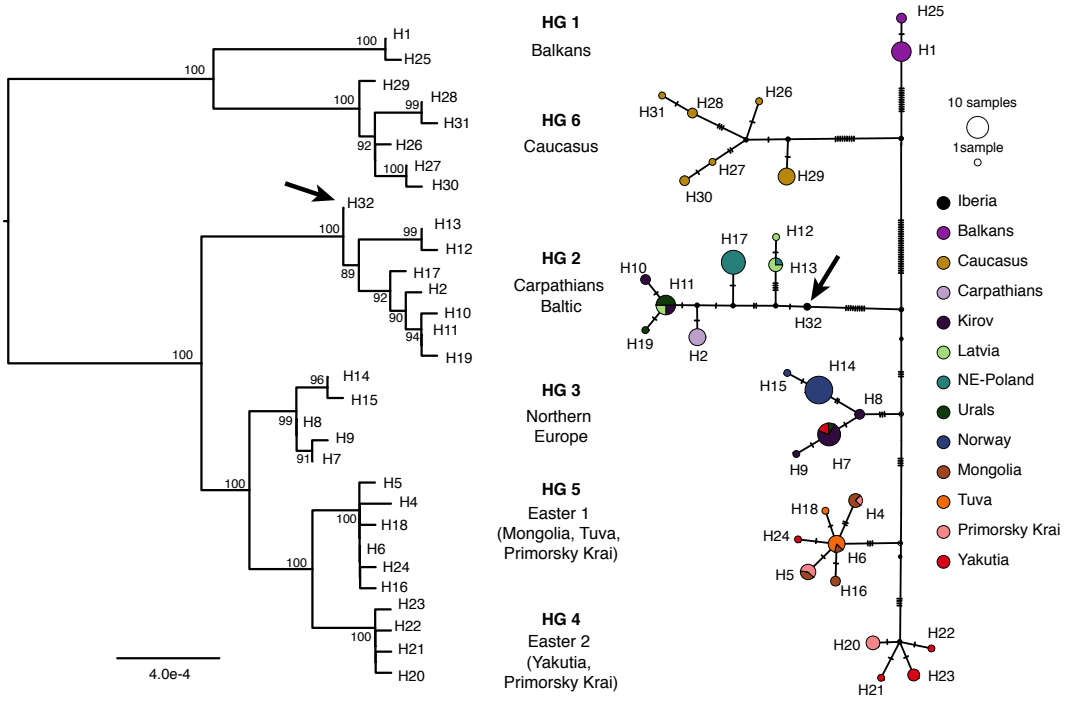


Fig. 4. ML tree depicting the phylogenetic relationships of the ancient H32 mitogenome to other extant mitogenomes reported by Lucena et al. (2020). The haplogroups inferred from whole mitogenome sequences are indicated along with their current distribution. On the right, ML phylogenetic tree inferred with RaXML. Numbers above branches indicate the bootstrap support of the corresponding node. On the left, Median-joining haplotype network reconstructed with PopART. The arrow identifies the haplotype reconstructed from the Iberian sample.

This same haplotype is retrieved from samples from Serpenteko (Navarra, Spain), Coulet des Roches (Pyrenees, France) and two localities in Denmark. Although the phylogeographic ascription of this haplotype is thus ambiguous with this reduced length dataset, this haplotype is clearly related to the Carpathian-Baltic haplogroup and not to the Eastern haplogroups, based on whole mitogenome sequences. Other ancient short haplotypes group with either our northeastern HG3, currently distributed in Kirov, Norway, Urals and Yakutia (R-V's haplotype 4 recorded in

Sueve, Asturias, Spain), or with unresolved relationships to HG3, 4 and 5 (R-V's haplotype 2, retrieved from a sample from Cueva de los Cinchos, Asturias, Spain). The rest of ancient haplotypes form a separate well supported clade clustering with, but with unresolved relationships to, Balkan and Caucasian haplogroups HG1 and HG6, respectively, suggesting an additional haplogroup closely related to other extant southern haplogroups, anciently distributed across Italy, Iberia, and Denmark.

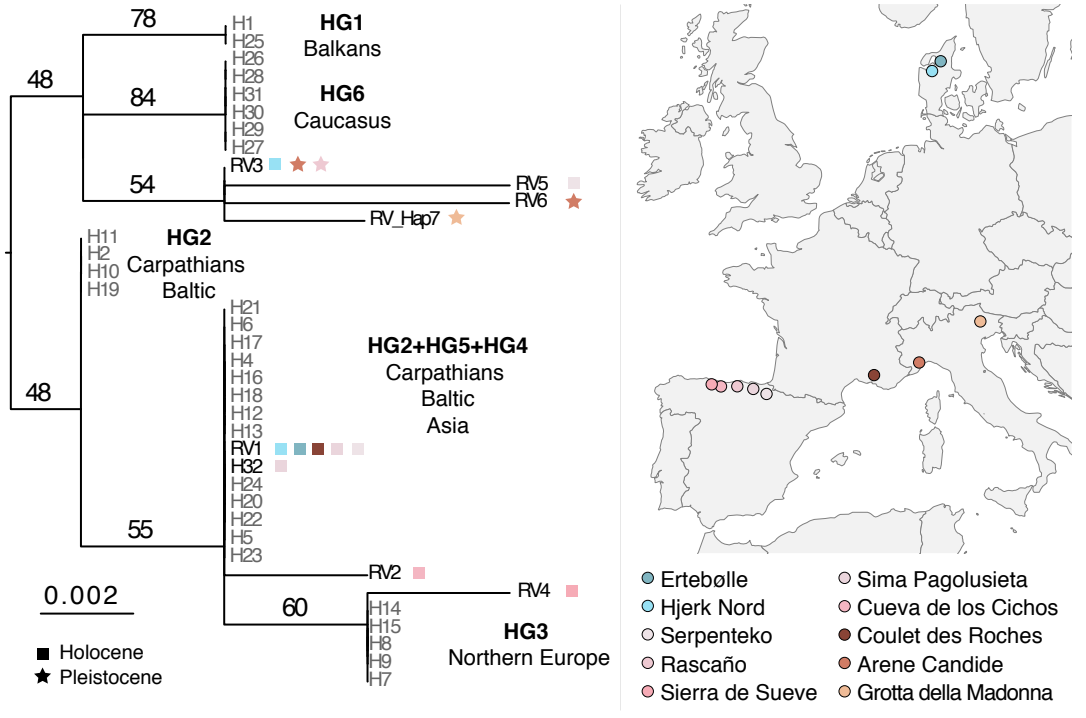


Fig. 5. Phylogenetic relationships among ancient and modern cytb+CR haplotypes. Ancient haplotypes reported by Rodríguez-Varela (identified with RV prefix) are aligned to overlapping sequences in current mitogenomes reported by Lucena et al. (2020) and the ancient mitogenome haplotype reported here. A maximum-likelihood tree was constructed with RAxML. The haplogroups inferred from whole mitogenome sequences are indicated along with their current distribution. Numbers above branches indicate the bootstrap support of the corresponding node. For ancient haplotypes the origin (colour) and period (shape) of the samples carrying the haplotype is shown (only shown once for samples with same origin and period). On the right a map showing the different sites.

DISCUSSION

The successful sequencing of an individual of the now extinct Iberian Eurasian lynx population helps to elucidate the relationships of this population with the extant populations and sheds light on some of the possible reasons of its extinction.

Eurasian lynx populations were genetically homogeneous until 100 kya, when they entered a demographic decline that promoted divergence among populations (Lucena-Perez et al., 2020; Lucena-Perez et al. *in prep.*). In Europe, Eurasian lynx arrived around the early Late Pleistocene (Werdelin, 1981; Kurten & Werdelin, 1984; Kahlke, 1999), matching the time of the emergence of different mitochondrial lineages (Lucena-Perez et al., 2020; Lucena-Perez et al. *in prep.*). The ancient Iberian mitogenome sequenced here do not reveal a new Iberian mitochondrial lineage, but rather clusters with Central European mitogenomes as part of a haplogroup currently distributed in Carpathians, Urals, Kirov, Latvia and NE-Poland (HG2, after Lucena-Perez et al. 2020), where it occupies a basal position suggestive of an ancestral status. The emergence of this haplogroup was coincident with a global climate cooling (48-54 kya; Kindler et al., 2014), and it has been suggested that it diverged in the Carpathians refugium and then spread over the central and northern part of continent (Lucena-Perez et al., 2020). Our results and previous data by Rodriguez-Varela reanalysed here suggest a wider presence of HG2 and of an extinct lineage related to modern southern haplogroups (HG1, HG6) in the Mediterranean area and across central parts of Europe, Denmark, Italy and Iberia. Lynx with southern haplotypes are generally older (11 - 28 kya; with the exception of one sample from Denmark dated in the Late Atlantic period (about 6 kya) and one from Navarra, Spain dated 0,4 kya), than those ancient lynx showing central European haplotypes (1.7 - 8 kya). The absence of records of the Central European lineage in western Europe during the Pleistocene is suggestive of a partial replacement of the southern-like lineage by the central Europe lineage after the Holocene, which may have coexisted and admixed, as suggested by modern nuclear genome data (Lucena-Perez et al., 2020). The observed pattern is also compatible with long-term coexistence of both lineages in Europe, or even lack of clear phylogeographic structure before the Holocene, as suggested for other temperate mammals (Valdiosera et al., 2008; Horn et al., 2014; Rey-Iglesia et al., 2017, 2019; Ersmark et al., 2019; García-Vázquez, Pinto Llona & Grandal-d'Anglade, 2019). It must be noted, though, that these inferences are hampered by the limited information content of the fragments analysed, and that a more complete and robust assessment of the past history of Eurasian lynx will require whole mitogenomic sequences and, ideally, genome-wide nuclear data from additional ancient and historical sequences.

Differentiation analysis based on nuclear data also supports a clustering of the Iberian Eurasian lynx individual together with other European populations. The Iberian sample shares ancestry mostly with Central and Southern populations in Europe, and to a lesser extent to Asian ones. Accordingly it occupies an intermediate position in the PCA among European populations, between Caucasus, Carpathian, other Central European populations, but also closer to Asian samples than other European ones. As no other population shows a similar ancestry pattern or clusters with the Iberian sample in the PCA, we suggest that this Iberian individual may represent a lineage with no extant representation, or a representative of a lineage ancestral to modern ones, which still shows the traces of ancient Asian ancestry.

The extremely low observed heterozygosity (1.1e-05) supports a genetic drift acting upon the population, due to either low effective populations size maintained in time, or a sharp decline in the population. This diversity is half of the one observed in the eroded Balkans population and similar to that of the extremely eroded Iberian lynx (*L. pardinus*), even though this ancient Iberian population was still not affected by the last 2 kya of human impact and land use change in the Cantabrian Strip. The genetic diversity of the ancient Iberian sample is similar to that found in the Andújar Iberian lynx population (8.9e-6 to 1.54e-05, n=18), which in turn is similar to the values reported for threatened species such as the snow leopard (Panthera uncia) or the Tasmanian devil (Sarcophilus harrisii) (Abascal et al., 2016; Lucena-Perez et al., 2020). Although this different extremely low diverse lineage could have emerged due to different processes, we think that a series of historical and anthropogenic factors could be behind this pattern.

The Iberian Peninsula represents the westernmost population of the species distribution, and it is probably one of the latest territories to be occupied by Eurasian lynx (Altuna, 1980; Sommer & Benecke, 2006). Most authors claim that the Eurasian lynx reached the Iberian Peninsula during the Pleistocene-Holocene boundary (Altuna, 1980; Sommer & Benecke, 2006), and actually, the first molecular record of the species in Iberia is dated in the very late Pleistocene, 12-11.8 kya (haplotype RV3) (Rodríguez-Varela et al., 2016). However, the fact that the Iberian lynx occupied the Iberian Peninsula at that time together with the difficulty of discriminating between the two lynx inhabiting Europe during the Late Pleistocene and the Holocene based on the morphology of the remains (Boscaini et al., 2016), claim for caution regarding the date of arrival of the Eurasian lynx in the Iberian Peninsula. Nevertheless, during this boundary, also the brown bear (*Ursus arctos*), which was apparently absent from the Iberian Peninsula from 20 kya to 10 kya, is recorded in the fossil records again (García-Vázquez, Pinto Llona & Grandal-d'Anglade, 2019). Many different explanations proposed in the literature for the absence

of bears in pre-Holocene periods (García-Vázquez, Pinto Llona & Grandal-d'Anglade, 2019), such as the glacier dynamics in the Pyrenean-Cantabrian basins (Jambrina-Enríquez et al., 2014), and the occupations of warmer areas by humans (Straus, Bicho & Winegardner, 2000), may also explain the late arrival of lynxes. Both bears and lynxes may have benefited from the expansion of boreal forest creating suitable habitat in areas of low human occupation during the Holocene (Carrión et al., 2010). Regardless of the exact date of arrival to the peninsula by Eurasian lynx, it seems likely that it followed a series of bottlenecks during westward expansion, which continued during its occupancy of the narrow Cantabrian Coast corridor. Therefore, we expect that genetic diversity was lower in this small peripheral western population. Indeed, the Eurasian lynx in the Iberian Peninsula inhabited a relatively small area that could have remained isolated from the rest of the European populations.

Anthropogenic pressures could have contributed to the isolation and decline of the Eurasian lynx in the Iberian Peninsula since the arrival of the species, increasing more and more throughout time. The remains analysed here were found at Sima de Pagolusieta in an area that has been occupied by hominids since the Middle Pleistocene, in such density that had an impact on carnivores communities inhabiting there during that period (Villaluenga et al., 2012). More recently, during the Pleistocene-Holocene boundary, human activities in the Cantabrian strip intensified, producing an overhunting on potential lynx prey such as the red deer (*Cervus elaphus*), Iberian Ibex (*Capra pyrenaica*) and roe deer (*Capreolus capreolus*) (Marín-Arroyo, 2013). Human impact strengthened after the introduction of agriculture and farming during the initial phases of the Neolithic, around 5 kya in the region (Peña-Chocarro et al., 2005; Cubas et al., 2016, 2020; Pérez-Díaz et al., 2018). Lynx habitat, i.e. forest, suffer mayor regression due to human activities in the north of the Iberian Peninsula starting around 6 kya with an intensification at the transition from Roman to Germanic period around 2 kya (Kaal et al., 2011). This trend continued during middle and modern ages, when forest were cleared for shipbuilding, and also summer grasslands exploitation after the introduction of large distance transhumance (Valbuena-Carabaña et al., 2010; García-Ruiz et al., 2020). Also, Eurasian lynx in the peninsula have suffered from direct persecution due to conflicts with cattle (Clavero & Delibes, 2013). More recently, the introduction of the industry in the area and the subsequent urbanization and migration from other areas, could have finally led to the extinction of the remaining lynx populations (Vilallonga, 1998).

In summary, the use of ancient genomics provided insights into the evolutionary history of the Eurasian lynx in Europe. Our work suggests that a combination of human direct

(prosecution) and indirect impacts (on habitats and prey communities), together with historical factors (i.e. little habitat availability, small population size, low genetic diversity), enhanced the genetic impoverishment of the population, and could have contributed to precipitate the extinction of the westernmost population of Eurasian lynx.

AUTHOR CONTRIBUTION

J.A.G. and M.L.P. conceived the project and designed the study; M.L.P. performed the lab work under J.P. supervision; M.L.P., E.B. and D.K. and J.A.G. analysed data; J.A.G., and M.L.P. interpreted the results; M.L.P. drafted the manuscript with support from J.A.G., M.C., and M.D; J.A.G. supervised the project. All authors approved the final version of the manuscript.

ACKNOWLEDGEMENTS

This project was funded by the Dirección General de Investigación Científica y Técnica of the Spanish Ministerio de Ciencia e Innovacion (MICINN) through projects CGL2013-47755-P, CGL2017-84641-P and CGL2014-55266-P, and by the Fundacion BBVA through a project on ancient lynx genetics to M.D. MLP was supported by PhD contracts from the Programa Internacional de Becas "La Caixa-Severo Ochoa" of the "La Caixa" Foundation (ID 100010434), under agreement LCF/BQ/SO14/52250035. Also, María Lucena-Perez received funding from the European Science Foundation Reseach Networking Programmes - ConGenOmics for a Exchange Visit Grant. The Sociedad de Ciencias Aranzadi (San Sebastián, Spain) preserved and kindly provided the ancient Eurasian lynx sample. Radicarbon dating was performed by Centro Nacional de Aceleradores in Seville, Spain. Calculations were carried out in the Genomics servers of Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD) and in Fundación Pública Galega Centro Tecnolóxico de Supercomputación de Galicia (CESGA). EBD-CSIC received support from the Spanish Ministry of Economy and Competitiveness under the 'Centro de Excelencia Severo Ochoa 2013-2017' program, grants SEV-2012-0262. Besides the free software mentioned in the text, this work was possible thanks to the free, open source packages: dplyr (Wickham, Francois, Henry, & Müller, 2017), ggplot2 (Wickham, 2009), tidyr (Wickham & Henry, 2018), broom (D. Robinson, 2018), RColorBrewer (Neuwirth, 2014), viridis (Garnier, 2018), stringr (Wickham & Henry, 2018), knitr (Xie, 2014, 2015, 2017), reshape2 (Wickham, 2007), data.table (Dowle & Srinivasan, 2017), rlang (Henry & Wickham, 2018).

02

APÉNDICE

ANCIENT GENOME PROVIDES INSIGHTS INTO THE HISTORY OF EURASIAN LYNX IN IBERIA AND WESTERN EUROPE

Table S1. Detailed information about the *L. lynx* samples analysed in this study.

Sample	Dataset	Population	Assigned date
a_ll_pv_0223	ancient	Basque Country	2,570+/-30BP
c_ll_ba_0224	contemporary	Balkans region	2017
c_ll_ba_0226	contemporary	Balkans region	2017
c_ll_ba_0227	contemporary	Balkans region	2018
c_ll_ba_0228	contemporary	Balkans region	2012
c_ll_ba_0229	contemporary	Balkans region	2012
c_ll_ba_0230	contemporary	Balkans region	2016
c_ll_ba_0233	contemporary	Balkans region	2019
c_ll_ca_0240	contemporary	Caucasus	2016
c_ll_ca_0241	contemporary	Caucasus	2019
c_ll_ca_0242	contemporary	Caucasus	2019
c_ll_ca_0243	contemporary	Caucasus	2019
c_ll_ca_0244	contemporary	Caucasus	2012
c_ll_ca_0245	contemporary	Caucasus	2010
c_ll_ca_0247	contemporary	Caucasus	2016
c_ll_ca_0248	contemporary	Caucasus	2012
c_ll_ca_0252	contemporary	Caucasus	2019
c_ll_ca_0254	contemporary	Caucasus	2016
c_ll_ca_0259	contemporary	Caucasus	2008
c_ll_ca_0260	contemporary	Caucasus	2018
c_ll_cr_0205	contemporary	Carpathian montains	2007
c_ll_cr_0206	contemporary	Carpathian montains	2007
c_ll_cr_0207	contemporary	Carpathian montains	2007
c_ll_cr_0208	contemporary	Carpathian montains	2010
c_ll_cr_0209	contemporary	Carpathian montains	2011
c_ll_cr_0211	contemporary	Carpathian montains	2011
c_ll_cr_0212	contemporary	Carpathian montains	2012
c_ll_ka_0184	contemporary	Mongolia (Central/Khentii aymag)	2016
c_ll_ka_0186	contemporary	Mongolia (Central/Khentii aymag)	2015
c_ll_ka_0188	contemporary	Mongolia (Central/Khentii aymag)	2015
c_ll_ka_0189	contemporary	Mongolia (Central/Khentii aymag)	2015
c_ll_ki_0086	contemporary	Kirov region	NA
c_ll_ki_0088	contemporary	Kirov region	NA
c_ll_ki_0090	contemporary	Kirov region	2011
c_ll_ki_0091	contemporary	Kirov region	2011
c_ll_ki_0092	contemporary	Kirov region	2011
c_ll_ki_0093	contemporary	Kirov region	2011
c_ll_ki_0094	contemporary	Kirov region	2011
c_ll_ki_0095	contemporary	Kirov region	2011
c_ll_ki_0096	contemporary	Kirov region	2011
c_ll_ki_0097	contemporary	Kirov region	2011

Reference	Sex	Latitude	Longitud	Mitogenome	Nuclear WG	Original depth Lynx pardinus	Original depth Felis catus
This study	M	-2,7800	43,1006	Yes	Yes	5,6	4,62
Lucena-Perez et al. in prep	F	20,7739	41,4805	Yes	Yes	31,9	NA
Lucena-Perez et al. in prep	F	20,9602	41,5038	Yes	Yes	15,0	NA
Lucena-Perez et al. in prep	F	20,8176	41,3752	Yes	Yes	16,5	8,97
Lucena-Perez et al. in prep	M	20,9369	41,6342	Yes	Yes	26,8	NA
Lucena-Perez et al. in prep	M	20,7937	41,4794	Yes	Yes	21,1	11,27
Lucena-Perez et al. in prep	M	20,7937	41,4794	Yes	Yes	19,8	NA
Lucena-Perez et al. in prep	M	NA	NA	Yes	Yes	21,0	NA
Lucena-Perez et al. in prep	M	42,4921	41,4559	Yes	Yes	35,7	NA
Lucena-Perez et al. in prep	F	46,3772	41,8718	Yes	Yes	22,2	12,09
Lucena-Perez et al. in prep	F	44,9274	40,0347	Yes	Yes	22,4	12,21
Lucena-Perez et al. in prep	M	47,8371	41,5720	Yes	Yes	26,9	NA
Lucena-Perez et al. in prep	F	47,8938	41,7332	Yes	Yes	18,9	NA
Lucena-Perez et al. in prep	F	46,6723	42,3942	Yes	Yes	9,2	NA
Lucena-Perez et al. in prep	M	44,3492	41,6888	Yes	Yes	13,1	NA
Lucena-Perez et al. in prep	F	42,8088	43,0395	Yes	Yes	11,4	NA
Lucena-Perez et al. in prep	F	43,2502	43,1969	Yes	Yes	17,3	NA
Lucena-Perez et al. in prep	M	42,0562	41,5716	Yes	Yes	14,7	NA
Lucena-Perez et al. in prep	M	46,3864	42,6917	Yes	Yes	12,4	NA
Lucena-Perez et al. in prep	M	46,4468	42,1021	Yes	Yes	15,7	NA
Lucena-Perez et al. 2020	F	19,7585	49,6714	Yes	Yes	8,0	4,96
Lucena-Perez et al. 2020	M	25,5994	45,6407	Yes	Yes	7,5	NA
Lucena-Perez et al. 2020	F	22,2012	45,3975	Yes	Yes	8,7	NA
Lucena-Perez et al. 2020	F	22,2989	49,7508	Yes	Yes	7,6	NA
Lucena-Perez et al. 2020	F	22,2012	45,3975	Yes	Yes	8,6	NA
Lucena-Perez et al. in prep	M	26,9146	46,5670	No	Yes	23,8	NA
Lucena-Perez et al. 2020	M	22,4348	49,7078	Yes	Yes	19,2	9,94
Lucena-Perez et al. 2020	M	108,6364	48,3786	Yes	Yes	7,0	NA
Lucena-Perez et al. 2020	F	108,6364	48,3786	Yes	Yes	6,9	NA
Lucena-Perez et al. 2020	M	110,4864	48,6225	Yes	Yes	6,0	NA
Lucena-Perez et al. 2020	M	110,5047	49,6247	Yes	Yes	5,9	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	F	50,3964	59,8147	Yes	Yes	21,9	11,43
Lucena-Perez et al. 2020	M	50,3964	59,8147	Yes	Yes	6,1	4,05
Lucena-Perez et al. 2020	F	50,3964	59,8147	Yes	Yes	5,3	NA
Lucena-Perez et al. 2020	M	50,3964	59,8147	Yes	Yes	5,3	NA
Lucena-Perez et al. 2020	M	50,3964	59,8147	Yes	Yes	5,8	NA
Lucena-Perez et al. 2020	M	48,4164	60,8122	Yes	Yes	5,3	NA
Lucena-Perez et al. 2020	M	48,4164	60,8122	Yes	Yes	6,1	NA
Lucena-Perez et al. 2020	F	48,4164	60,8122	Yes	Yes	6,0	NA

c_ll_ki_0098	contemporary	Kirov region	2011
c_ll_ki_0099	contemporary	Kirov region	2011
c_ll_ki_0100	contemporary	Kirov region	2011
c_ll_ki_0101	contemporary	Kirov region	2011
c_ll_ki_0102	contemporary	Kirov region	2011
c_ll_la_0044	contemporary	Latvia	2010
c_ll_la_0045	contemporary	Latvia	2009
c_ll_la_0047	contemporary	Latvia	2009
c_ll_la_0048	contemporary	Latvia	2010
c_ll_la_0052	contemporary	Latvia	2010
c_ll_la_0053	contemporary	Latvia	2010
c_ll_la_0054	contemporary	Latvia	2010
c_ll_no_0063	contemporary	Norway	NA
c_ll_no_0064	contemporary	Norway	NA
c_ll_no_0065	contemporary	Norway	2007
c_ll_no_0065	contemporary	Norway	2007
c_ll_no_0068	contemporary	Norway	NA
c_ll_no_0069	contemporary	Norway	NA
c_ll_no_0070	contemporary	Norway	NA
c_ll_no_0071	contemporary	Norway	NA
c_ll_no_0072	contemporary	Norway	NA
c_ll_no_0074	contemporary	Norway	NA
c_ll_no_0075	contemporary	Norway	2008
c_ll_no_0076	contemporary	Norway	2008
c_ll_no_0077	contemporary	Norway	2008
c_ll_no_0078	contemporary	Norway	2008
c_ll_no_0079	contemporary	Norway	2008
c_ll_no_0080	contemporary	Norway	2008
c_ll_no_0081	contemporary	Norway	2008
c_ll_no_0082	contemporary	Norway	2008
c_ll_og_0181	contemporary	Mongolia (Omnogovi)	2015
c_ll_og_0187	contemporary	Mongolia (Omnogovi)	2015
c_ll_po_0001	contemporary	Białowieża Primeval Forest	2001
c_ll_po_0002	contemporary	Białowieża Primeval Forest	1994
c_ll_po_0003	contemporary	Białowieża Primeval Forest	2002
c_ll_po_0004	contemporary	Białowieża Primeval Forest	NA
c_ll_po_0005	contemporary	Białowieża Primeval Forest	NA
c_ll_po_0007	contemporary	Białowieża Primeval Forest	NA
c_ll_po_0010	contemporary	Białowieża Primeval Forest	NA
c_ll_po_0011	contemporary	Białowieża Primeval Forest	2007
c_ll_po_0014	contemporary	Białowieża Primeval Forest	2008
c_ll_po_0017	contemporary	Białowieża Primeval Forest	NA
c_ll_po_0019	contemporary	Białowieża Primeval Forest	2010
c_ll_po_0105	contemporary	Knyszyn Primeval Forests	2004
c_ll_po_0106	contemporary	Knyszyn Primeval Forests	2006
c_ll_po_0150	contemporary	Białowieża Primeval Forest	2015

Lucena-Perez et al. 2020	F	48,4164	60,8122	Yes	Yes	5,8	NA
Lucena-Perez et al. 2020	F	48,4164	60,8122	Yes	Yes	6,0	NA
Lucena-Perez et al. 2020	M	46,7272	61,1939	Yes	Yes	6,2	NA
Lucena-Perez et al. 2020	F	46,7272	61,1939	Yes	Yes	6,3	NA
Lucena-Perez et al. 2020	M	46,7272	61,1939	Yes	Yes	6,1	NA
Lucena-Perez et al. 2020	M	22,5279	56,6426	Yes	Yes	7,7	NA
Lucena-Perez et al. in prep	F	27,5770	56,3144	No	Yes	24,3	NA
Lucena-Perez et al. 2020	F	28,1454	56,3767	Yes	Yes	6,9	NA
Lucena-Perez et al. 2020	F	22,8189	57,3586	Yes	Yes	5,3	NA
Lucena-Perez et al. 2020	M	22,5967	57,3506	Yes	Yes	6,8	NA
Lucena-Perez et al. 2020	F	22,8168	57,0459	Yes	Yes	7,8	NA
Lucena-Perez et al. 2020	M	25,5718	57,6727	Yes	Yes	8,2	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	F	21,3435	69,5098	Yes	No	23,7	NA
Lucena-Perez et al. in prep	F	21,3435	69,5098	Yes	Yes	23,7	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	M	12,3611	64,5075	Yes	Yes	5,4	NA
Lucena-Perez et al. 2020	M	10,6792	63,8246	Yes	Yes	5,8	NA
Lucena-Perez et al. 2020	M	10,9543	63,5673	Yes	Yes	5,1	NA
Lucena-Perez et al. 2020	M	12,1471	61,3841	Yes	Yes	5,3	NA
Lucena-Perez et al. 2020	M	8,4118	59,3987	Yes	Yes	5,1	NA
Lucena-Perez et al. 2020	F	8,7274	60,6649	Yes	Yes	5,2	NA
Lucena-Perez et al. 2020	F	8,9674	60,2747	Yes	Yes	5,5	NA
Lucena-Perez et al. 2020	M	9,0288	60,1445	Yes	Yes	5,4	NA
Lucena-Perez et al. 2020	M	101,0478	43,2250	Yes	Yes	6,8	NA
Lucena-Perez et al. 2020	F	103,7781	48,6367	Yes	Yes	5,6	NA
Lucena-Perez et al. 2020	M	23,5955	52,9544	Yes	Yes	6,5	NA
Lucena-Perez et al. 2020	F	23,8431	52,8230	Yes	Yes	6,6	NA
Lucena-Perez et al. 2020	F	23,7202	52,7968	Yes	Yes	6,4	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	M	23,5143	52,8918	Yes	Yes	6,4	NA
Lucena-Perez et al. 2020	M	23,9455	52,7685	Yes	Yes	6,4	NA
Lucena-Perez et al. 2020	NA	NA	NA	Yes	No	NA	NA
Lucena-Perez et al. 2020	M	23,7435	52,6807	Yes	Yes	6,4	NA
Lucena-Perez et al. 2020	M	23,4738	53,0907	Yes	Yes	6,1	NA
Lucena-Perez et al. 2020	F	23,6591	53,1045	Yes	Yes	6,1	NA
Lucena-Perez et al. in prep	M	23,6017	52,7633	No	Yes	24,0	NA

c_ll_to_0190	contemporary	Mongolia (Central/Khentii aymag)	2016
c_ll_to_0191	contemporary	Mongolia (Central/Khentii aymag)	2015
c_ll_tu_0153	contemporary	Tuva Republic	2015
c_ll_tu_0154	contemporary	Tuva Republic	2015
c_ll_tu_0157	contemporary	Tuva Republic	2015
c_ll_tu_0158	contemporary	Tuva Republic	2015
c_ll_tu_0159	contemporary	Tuva Republic	2015
c_ll_tu_0165	contemporary	Tuva Republic	2014
c_ll_tu_0166	contemporary	Tuva Republic	2014
c_ll_ur_0194	contemporary	Ural Mountains	2016
c_ll_ur_0195	contemporary	Ural Mountains	2016
c_ll_ur_0196	contemporary	Ural Mountains	2016
c_ll_ur_0199	contemporary	Ural Mountains	2016
c_ll_ur_0200	contemporary	Ural Mountains	2016
c_ll_ur_0202	contemporary	Ural Mountains	2016
c_ll_ur_0203	contemporary	Ural Mountains	2016
c_ll_vl_0107	contemporary	Primorsky Krai	2010
c_ll_vl_0108	contemporary	Primorsky Krai	2013
c_ll_vl_0109	contemporary	Primorsky Krai	2013
c_ll_vl_0110	contemporary	Primorsky Krai	2014
c_ll_vl_0112	contemporary	Primorsky Krai	2014
c_ll_vl_0113	contemporary	Primorsky Krai	2015
c_ll_vl_0114	contemporary	Primorsky Krai	2014
c_ll_vl_0128	contemporary	Primorsky Krai	2014
c_ll_vl_0132	contemporary	Primorsky Krai	2014
c_ll_vl_0137	contemporary	Primorsky Krai	2014
c_ll_ya_0138	contemporary	Yakutia Republic	2015
c_ll_ya_0139	contemporary	Yakutia Republic	2015
c_ll_ya_0140	contemporary	Yakutia Republic	2015
c_ll_ya_0141	contemporary	Yakutia Republic	2015
c_ll_ya_0142	contemporary	Yakutia Republic	2015
c_ll_ya_0143	contemporary	Yakutia Republic	2015
c_ll_ya_0145	contemporary	Yakutia Republic	2015
c_ll_ya_0146	contemporary	Yakutia Republic	2015
c_ll_ya_0147	contemporary	Yakutia Republic	2015

Lucena-Perez et al. 2020	F	108,2731	47,3986	Yes	Yes	7,3	NA
Lucena-Perez et al. 2020	M	108,6836	47,6300	Yes	Yes	7,7	NA
Lucena-Perez et al. 2020	F	96,5333	51,4833	Yes	Yes	8,1	NA
Lucena-Perez et al. in prep	F	96,5333	51,4833	No	Yes	22,4	NA
Lucena-Perez et al. 2020	M	96,5333	51,4833	Yes	Yes	7,8	NA
Lucena-Perez et al. 2020	F	96,5333	51,4833	Yes	Yes	8,0	NA
Lucena-Perez et al. 2020	F	96,5333	51,4833	Yes	Yes	7,7	NA
Lucena-Perez et al. 2020	F	96,5333	51,4833	Yes	Yes	7,8	NA
Lucena-Perez et al. 2020	M	96,5333	51,4833	Yes	Yes	8,1	NA
Lucena-Perez et al. 2020	F	59,6425	56,0481	Yes	Yes	11,0	NA
Lucena-Perez et al. 2020	M	60,1306	55,1494	Yes	Yes	11,8	NA
Lucena-Perez et al. 2020	M	60,1306	55,1494	Yes	Yes	12,4	NA
Lucena-Perez et al. 2020	M	59,7886	55,1817	Yes	Yes	13,2	NA
Lucena-Perez et al. 2020	M	59,0006	55,0006	Yes	Yes	12,7	NA
Lucena-Perez et al. in prep	F	60,8922	54,3800	No	Yes	23,7	NA
Lucena-Perez et al. 2020	F	57,3625	54,9994	Yes	Yes	13,4	NA
Lucena-Perez et al. 2020	F	136,4727	44,9357	Yes	Yes	6,5	NA
Lucena-Perez et al. 2020	F	135,5792	45,2938	Yes	Yes	11,9	NA
Lucena-Perez et al. 2020	M	136,7313	45,6667	Yes	Yes	6,0	NA
Lucena-Perez et al. 2020	M	132,9141	49,0148	Yes	Yes	7,1	NA
Lucena-Perez et al. 2020	F	137,0093	45,8572	Yes	Yes	28,9	14,10
Lucena-Perez et al. 2020	M	137,0093	45,8572	Yes	Yes	8,4	5,24
Lucena-Perez et al. in prep	F	137,6066	47,2945	No	Yes	23,8	NA
Lucena-Perez et al. 2020	M	137,4308	45,9075	Yes	Yes	8,0	NA
Lucena-Perez et al. 2020	M	137,4308	45,9075	Yes	Yes	8,3	NA
Lucena-Perez et al. in prep	F	133,4556	51,3910	No	Yes	25,8	NA
Lucena-Perez et al. 2020	M	132,0694	59,9010	Yes	Yes	7,9	4,82
Lucena-Perez et al. 2020	M	129,2049	61,1892	Yes	Yes	8,0	NA
Lucena-Perez et al. 2020	M	129,1618	61,1594	Yes	Yes	8,2	NA
Lucena-Perez et al. in prep	F	136,1769	66,8881	No	Yes	23,9	NA
Lucena-Perez et al. 2020	M	136,1769	66,8881	Yes	Yes	8,3	NA
Lucena-Perez et al. 2020	M	136,1769	66,8881	Yes	Yes	7,9	NA
Lucena-Perez et al. 2020	M	129,4453	61,8933	Yes	Yes	8,3	NA
Lucena-Perez et al. 2020	M	130,6791	60,7762	Yes	Yes	22,9	12,32
Lucena-Perez et al. 2020	F	127,3045	60,7515	Yes	Yes	8,1	NA

Table S2. Observed heterozygosity of all individuals analysed and standard deviation calculated over 500 Kb windows.

Sample	Populations	Observed heterozygosity	Standard_deviation
a_ll_pv_0223	Ancient_Basque_Country	9,97E-06	1,96E-07
c_ll_ba_0227	Balkans	1,59E-05	4,22E-07
c_ll_ba_0229	Balkans	1,94E-05	4,65E-07
c_ll_ca_0241	Caucasus	2,57E-05	4,67E-07
c_ll_ca_0242	Caucasus	3,56E-05	5,23E-07
c_ll_cr_0205	Carpathians	3,12E-05	5,61E-07
c_ll_cr_0212	Carpathians	2,62E-05	3,79E-07
c_ll_ki_0090	Kirov	3,17E-05	3,74E-07
c_ll_ki_0091	Kirov	3,47E-05	5,23E-07
c_ll_vl_0112	Primorsky_Krai	3,50E-05	3,78E-07
c_ll_vl_0113	Primorsky_Krai	4,88E-05	5,68E-07
c_ll_ya_0138	Yakutia	4,31E-05	6,12E-07
c_ll_ya_0146	Yakutia	3,80E-05	3,73E-07

03 CAPÍTULO

GENETIC RESTORATION AFTER INTERSPECIFIC INTROGRESSION IN AN ENDANGERED FELID

Maria Lucena-Perez

Johanna Paijmans, Francisco Nocete, Jordi Nadal, Cleia Detry, Love Dalen, Michael Hofreiter, Axel Barlow, José A. Godoy.

ABSTRACT

The study of ancient DNA can greatly improve the reconstruction of the evolutionary history of a species or population by providing a snapshot of the past that may reveal patterns and processes that might remain undetected with modern samples. Here, we use ancient whole genome data to characterize the Iberian lynx (*Lynx pardinus*) population inhabiting the Iberian Peninsula 2-4 kya (n=3), and we compare it with the two populations remaining by 2000, Andújar (n=18) and Doñana (n=12). The ancient population, which is relatively homogeneous and closer to Andújar population, harboured lower diversity than the highly eroded contemporary populations, being the lowest diversity ever reported for a mammal. The increase in diversity in the species is concomitant with an inferred gene flow event with European Eurasian lynx that is equally recorded in the two remnant Iberian lynx populations. The signal of introgression into Iberian lynx is higher from current western than eastern Eurasian lynx populations, but is similar across the former. A 2 kyo Eurasian lynx from the Iberian Peninsula showed less shared allele with current Iberian lynx than modern European Eurasian lynx populations, indicating that this particular Eurasian lynx lineage was not involved in the recent admixture process. Also, we found an increasing signal of introgression through time, with more recent individuals showing more shared alleles with Eurasian lynx. The admixture with Eurasian lynx could have increased the extremely low genetic diversity of the species, but with unknown consequences on fitness and demography.

INTRODUCTION

Evolutionary, demographic, and genetic processes in natural populations are usually inferred from the analysis of contemporary genetic data. Traditionally, methods for monitoring genetic changes relied on neutral genetic markers, mostly microsatellites, and mitochondrial DNA, and focused on comparing the extant genetic status among closely-related populations and/or species (e.g. Casas-Marce, Soriano, López-Bao, & Godoy, 2013). This approach has two main limitations. First, patterns in a few neutral markers (e.g. microsatellites) do not necessarily reflect genome-wide patterns, and especially those of the most relevant functional variation. Recent developments on next generation sequencing technics have partially overcome this limitation by making the genome sequencing of non-model species more affordable, enabling more accurate estimates of genome-wide diversity as well as the assessment of adaptive and deleterious variation (Murchison et al., 2012; Li et al., 2014; Abascal et al., 2016; Feng et al., 2019). Second, the evaluation of one target population in comparison with a closely related population or species provide some inferences about the evolutionary processes that could have occurred in the population of interest, but does not provide direct evidences that the process inferred produced the pattern observed. Unnoticed demographic common history, or unknown differences in post-split demographic trends could lead to partial or inaccurate inferences. To overcome this limitation, the study of historical and ancient DNA provides a better approach for reconstructing the evolutionary process by providing comparison of the same population through time (Leonard, 2008; Navascués, Depaulis & Emerson, 2010; Díez-del-Molino et al., 2017; Leonardi et al., 2017). This approach has been extensively used in humans (e.g. Nielsen et al., 2017; Racimo, Sankararaman, Nielsen, & Huerta-Sánchez, 2015), and also in some non-model extant and extinct species, such as bears (e.g. Barlow et al., 2018; Valdiosera et al., 2007), wolfs (e.g. Skoglund, Ersmark, Palkopoulou, & Dalen, 2015), mammoths (e.g. Palkopoulou et al., 2018; Palkopoulou, Mallick, Skoglund, Enk, Rohland, Li, Omrak, Vartanyan, Poinar, Götherström, et al., 2015; Rogers & Slatkin, 2017), hyena (e.g. Westbury et al., 2020), horses (e.g. Fages et al., 2019; Gaunitz et al., 2018; Librado et al., 2015), and lynx (e.g. Casas-Marce et al., 2017; Rodríguez-Varela et al., 2015, 2016). In many of these examples, the analysis of ancient DNA data allowed the calibration of the loss of genetic diversity through time and the evaluation of the relative contribution of human impacts or climate change to the demographic decline or extinction of the species. However, even though there is a tendency to assume that ancient populations harbour higher diversity than contemporary ones, there are a number of processes that could lead to the opposite pattern. For instance, events of admixture or gene flow with closely

related populations or species may lead to a global increase in genetic diversity, even in populations undergoing dramatic demographic declines. Indeed, the restoration of gene flow has been widely proposed and eventually implemented to increase genetic diversity in highly eroded isolated populations, with the expectation that this will also result in increased fitness and improved demographic trends (i.e. a proper genetic rescue).

The Iberian lynx (*Lynx pardinus*) was recognized as the most endangered felid in the word (Nowell & Jackson, 1996), and was classified by the IUCN as critically endangered in 2002, 2006, 2008 red lists, until 2015, when thanks to conservation actions the species was downlisted to endangered (Rodríguez & Calzada, 2015). The availability of resources, such as ancient, historical and modern samples and an annotated genome (Abascal et al., 2016; Casas-Marce et al., 2017), and the profound knowledge of the species accumulated during decades of studies make of this species an outstanding model to study the consequences of a demographic declines in the genome. In fact, a previous work has reconstructed the loss of mitogenomic -using ancient and historical data-, and microsatellite diversity -using historical data-, and the increase in genetic differentiation of the populations that accompanied the decline and isolation of populations (Casas-Marcé et al, 2017). However, methodological challenges and costs associated with sequencing of ancient (subfossil), or historical (museum) material prevented then the analysis of whole genome sequences.

Here, we report an unexpected increase in genetic diversity thought time in the Iberian lynx by using whole genome data from three ancient individuals that inhabited the Iberian Peninsula from 2 to 4 kya, compared to whole genome data from contemporary populations, and relate this increase with an admixture event with its sister and species, the Eurasian lynx (*Lynx lynx*).

MATERIALS AND METHODS

Samples

Ancient DNA data was obtained from Iberian lynx archeological and paleontological remains (Table S1). We tested 20 ancient samples selected based on sequencing performance of the 58 samples analysed in Casas-Marce et al. (2017). Data from modern samples was obtained from Lucena-Perez et al. submitted. For this project we used 30 contemporary Iberian lynx, from two different populations, Andújar (n=18) and Doñana (n=12), 12 Eurasian lynx from six different populations distributed in a west-east gradient (Balkans, Carpathians, Kirov, Caucasus, Yakutia and Primorsky Krai, n=2 for every population, Table S2), and one bobcat to be used as outgroup for ancestral state identification and admixture analysis.

Laboratory methods

All laboratory work was carried out in dedicated ancient DNA facilities at the University of Potsdam, following established procedures to prevent contamination with modern or synthetic DNA (Fulton, 2012). We performed 70 DNA extractions from 20 archeological samples, selected based on mitogenome sequence yield (Casas-Marcé et al, 2017). Negative controls were included in all experiments. We drilled the remains to get a variable amount of bone powder up to a maximum of 50 mg. For Algarve and Catalonia sample extraction, we pre-treated the samples with 1 ml 0.5% bleach for 15 min at room temperature to minimize DNA contamination (Korlević et al., 2015). Bone powder was digested for ~18 h at 37°C in 1 ml extraction buffer (0.45M EDTA, 0.25mg/ml Proteinase K) with constant rotation. DNA in the resulting supernatant was purified following Dabney et al. (2013), which maximize degraded DNA recovery by increasing the ratio of binding buffer to sample (13:1).

DNA extracts were converted into sequencing libraries using the single-stranded protocol described in (Gansauge & Meyer, 2013). Prior to library preparation DNA extracts were treated with UDG (uracil-DNA glycosylase) and endonuclease VIII to remove deoxyuracils resulting from postmortem DNA damage. Then double-stranded DNA was denatured and a biotinylated adapter oligo was ligated to the 3'end of each molecule. Resulting products were immobilized on Streptavidin-coated beads to repair blunt ends, and the P5 adapter was ligated to the template molecule. We performed a qPCR to evaluate library concentration and used this information to estimate the minimal number of amplification cycles to use in the indexing PCR to avoid clonality. Indexing of single-stranded libraries incorporated an index sequence next to the P7 and P5 adapter using PCR with AccuPrime Pfx DNA polymerase (ThermoFisher) for its ability to read over uracil (Seguin-Orlando et al., 2015). Indexed libraries were quantified using Qubit 2.0 fluorometer (ThermoFisher Scientific) and 2200 TapeStation Instrument, and pooled in an equimolar ratio.

A test sequencing to evaluate endogenous content was done using Illumina MiSeq sequencing platform, producing 2x70pb paired-end (PE) reads. Out of the 70 libraries prepared, 4 libraries, coming from 3 samples (Fig. 1A), showed an endogenous content above 10%. These libraries were checked for complexity using Preseq software (http://smithlabresearch.org/software/preseq/), before major sequencing using Illumina HiSeq X sequencing platform in the Swedish Museum of Natural History.

Data processing

Contemporary and ancient data were processed following the same pipeline. Whole genome re-sequencing data was quality-checked using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Overlapping reads were merged using SeqPrep (https://github.com/jstjohn/SeqPrep), with default parameters. Merged reads coming from ancient samples, and merged and unmerged contemporary data were mapped to the 2.8 Gb *Lynx pardinus* genome (Abascal et al., 2016), and the 2.5Gb *Felis catus* genome v9.0 (GCF_000181335.3) (Buckley et al., 2020) using bwa aln 7.17 (Li & Durbin, 2009). After filtering reads with mapping quality above 30 and sorting using samtools (Li et al., 2009), we added read groups (https://broadinstitute.github.io/picard), merged sequencing runs on an individual basis (Li et al., 2009), marked duplicates using picard-tools (https://broadinstitute.github.io/picard), and realigned using GATK (McKenna et al., 2010). Coverage, mapping statistics, and read length distributions were calculated using samtools (Li et al., 2009). The authenticity of the ancient data was checked by the small fragment length distribution, and the existence of ancient DNA damage -excess of C to T substitutions at the ends of the reads-, as revealed by mapDamage (Jónsson et al., 2013) (Fig. S1). Mapping to Iberian lynx reference genome was used for structure and diversity analysis, while mapping to domestic cat reference genome was used for ABBA-BABA tests to avoid previously reported reference bias (Sheng et al., 2019). For all our analysis we excluded the previously identified low complexity and low mappability regions in lynx or in cat, as well as X and Y chromosome when specified (Abascal et al., 2016; Buckley et al., 2020).

We also sexed the samples by calculating the ratio of depth of coverage in X chromosome to A1 autosome. This method was validated using contemporary samples of known sex.

Population structure

Principal component analysis (PCA) was done for contemporary Iberian lynx data using ANGSD. To avoid coverage depth bias, we subsampled high coverage Iberian lynx samples coming from (Abascal et al., 2016) to ~6x using samtools (Li et al., 2009). Then, we calculated the genotype posterior probabilities using ANGSD (Li, 2011; Kim et al., 2011) and NGSTools/ngsPopGen/ngsCovar (Fumagalli, 2013;

Fumagalli et al., 2013) with the following filters (-uniqueOnly 1 -remove_bads 1 -only_proper_pairs 1 -baq 1 -C 50 -minMapQ 30 -minQ 20 -doCounts 1 –minInd (number of individuals in the population/2) –setMaxDepth (3*sqrt(average (AVR) depth for the population)) -skipTriallelic 1 –SNP_pval 1e-3. To infer the ancestral state we used a fasta sequence of *L. rufus* covering 97% of our bases obtained after mapping reads to the Iberian lynx genome, calling variants using SAMtools mpileup (q-30) (Li et al., 2009), and pseudohaploidizing using pu2fa (-C45) (https://github.com/Paleogenomics/Chrom-Compare). To incorporate ancient data, we did 3 separate PCA including one ancient sample at each time. Then we projected each individual ancient PCA, i.e. contemporary and ancient coordinates, onto the coordinates obtained in the PCA using only contemporary data using (Procrustes analysis, following Borg and Groenen 1997), using the package MCMCpack (Martin, Quinn & Park, 2011) in R (R Core Team, 2019). Finally, we plotted the mean of all the transformed and the original contemporary coordinates, and the ancient coordinates together.

Genotype likelihoods calculated with ANGSD (Kim et al., 2011; Li, 2011) were used for a genomic structure analysis using NGSadmix (Li, 2011; Skotte, Korneliussen & Albrechtsen, 2013), with the same filters used for PCA analysis. NGSadmix was run 10 times from K=1 to K=5, and the results were plotted using R. Optimal K was evaluated following (Evanno, Regnaut & Goudet, 2005), using CLUMPAK.

Nuclear genomic diversity

Two measures of genetic diversity per site (nucleotide diversity (π), and Watterson estimator (θ_W)) was calculated for each population using ANGSD (Li, 2011; Korneliussen et al., 2013; Korneliussen, Albrechtsen & Nielsen, 2014), and realSFS (Korneliussen et al., 2013). To account for differences in coverage and sampling size between contemporary and ancient samples we subsampled the contemporary samples to a depth similar to the ancient one, and ran the analysis for 5 different groups of 3 random individuals in the contemporary populations (subsamples). Using thetaStat (Korneliussen et al., 2013), we performed a sliding-window approach with a window size set to 50,000 bp and a step size of 10,000. All filters described for PCA were also used for diversity calculation except –SNP_pval (not applicable). We calculated mean diversity weighted by the informative sites of each window, and standard deviation by bootstrapping over windows as implemented in boot package

for R (Canty & Ripley, 2017), to account for the correlation among nearby sites due to linkage disequilibrium (LD). Finally, we computed the average of all iterations for the contemporary populations and calculated the standard deviation.

We calculated individual heterozygosity by following the same procedure used for population diversity until the SFS calculation, but in an individual-by-individual basis instead of population-wise. Individual heterozygosity is the second value of the SFS when calculated for a diploid individual. To account for differences in diversity due to variation in depth of coverage among samples we plotted the correlation between heterozygosity and depth.

To check whether diversity differences were conspicuous along the genome, we calculated average diversity (π, and θ_W) across different genomic features representing neutral vs. evolutionary constrained regions, namely intergenic regions, coding gene promoters, 5'UTR, CDS, introns, 3'UTR, long-non-coding RNA promoters, long-non-coding RNA exons, long-non-coding RNA introns, non-coding RNA (mostly miRNAs, snRNAs and snoRNAs), and ultra-conserved-non-coding-elements (UCNE), for the ancient and the contemporary populations. To avoid biases due to differences in sample size we randomly selected one of the five contemporary iterations of size n=3 for this analysis. Standard deviation was calculated using a bootstrapping procedure implemented in the R package boot (Canty & Ripley, 2017) with 100 iterations. Besides, we tested whether topological widows with largest differences in diversity (π, and θ_W) in contemporary versus ancient comparisons were enriched or depleted in CDS sites. To do so, we calculated differences in diversity between the ancient and the contemporary population by windows and calculated mean and standard deviation of the difference distribution. Diversity outliers windows were defined as windows with a difference diversity value lower than mean - 2*standard deviation. Then, we calculated the percentage of CDS sites for each window, and compared the distribution and mean of CDS percentage in outliers vs. non-outlier windows.

Admixture between species

We tested for admixture between Iberian lynx and Eurasian lynx populations with the D statistic (Green et al., 2010; Durand et al., 2011), using the data mapped to the domestic cat genome to avoid reference bias, especially in the ancient samples (Barlow et al., 2018; Günther & Nettelblad, 2019). This method uses haploidized

sequences of 4 individuals: two sister populations (P1 and P2), a potential source of introgression (P3), and an outgroup (P4) that help establishing the ancestral alleles (A). Then, it identifies how many derived alleles (B) are shared between P1 or P2 and the donor P3, leading to two different patterns: ABBA – derived allele shared by P2 and P3; and BABA – derived allele shared between P1 and P3. The D statistics then is calculated as [sum(ABBA) – sum(BABA)] / [sum(ABBA) + sum(BABA)]. Under incomplete lineage sorting we expect that ABBA occurred at the same frequency as BABA and hence, D value would be 0. Excess of ABBA, or BABA are interpreted as admixture between P2-P3, and P1-P3, respectively.

For our analysis we first used Consensify to generate a haploidized sequence (Barlow et al., 2020). This method is specially developed for reducing biases resulting from differences in sequencing coverage and lab protocols in low to medium coverage samples. We then tested for admixture using different meaningful topologies. We performed a total of ~9000 comparisons testing different combination of one ancient Iberian lynx and a single ancient Eurasian lynx inhabiting the Iberian peninsula 2 kya, and contemporary Eurasian and Iberian lynx populations. For all D statistics test, we used the domestic cat (*Felis catus*) as an outgroup. Significance was assessed using a weighted block jack-knife (Durand et al., 2011).

RESULTS

Genome sequencing of the ancient Iberian lynx samples

Shotgun sequencing of the ancient samples resulted in a total of 428,036,688, 889,284,660, and 854,907,829 reads for Algarve, Catalonia and Andújar respectively. Fragment size mode was around 40bp for all samples after excluding adapters and merging the reads (Fig. S1). Samples showed and endogenous content (calculated as the number of unique mapped reads from the total of sequenced reads) of 18%, 21%, and 13% for Algarve, Catalonia, and Andújar sample, when mapped to lynx and 15%, 18% and 12% when mapped to cat. Duplication level (calculated as the number of duplicates from the number of reads mapped) was 59%, 41%, and 32% when mapped to lynx and cat. After trimming, mapping and removing duplicates, depth of coverage was 2.3x (Algarve), 4.3x (Catalonia) and 2.5x (Andújar), and 1.8 (Algarve), 3.5 (Catalonia), and 2 (Andújar) when mapped to Iberian lynx and domestic cat reference genomes, respectively.

Two of the processed ancient Iberian samples were a male (Andújar and Catalonia) and one a female (Algarve), as inferred from a ratio of depth of coverage between X to A1 domestic cat chromosome (0.4, 0.5 and 0.9, respectively). Casas-Marce et al. (2017) reported the dates of the samples to be 2,710 BP for the Catalonian, 4,845 BP for the Andújar, and 2,070 BP for the Algarve sample, accordingly all samples show typical damage pattern C to T at the ends of the reads (Fig. S1).

Description of the ancient Iberian lynx population

PCA analysis and individual-based clustering confirm the relative homogeneity of the ancient samples, hereafter the "ancient population" due to the limited geographical or temporal structure in ancient lynx data, when analysed together with Andújar and Doñana contemporary populations (Fig. 1C-D).

Ancient samples are closer to Andújar population, both in the PCA and individual-based clustering analysis (Fig. 1C-D). When K=3 in the individual-based clustering analysis, the ancient populations emerged as a separate cluster in all runs, with more or less shared ancestry with Andújar contemporary population depending on the run considered (Fig. 1C). However, we observe that PC3 separates the ancient Catalonian sample (dated 2710 ya) from the other two Iberian lynx samples (Algarve, 2070 ya; Andújar 4845 ya) (Fig. 1D), an evidence of some ancient geographical or temporal population structure. This ancient structure is, in any case, shallower than that observed nowadays between the two contemporary populations Andújar and Doñana. And even shallower than the structure observed within Andújar population, as shown by the dispersion of contemporary Andújar samples in PC2 and PC3, and subsequent partitions on the individual clustering analysis, which split Andújar contemporary population at K=4 and Doñana populations at K=5, but maintain the ancient populations as a single cluster (Fig. S2).

Unexpectedly, ancient Iberian lynx genomic diversity is lower than that of the contemporary populations, for both π and θ_W (Fig. 1B), independently of the subsample considered and of whether transitions are included or not in the analysis (Fig. 1B, S3). Precisely, θ_W diversity is 9.50e-06 in the ancient population, which represents 62%, and 38% of θ_W in Doñana and Andújar, respectively (θ_W Ancient=5,24e-05, 85%, and 51% of Doñana and Andújar respectively when considering both transversions and transitions), and ancient π diversity is 1,08e-05, which represents 55% of Doñana

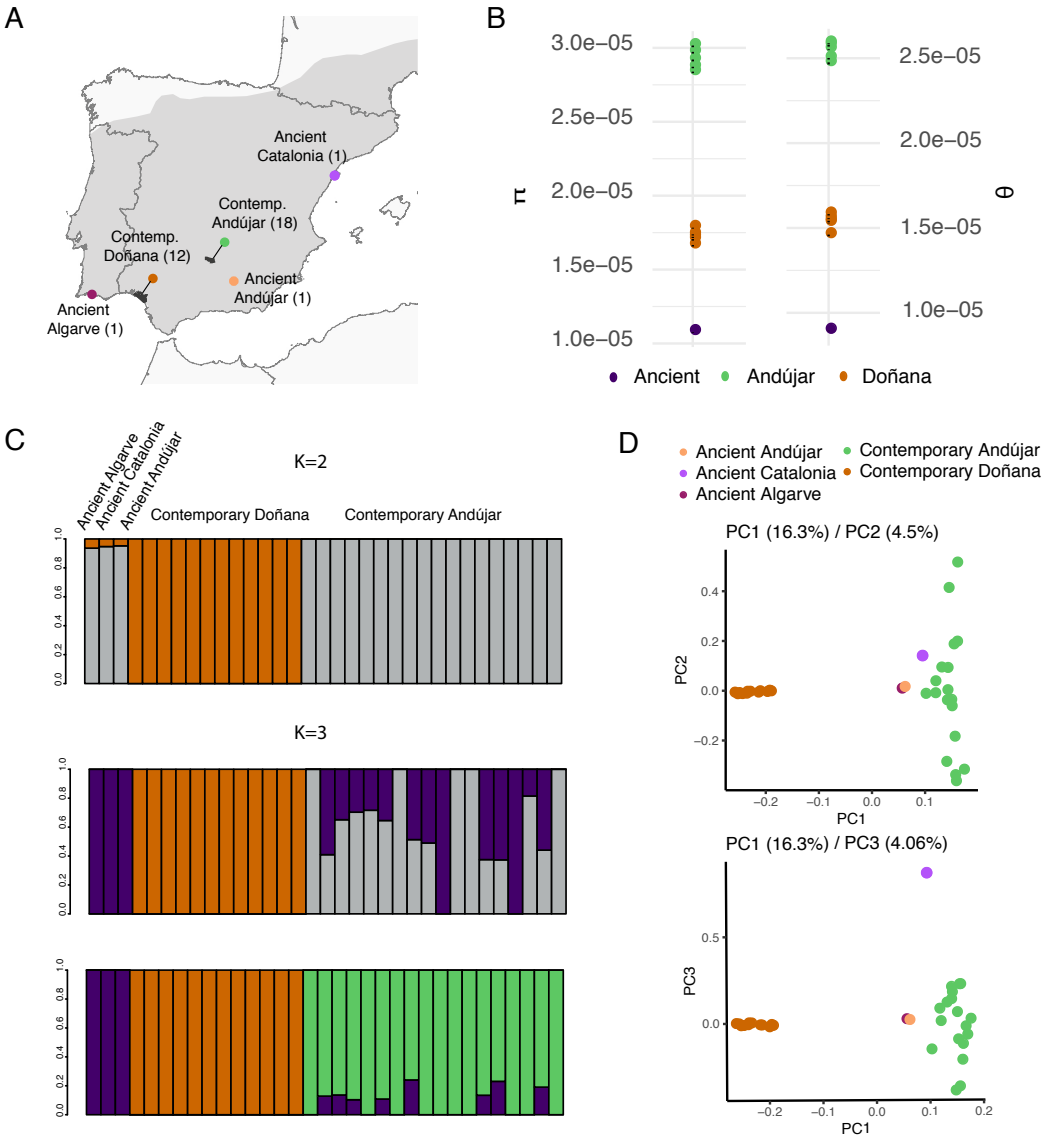


Fig. 1. A. Contemporary and ancient sampling, with the number of individuals sampled in brackets. Ancient Iberian lynx distribution is shown in grey and contemporary distribution is shown in black. B. Diversity of the ancient population (n=3), and different iterations (n=3) of individuals from the contemporary populations Andújar and Doñana considering only transversions. C. Relationships among individuals based on individual-based clustering. Results are shown for K=2 and K=3. For K=2 all runs converge, while for K=3 runs results in two different solutions represented here. D. Principal component analysis. PC1-PC2, and PC2–PC3 results are shown, including the percentage of the variance explained by each axis (in brackets). Color code is maintained in the map and PCA results.

and 32% of Andújar diversity (π Ancient=6,19e-05, 89% and 52% when including transitions and transversions), with subtle differences among distinct contemporary subsamples (Fig. S3).

Genetic diversity differences between ancient and contemporary populations are similar among features for both π, and θ_W, irrespective of whether they are selectively-constrained (e.g. CDS), or neutral (e.g. intergenic or introns) (Fig. 2A). Also, windows with the largest diversity differences between the ancient and any of the contemporary pops are not enriched for coding regions, either for π or θ_W (Fig. 2B).

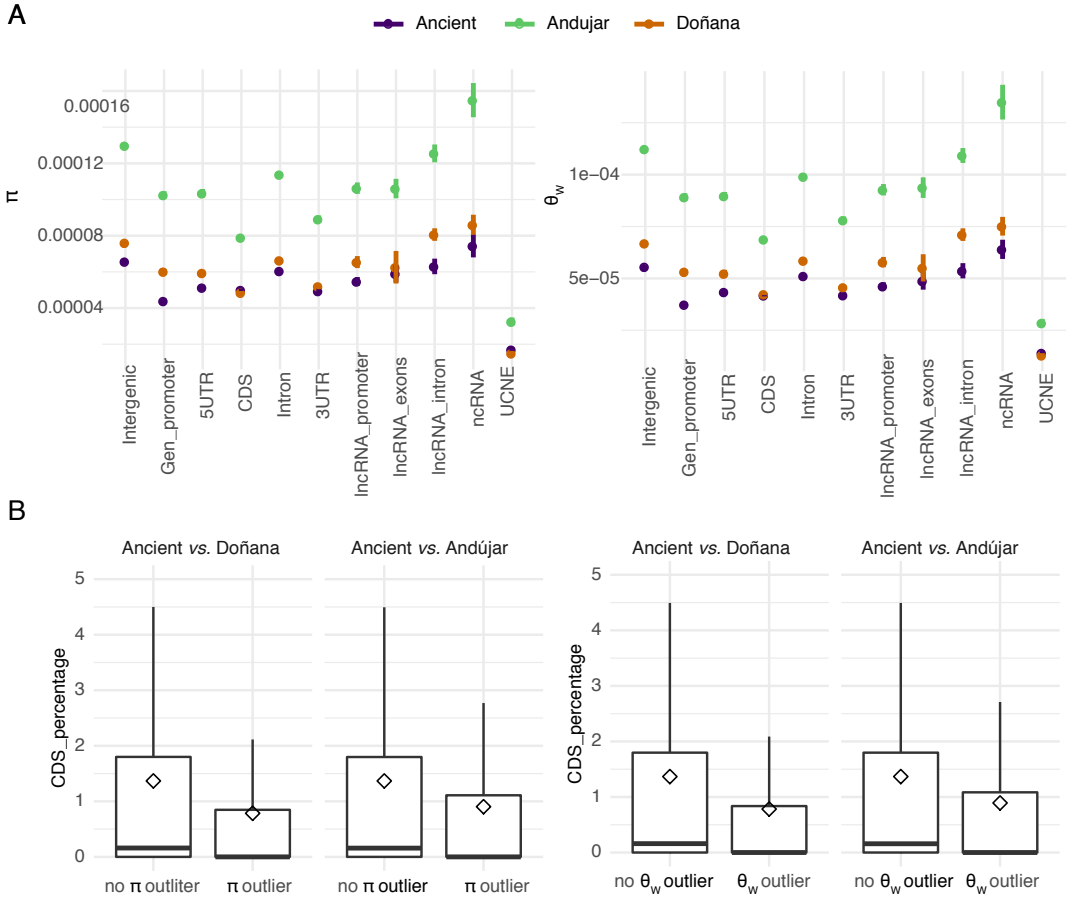


Fig. 2. A. Diversity (nucleotide diversity (π) in the right, and θ_W in the left) amongst the different features for contemporary (Doñana, and Andújar) and ancient populations. B. CDS percentage for outlier diversity difference windows compared to non-outlier ones. Outliers windows were defined as those windows that showed values of diversity difference lower than average diversity difference-2*standard deviation. Comparisons include Ancient vs. Doñana and vs. Andújar, for both π, and θ_W.

Also, individual diversity, i.e. observed heterozygosity, shows pattern similar to population diversity. The heterozygosity of ancient individuals ranges from 95% to 33% of that observed in Doñana individuals, and from 51% to 18% of the Andújar individuals, depending on the samples considered (Fig. S4). Although we detect a positive correlation between observed heterozygosity and depth of coverage, the difference in diversity among Iberian lynx populations is not explained by differences in coverage, as all samples were subsampled to similar depth (Fig. S4).

Admixture between species

Contemporary Iberian lynx samples show a significant excess of shared derived alleles with Eurasian lynx relative to ancient samples, supporting a gene flow event(s) between Eurasian lynx and Iberian lynx in the last 2 ky (Fig. 3A). This gene flow event(s) permeated the whole contemporary Iberian lynx distribution, as the admixed alleles in contemporary individuals are equally distributed between the two contemporary populations (Fig. 3B). Among the plausible Eurasian lynx contemporary populations involved in the admixture process, we find that western European lynx share more alleles with contemporary Iberian lynx than Eurasian lynx from Asia (Fig. 3C). We also find a relative homogeneity in admixture levels within western or eastern Eurasian lynx populations (Fig. S5A, S5B). All in all, our results indicate that contemporary Iberian samples admixed with Eurasian lynx during the last 2 kya.

To investigate further the timing and the source of admixture we used an ancient Eurasian lynx that inhabited the Iberian Peninsula 2 kya (Lucena-Perez, *in prep.*). We find that contemporary Iberian lynx samples share more alleles with contemporary Eurasian lynx than with the ancient Eurasian lynx inhabiting the Iberian Peninsula (Fig. S5C). Also, the ancient Iberian lynx samples do show evidence of admixture with contemporary Eurasian lynx but less so with the ancient Eurasian lynx from Iberia: more alleles are shared with western contemporary individuals than with the ancient Eurasian lynx, the latter being similar to those of the eastern contemporary populations (Fig. S5D, S5E). Also, we found varying levels of admixture in our ancient Iberian lynx samples, with the most ancient sample from Andújar (4.8 kya) showing less shared alleles with Eurasian lynx than samples from Catalonia (2.7 kya) and Algarve (2 kya), further supporting a trend for more recent samples showing more introgression (Fig. S5F, S5G).

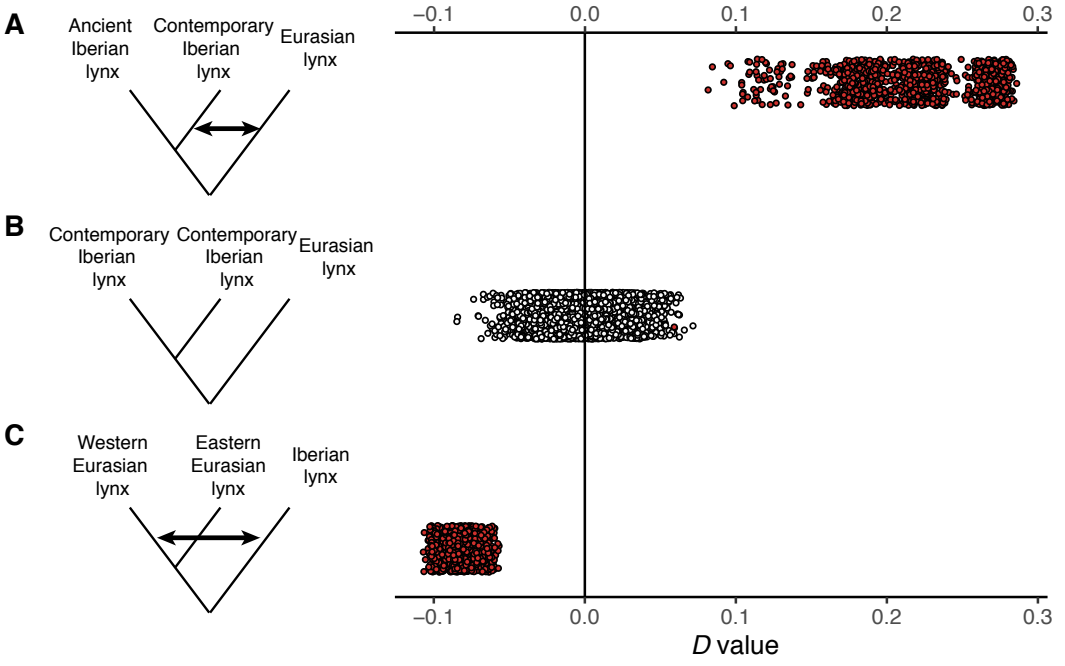


Fig. 4. Results of D statistical tests. Red and grey points show significant and non-significant D values, respectively. In the right, the three different tree topology tested (A-C), with an arrow indicating the gene flow exchange if it was significant.

DISCUSSION

Here, we describe the genetic patterns of three ancient individuals of Iberian lynx based on whole genome data, and we report an unexpected low genetic diversity in ancient Iberian lynx population. Also, we find evidence of gene flow with Western Eurasian a lynx population that was concomitant with, and thus may explain, the observed increase in Iberian lynx genetic diversity within the last 2 ky.

The Iberian lynx, a species currently endemic to the Iberian peninsula, had a wider distribution throughout the Mediterranean biogeographic region, including the North of Italy and France, during the Pleistocene and the beginning of the Holocene (Rodríguez et al., 2011; Rodríguez-Varela et al., 2015; Boscaini et al., 2016; Casas-Marce et al., 2017). Since 100 kya, the species has been declining, with a series of population bottlenecks, the last one occurring during the 20th century (Abascal et al., 2016). As a result, the species was by the beginning of the 21st century confined to two small populations in Southern Spain – Andújar and Doñana. Although the species

was claimed to have no mitochondrial diversity since the Late Pleistocene (Rodríguez et al., 2011), a more recent study using whole mitogenome data from ancient (from 43 kya to 2 kya, n=10), historical (since 1770 till 1990, n=245) and contemporary samples (1990 till 2013, n=172) showed a low ancient diversity that was further reduced by the decline and fragmentation occurring during the last few centuries (Casas-Marce et al., 2017). Also, nuclear microsatellite diversity was progressively reduced since 1770 in most populations and globally, but genetic diversity was not assessed for ancient samples due to methodological constrains. As a result of small long-term population sizes and demographic declines, contemporary Iberian lynx populations have one of the lowest diversity in whole genome data reported to date (Abascal et al., 2016).

With this information we predicted that loss of genetic diversity during the last millennia must have also occurred in the nuclear genome. However, the genome-wide genetic diversity found in ancient Iberian lynx samples is lower than that in extant populations, and to our knowledge the lowest ever reported for any mammal (Fig. 4). Ancient genetic diversity is lower than the one reported for critically endangered species such as baiji (Lipotes vexillifer) and even highly endangered populations like the one of Channel Island fox in San Miguel (Urocyon littoralis). This low ancient diversity suggests small long-term population sizes throughout the species history, likely associated to small ranges, low densities, and recurrent bottlenecks.

The observation of a larger current diversity in Iberian lynx requires a process that increases diversity along the genome occurring in the last 2 ky. Here, we provide conclusive evidence of admixture between Eurasian lynx and Iberian lynx occurring during the last 2 kya that was at least concomitant with, and the most likely ultimate cause of, the inferred increase in genetic diversity in Iberian lynx. Signatures of this admixture are found equally in the two remnant populations –Andújar and Doñana-, suggesting that it was prior to their divergence around 200 ya (Casas-Marcé et al, 2017). We also found evidence of increasing admixture throughout time with the oldest Iberian sample (4.8 kya) showing less share alleles with Eurasian lynx than samples from Catalonia (2.7 kya) and Algarve (2 kya). All European populations, but not Asian populations, of Eurasian lynx show similar proportions of alleles shared with Iberian lynx. However, the one ancient Eurasian sample inhabiting the Iberian Peninsula 2 kya shows levels of allele sharing with contemporary Iberian lynx individuals lower than current European populations, indicating that this particular Eurasian lynx lineage was not involved in the recent admixture process (Fig. S5).

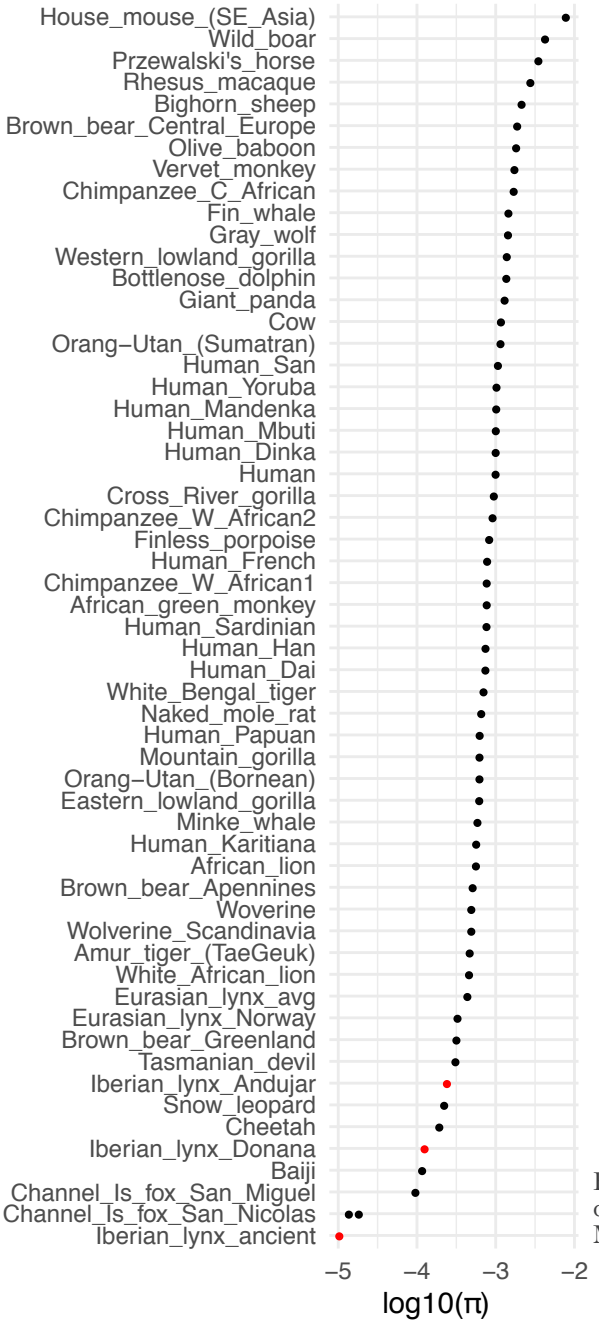


Fig. 4. Comparison of Π diversity among different mammals species or populations. Modified from Robinson et al. (2016).

The Eurasian lynx inhabited Western Europe, including Mediterranean coastal regions, during the Late Pleistocene and the Holocene, reaching the Iberian peninsula during the Pleistocene-Holocene boundary (Altuna, 1980; Sommer & Benecke, 2006; Rodríguez-Varela et al., 2016). During the Pleistocene and at least until 3.7 kya, the Iberian lynx distribution included northern Iberia, southern France and even north-western Italy (Rodríguez-Varela et al., 2015; Boscaini et al., 2016), where the two lynx species have been recorded at the same site (Arene Candide, north Italy) during the Last Glacial Maximum (Cassoli & Tagliacozzo, 1994; Rodríguez-Varela et al., 2015). The plausible past coexistence of the Eurasian and Iberian lynx could have resulted in interbreeding, as widely observed among felids (Li et al., 2016, 2019). However, how is it possible than a 2 ky Eurasian lynx in the Iberian Peninsula show less evidence of admixture with Iberian lynx than modern European populations? Although we can only speculate about this result, some explanations might be behind the pattern. First, both species may have remained ecologically separated in different biogeographic regions: Iberian lynx in Mediterranean and sub-Mediterranean areas and Eurasian lynx in Eurosiberian Atlantic areas (Clavero & Delibes, 2013; Rodríguez-Varela et al., 2016). Second, the Eurasian lynx in Iberia might be a different lineage from those occupying the rest of Europe during the last few millennia. Evidences based on ancient DNA support two different events of colonisation of western Europe: first, during the Late Pleistocene a migration by Eurasian lynx carrying a southern mitogenomic lineage, related to the one in Balkans and Caucasus; second, in the beginning of the Holocene, a colonization with lynx carrying haplotypes related to the ones found in Central European, Asiatic and Baltic populations today (Lucena-Pérez et al, *in prep.*). Modern nuclear genome data suggest that theses mitochondrial lineages, probably originated during periods of isolation in glacial refugia, became largely admixed in Europe in postglacial periods (Lucena-Perez et al., 2020). The fact that all modern populations, from Balkans to Carpathian show similar signals of admixture indicates that the source was this admixed ancestral population, prior to their subsequent differentiation during the Holocene, largely driven by genetic drift. The Eurasian lynx sample from the Iberian Peninsula might be part of a population that entered Iberia prior to the admixture of the two lineages in Europe, and remained isolated there.

The admixture with the Eurasian lynx seems to have played a prominent role in the history of the Iberian lynx, as previously suggested by Abascal et al. (2016) and recently confirmed by Li et al. (2019). The species, whose effective population size has been relatively low throughout its history, may have periodically increased its diversity by pulses of admixture with the Eurasian lynx. The fitness consequences of the inferred admixture events between the Eurasian and the Iberian lynx are hard to predict.

Theoretically, we expect that the introduction of new alleles via gene flow in a population could lead to an increase or decrease in the fitness of the individuals of the populations, depending on different variables, such as the genetic load of the populations involved in the process, the time of divergence of the two populations, or the existence of local adaptations (Frankham et al., 2011). The admixture between the two lynx species could lead to an increase in fitness given that the two species show the same chromosomal number and large chromosomal reorganizations are presumably very scarce (Abascal et al., 2016), and admixture have been extensive and common in the history of the species since their divergence (Abascal et al., 2016; Li et al., 2019). Also, new alleles could increase the Iberian lynx low genetic diversity for adaptation to environmental change, i.e. evolutionary potential. However, we are dealing with two different species, that evolved mostly in allopatry, and present notable morphological, behavioural and ecological differences (Sunquist & Sunquist, 2002; Nowak, 2005), which could lead to outbreeding depression. As the Iberian lynx has dramatically declined during the last 2 kya, likely due to direct and indirect human impacts and with likely but unknown contribution of genetic factors, the evaluation of the impact of the recent admixture on the species is impossible.

In conservation genetics theory and practice, establishing gene flow with a closely-related genetically healthier population is seen as a mitigation strategy that can decrease extinction risks through the restoration of genetic diversity and the reduction in inbreeding levels (Hedrick & Fredrickson, 2010; Hedrick & Garcia-Dorado, 2016). This type of conservation action, known as genetic rescue, is usually taken as an urgent and extreme measure, as it has some risk derived from the possibility of outbreeding depression, i.e. that the admixture results in a reduction of fitness due to either the influx of maladaptive alleles or the disruption of co-adapted gene complexes or fixation of chromosomal variants (Frankham et al., 2011). Nevertheless, accumulated evidences have shown that genetic rescue has more positive than negative effects in general, with the reduction of inbreeding depression usually compensating possible outbreeding risks (Frankham, 2015). Using a closely related species as the source of novel alleles has remained mostly out of discussion due to anticipated outbreeding depression. However, increasing evidences suggest that natural introgression is a far more common and widespread process than previously thought, to the point that current species cannot longer be pictured as hermetic static compartments, but rather as dynamic entities interconnected with others by episodic events of introgression. With this background, genetic rescue across species should be carefully considered in those extreme cases in which species persistence is severely compromised by genetic factors and no source population of the same species exists.

Looking backwards to plan the future

Currently Iberian lynx is being managed at a species level to reduce inbreeding and maximize genetic diversity. Genetic rescue of the eroded Doñana population in recent years by the genetically healthier Andújar, via translocation, has restored diversity and apparently increased fitness (Lopez-Zamora *in prep.*). Also, the admixed captive population has restored the diversity of the species close to historical levels (Kleinman-Ruiz et al., 2019), and the species has increased from 100 individuals in two populations at the beginning of this century to almost 1000 distributed in the two remnant and several reintroduced populations, that are being genetically managed to preserve genetic diversity and reduce inbreeding. However, even though genetic factors might no longer be a risk for the species short-term viability if genetic management continues, its long-term viability, as determined by its adaptive potential, will remain relatively low due to a low intrinsic genetic diversity. Our results reporting admixture with Eurasian lynx over the last 2 kya, encourage interspecific hybridization to be carefully considered as a measure to restore the adaptive potential of the specie. Furthermore, plans to reintroduce Eurasian lynx in the Pyrenees are being discussed, along with the introduction of Iberian lynx in Catalonia, relatively close to Pyrenees area, what would make natural hybridization eventually possible. As the outcome and associated risks are unknown, we call for a careful consideration of interspecific gene flow both as a mitigation strategy and a threat for species survival.

AUTHOR CONTRIBUTION

J.A.G. conceived the project, A.B., J.A.G. and M.L.P., designed the study; C.D., F.N., and J.N. provided ancient samples; M.L.P. performed the lab work under J.P. supervision; A.B., J.P., M.L.P., analysed data; A.B., J.A.G., J.P., and M.L.P. interpreted the results; M.L.P. drafted the manuscript with support from J.A.G.; J.A.G. supervised the project. All authors approved the final version of the manuscript.

ACKNOWLEDGEMENTS

This project was funded by the Dirección General de Investigación Científica y Técnica of the Spanish Ministerio de Ciencia e Innovacion (MICINN) through projects CGL2013-47755-P and CGL2017-84641-P, and by the Fundación BBVA through a project on ancient lynx genetics to M.D. MLP was supported by PhD contracts

from the Programa Internacional de Becas "La Caixa-Severo Ochoa" of the "La Caixa" Foundation (ID 100010434), under agreement LCF/BQ/SO14/52250035. Also, María Lucena-Perez received funding from the European Science Foundation Reseach Networking Programmes - ConGenOmics for a Exchange Visit Grant. Radicarbon dating was performed by Centro Nacional de Aceleradores in Seville, Spain. Calculations were carried out in the Genomics servers of Doñana's Singular Scientific-Technical Infrastructure (ICTS-RBD) and in Fundación Pública Galega Centro Tecnolóxico de Supercomputación de Galicia (CESGA). EBD-CSIC received support from the Spanish Ministry of Economy and Competitiveness under the 'Centro de Excelencia Severo Ochoa 2013-2017' program, grants SEV-2012-0262. Besides the free software mentioned in the text, this work was possible thanks to the free, open source packages: dplyr (Wickham et al., 2017), ggplot2 (Wickham, 2009), tidyr (Wickham y Henry, 2018), broom (Robinson, 2018), RColorBrewer (Neuwirth, 2014), viridis (Garnier, 2018), stringr (Wickham y Henry, 2018), knitr (Xie, 2014, 2015, 2017), reshape2 (Wickham, 2007), data.table (Dowle y Srinivasan, 2017), rlang (Henry y Wickham, 2018).

08 APÉNDICE

GENETIC RESTORATION AFTER INTERSPECIFIC INTROGRESSION IN AN ENDANGERED FELID

SUPPLEMENTARY FIGURE

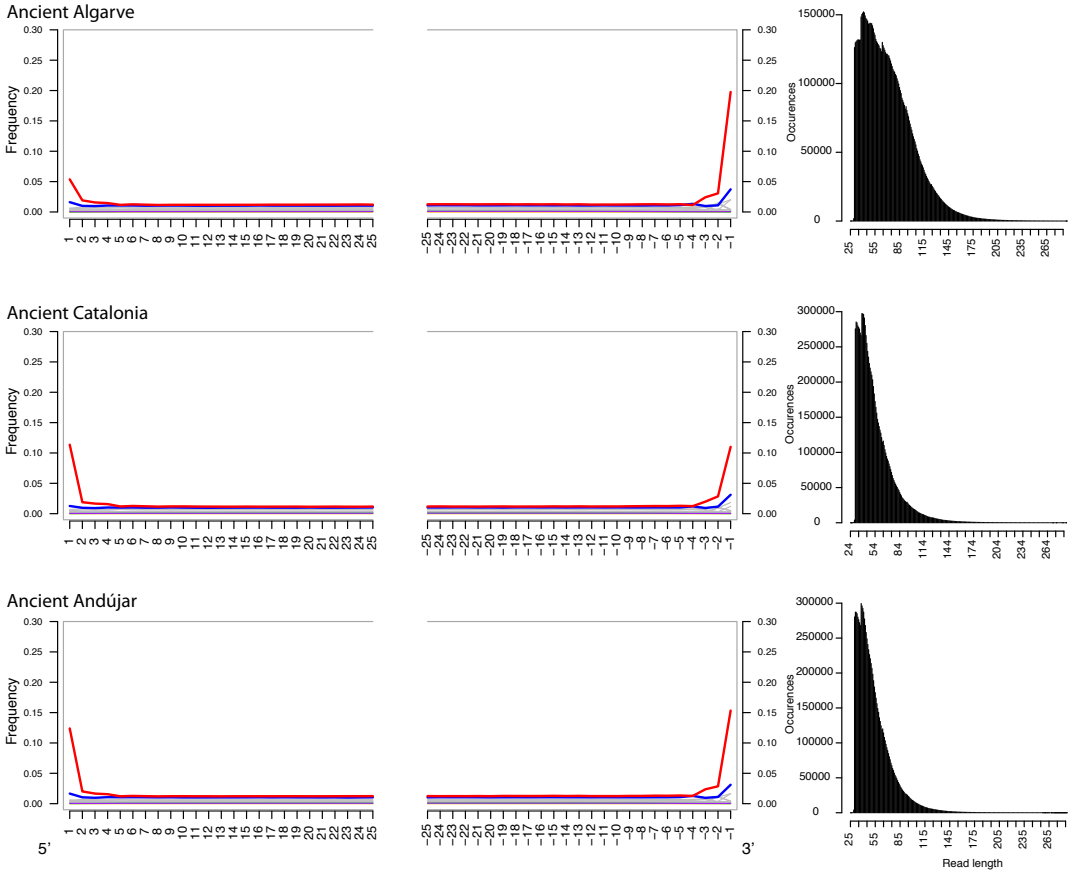


Fig. S1. Authentication of ancient DNA data using MapDamage. On the left, cytosine deamination patterns. X axis indicates the individual nucleotide position of the reads. Plots show the proportion of T where the reference genome possesses a C (red) and proportion of A where the reference possesses a G (blue). C to T changes are typical of ancient damage patterns. Plot on the right represents the fragment length distribution determined by the merging of the two reads. Minimum red length for mapping used was 30bp, resulting in a truncation of the plot.

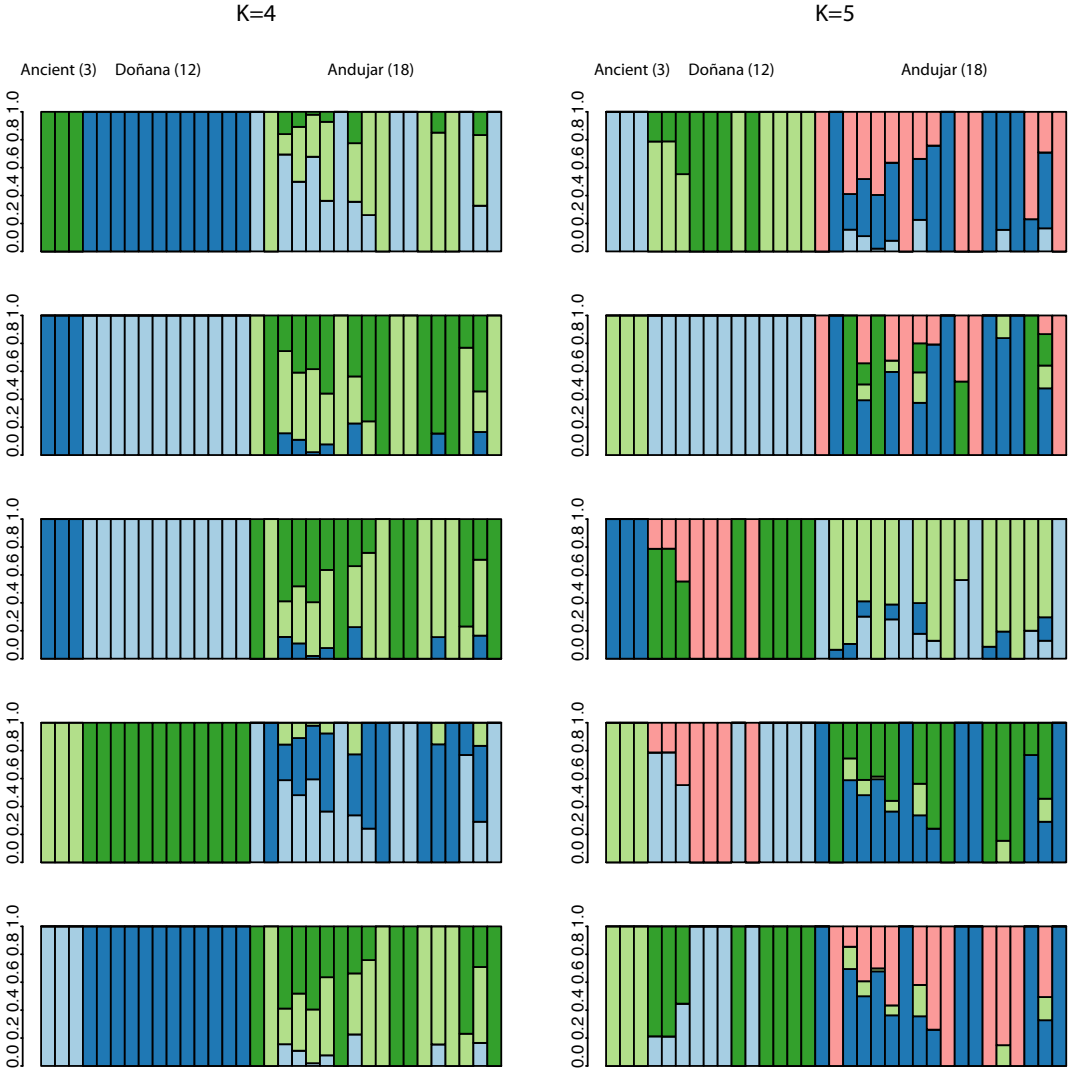


Fig. S2. Alternative clustering patterns obtained in 5 runs of NGSadmix for K=4 and K=5. Population names on top, with number of samples into brackets.

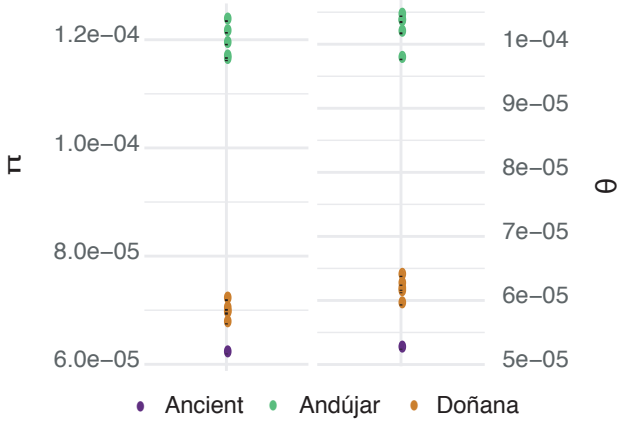


Fig. S3. Diversity of the ancient population (n=3), and different random subsamples (n=3) of individuals from the contemporary populations, Andújar and Doñana, considering transitions and transversion.

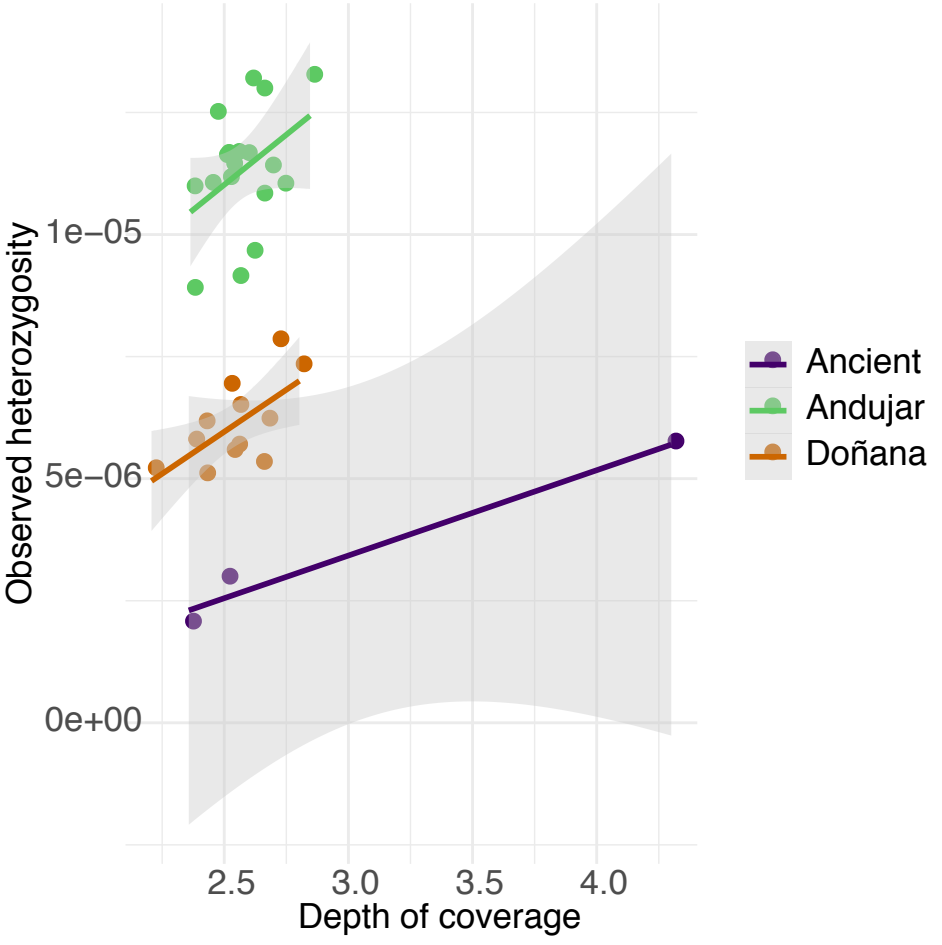


Fig. S4. Observed heterozygosity per samples vs. depth of coverage of the sample. Differences among the ancient samples might be attributed to differences in coverage, with the one with higher coverage showing smaller differences with contemporary populations, but differences between ancient and modern samples are not explained only by differences in coverage.

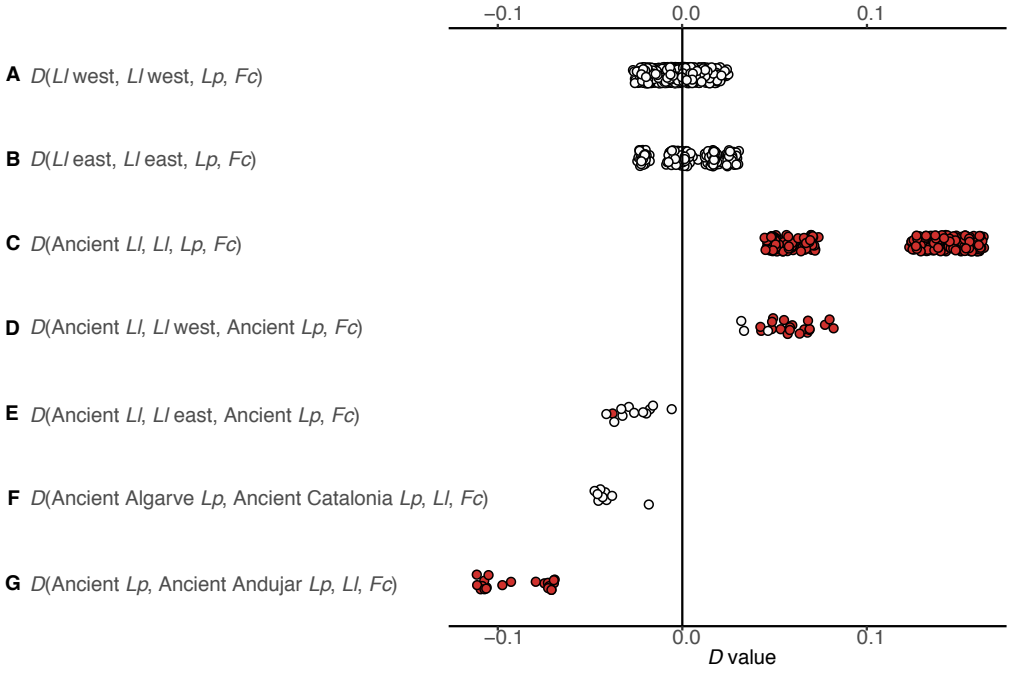


Fig. S5. Results of D statistical tests. Red and grey points show significant and non-significant D values, respectively. In the right, the three different tree topology tested (A-G). Ll=*Lynx lynx*, Fc=*Felis Catus*. Ll west included the following populations: Kirov, Caucasus, Balkans, and Carpathians, while Ll east included Vladivostok, and Yakutia.

SUPPLEMENTARY TABLES

Table S1. List of ancient samples used in this study.
* Dated by archeological context
Detry, C. and A. A. M. (2013). "A fauna da Idade do Ferro e da Época Romana de Monte Molião (Lagos, Algarve): continuidades e rupturas na dieta alimentar." Revista Portuguesa de Arqueologia 16: 213-226.

Sample	ID_code	Site	Locality	Province (Country)	AMS C14 dating			Assigned date	Period
					Lab code	Conventional age	2 sigma calibration		
a_lp_al_1311	273-4-A85	Monte Molião	Lagos	Algarve (Portugal)				2,070 BP*	Roman
a_lp_ca_1309	M88-302-52	La Moleta del Remei	Alcanar	Tarragona (Spain)	408255	2520 +/- 30 BP	BP 2740 to 2680; BP 2640 to 2610; BP 2600 to 2490	2,710 BP	Iberian
a_lp_sm_1310	UB-10/2006_1259	Eras del Alcázar	Úbeda	Jaen (Spain)	408254	4270 +/- 30 BP	BP 4865 to 4825	4,845 BP	Copper Age

Table S2. List of contemporary iberian lynx and Eurasian samples used in this study. Assigned birth date (Iberian lynx) or sampling year (Eurasian lynx). Date=birth date (Iberian lynx); date=sampling date (Eurasian lynx).

Sample	Specie	Population	Date	Reference	Sex
c_lp_sm_0325	L. pardinus	Andújar	1999	Lucena-Perez et al. in prep	F
c_lp_sm_0140	L. pardinus	Andújar	2000	Abascal et al. 2016	M
c_lp_sm_0138	L. pardinus	Andújar	2002	Abascal et al. 2016	M
c_lp_sm_0161	L. pardinus	Andújar	2002	Lucena-Perez et al. in prep	F
c_lp_sm_0134	L. pardinus	Andújar	2004	Lucena-Perez et al. in prep	F
c_lp_sm_0155	L. pardinus	Andújar	2004	Lucena-Perez et al. in prep	F
c_lp_sm_0156	L. pardinus	Andújar	2004	Lucena-Perez et al. in prep	F
c_lp_sm_0186	L. pardinus	Andújar	2004	Abascal et al. 2016	M
c_lp_sm_0359	L. pardinus	Andújar	2004	Abascal et al. 2016	M
c_lp_sm_0185	L. pardinus	Andújar	2005	Abascal et al. 2016	M
c_lp_sm_0208	L. pardinus	Andújar	2005	Lucena-Perez et al. in prep	F
c_lp_sm_0213	L. pardinus	Andújar	2005	Lucena-Perez et al. in prep	F
c_lp_sm_0206	L. pardinus	Andújar	2006	Lucena-Perez et al. in prep	F
c_lp_sm_0221	L. pardinus	Andújar	2006	Lucena-Perez et al. in prep	M
c_lp_sm_0226	L. pardinus	Andújar	2006	Lucena-Perez et al. in prep	M
c_lp_sm_0276	L. pardinus	Andújar	2006	Lucena-Perez et al. in prep	F
c_lp_sm_0298	L. pardinus	Andújar	2007	Abascal et al. 2016	M
c_lp_sm_0320	L. pardinus	Andújar	2008	Lucena-Perez et al. in prep	F
c_lp_sm_0450	L. pardinus	Andújar	2009	Lucena-Perez et al. in prep	M
c_lp_do_0007	L. pardinus	Doñana	1987	Abascal et al. 2016	M
c_lp_do_0144	L. pardinus	Doñana	2000	Lucena-Perez et al.	M
c_lp_do_0141	L. pardinus	Doñana	2001	Lucena-Perez et al. in prep	M
c_lp_do_0163	L. pardinus	Doñana	2001	Lucena-Perez et al. in prep	F
c_lp_do_0162	L. pardinus	Doñana	2002	Lucena-Perez et al. in prep	F
c_lp_do_0173	L. pardinus	Doñana	2003	Abascal et al. 2016	M
c_lp_do_0153	L. pardinus	Doñana	2004	Abascal et al. 2016	M
c_lp_do_0300	L. pardinus	Doñana	2007	Lucena-Perez et al. in prep	M
c_lp_do_0333	L. pardinus	Doñana	2008	Lucena-Perez et al. in prep	F
c_lp_do_0335	L. pardinus	Doñana	2008	Lucena-Perez et al. in prep	F
c_lp_do_0443	L. pardinus	Doñana	2010	Abascal et al. 2016	M
c_lp_do_0444	L. pardinus	Doñana	2010	Lucena-Perez et al. in prep	F
c_ll_ba_0229	L. lynx	Balkans	2012	Lucena-Perez et al. in prep	M
c_ll_ba_0227	L. lynx	Balkans	2018	Lucena-Perez et al. in prep	F
c_ll_cr_0205	L. lynx	Carpathians	2007	Lucena-Perez et al. 2020	F
c_ll_cr_0212	L. lynx	Carpathians	2012	Lucena-Perez et al. 2020	M
c_ll_ca_0241	L. lynx	Caucasus	2019	Lucena-Perez et al. in prep	F
c_ll_ca_0242	L. lynx	Caucasus	2019	Lucena-Perez et al. in prep	F
c_ll_ki_0090	L. lynx	Kirov	2011	Lucena-Perez et al. 2020	F
c_ll_ki_0091	L. lynx	Kirov	2011	Lucena-Perez et al. 2020	M
c_ll_vl_0112	L. lynx	Primorsky Krai	2014	Lucena-Perez et al. 2020	F
c_ll_vl_0113	L. lynx	Primorsky Krai	2015	Lucena-Perez et al. 2020	M
c_ll_ya_0138	L. lynx	Yakutia	2015	Lucena-Perez et al. 2020	M
c_ll_ya_0146	L. lynx	Yakutia	2015	Lucena-Perez et al. 2020	M

04 CAPÍTULO

BOTTLENECKED-ASSOCIATED CHANGES IN THE GENOMIC LANDSCAPE OF GENETIC DIVERSITY IN WILD LYNX POPULATIONS

Maria Lucena-Perez

Elena Marmesat, Daniel Kleinman-Ruiz, Alexander P. Saveljev, Krzysztof Schmidt, José A. Godoy.

ABSTRACT

Demographic bottlenecks generally reduce genetic diversity through more intense genetic drift, but their net effect may vary along the genome due to the local effects of recombination, mutation and selection. Here, we analysed the changes in genetic diversity following a bottleneck by comparing whole-genome diversity patterns in populations with and without severe recent documented declines of Iberian (*Lynx pardinus*, n=31) and Eurasian lynx (*Lynx lynx,* n=29). As expected, overall genomic diversity correlated negatively with bottleneck intensity and/or duration. Correlations of genetic diversity with several known correlates were observed in non-bottlenecked populations, but were weaker in bottlenecked populations. Also, the more selectively constrained features and the X chromosome showed an accumulation of low frequency alleles in bottlenecked populations, even resulting in higher θ_W diversity than in non-bottlenecked populations. This accumulation seems to be related to both a higher mutational input in those regions creating a large collection of low-frequency variants, a few of which increase in frequency during the bottleneck, and the relaxation of purifying selection, which could affect not only protein structure and function but also regulation of gene expression. This study provides novel insights into the genomic and fitness consequence of bottlenecks in wild populations with relevance for the conservation of endangered species.

INTRODUCTION

Genetic diversity in populations is the result of the interplay between mutation, genetic drift, recombination, selection and gene flow (Ellegren & Galtier, 2016; Ballenghien, Faivre & Galtier, 2017). In demographically stable populations, the site frequency spectrum (SFS) and genetic diversity varies along the genome and across genomic features, following variations in the strength of recombination, selection and mutation, and creating a genomic landscape of genetic diversity. Changes in effective population size alter the equilibrium patterns of genetic diversity across the genome expected at stable population sizes. In particular, a demographic bottleneck (i.e. a sudden reduction in Ne) will increase the action of genetic drift and cause a general loss of genetic diversity, in a magnitude determined by its severity and duration (Garza & Williamson, 2001), and will distort the SFS through the preferential loss of low frequency alleles (Nei & Maruyama, 1975). However, we might expect that the effect of demographic bottlenecks on genetic diversity will locally depart from predictions based on equilibrium or simple drift models because evolutionary forces interact in complex ways between them and with genetic drift (Duret & Arndt, 2008; Duret & Galtier, 2009; Pratto et al., 2014; Williams et al., 2015; Terekhanova et al., 2017; Smith, Arndt & Eyre-Walker, 2018; Halldorsson et al., 2019), A population bottleneck is therefore expected to affect differentially distinct genomic compartments (i.e. chromosomes, chromosomal regions, or features) as they differ in these evolutionary forces.

Genetic diversity is ultimately created by mutation and mutation rates vary extensively along and among chromosomes (Gonzalez-Perez, Sabarinathan & Lopez-Bigas, 2019). For example, it is known that GC-content is positively correlated with genetic diversity, partly due to the hyper-mutability of CpG dinucleotides (Duret & Arndt, 2008; Duret & Galtier, 2009; Smith, Arndt & Eyre-Walker, 2018). Recombination is also positively correlated with genetic diversity due to its associated mutagenic effect (Duret & Arndt, 2008; Duret & Galtier, 2009; Pratto et al., 2014; Terekhanova et al., 2017; Smith, Arndt & Eyre-Walker, 2018; Halldorsson et al., 2019).

Natural selection impacts genetic diversity at selected sites by either reducing (purifying or negative selection, and positive selection) or maintaining it (balancing selection). The effects of selection on diversity may extend to neighbouring linked neutral sites especially in regions of low recombination, a process often referred to as hitchhiking or linked selection, resulting in selective sweeps and background selection driven by positive and negative directional selection, respectively (Cutter

& Payseur, 2013). Furthermore, the effectiveness of selection may be hampered if multiple linked loci experience selective pressures simultaneously via Hill-Robertson interference (Hill & Robertson, 1966; Felsenstein, 1974).

A reduction in Ne would impact regions under selection leading to two predictions. Firstly, variants with selection coefficients lower than the inverse of Ne will behave as effectively neutral (Kimura, 1962) which results in the accumulation of moderately deleterious variation in small populations due to the relaxation of purifying selection (Charlesworth, Morgan & Charlesworth, 1993). Secondly, in small populations, higher rates of inbreeding and the increase in frequency of (partially) recessive deleterious variants may facilitate their elimination by selection (i.e. purging) (Hedrick, 1994; Garcia-Dorado, 2012; Hedrick & Garcia-Dorado, 2016).

Despite the increasing interest of the study of the evolutionary consequences of bottlenecks for the conservation of biodiversity, the lack of the required resources, such as a reference genome, and/or appropriate samples, in most endangered species has limited progress. We find, on one hand, a few empirical studies assessing differences in diversity along the genome in stable populations of non-model organisms (e.g. Corcoran et al., 2017; Dutoit, Burri, Nater, Mugal, & Ellegren, 2016; Wang, Street, Scofield, & Ingvarsson, 2016), and on the other, simulations exploring the dynamics of diversity in non-equilibrium scenarios, including demographic bottlenecks (Torres et al., 2020), but to our knowledge there are no empirical studies analysing the changes in the genomic landscape of diversity following a population bottleneck on wild conservation-relevant populations.

Here we characterize the changes in the genomic landscape of diversity brought about by a demographic bottleneck in the two lynx species present in Eurasia, the highly endangered Iberian lynx (*Lynx pardinus*) and the broadly distributed Eurasian lynx (*Lynx lynx*). By taking advantage of the recent de novo assembled and annotated Iberian lynx reference draft genome (Abascal et al., 2016) we measure genetic diversity across chromosomes, chromosomal regions, and genomic features, in populations with contrasting recent demographic histories. To do so, we assess genomic diversity patterns in populations that best represent an ancestral unaffected (i.e. pre-bottleneck) for each species or lineage (from now on referred to as non-bottlenecked populations, NB); and we compare them with those of genetically eroded populations of the same lineage (bottlenecked populations, B). We also assess the role of recombination, mutation, and selection in shaping post-bottleneck diversity patterns by comparing local values between regions or features with contrasting genetic diversity dynamics.

MATERIALS & METHODS

Study populations

The Iberian lynx was recognized as the most endangered felid in the world (Nowell & Jackson, 1996). In 2002, when the species was classify as critically endangered, there were only two remnant populations with less than 100 individuals that differed on genetic patterns and demographic history (Rodríguez & Delibes, 2003; Ferreras et al., 2010; Palomares et al., 2011; Casas-Marce et al., 2017). The highly eroded Doñana population was isolated during two centuries at an estimated Ne of 20 individuals, reaching 10 in the last few decades (Casas-Marce et al., 2013, 2017). In contrast, Andújar is the remnant of the Eastern Sierra Morena population, which remained large and genetically connected with other populations until ca 1980, being genetically close to the historical Iberian lynx population (Casas-Marce et al., 2017).

Unlike the Iberian lynx, the Eurasian lynx, is one of the most broadly distributed felids in the world (Breitenmoser et al., 2015). In Europe there are differences between populations in terms of neutral genetic differentiation and diversity (Lucena-Perez et al., 2020). Particularly, NE-Poland and the Scandinavian populations went through pronounced declines and remained isolated from other lynx populations during the last century, resulting in currently low genetic diversity and high differentiation. On the contrary, eastern European populations, represented here by the Kirov population, remained well connected and relatively large, and represents the population of the European lineage with the highest genetic diversity (Ratkiewicz et al., 2014; Lucena-Perez et al., 2020).

Sampling, DNA extraction, library preparation & sequencing

Whole genome resequencing data were obtained from 60 Iberian and Eurasian lynx (Fig. 1, Table S1). We sampled and processed 20 Iberian samples, which together with available sequence data from 11 additional individuals (Abascal et al., 2016), resulted in a total of 31 Iberian lynx from the two populations: Andújar (n=19), Doñana (n=12). We used 29 whole-genome sequences from Eurasian lynx (Lucena-Perez et al., 2020) from three different populations: Kirov region, Russia (n=13), NE-Poland (Białowieża and Knyszyn Primeval Forests) (n=8), and Norway (n=8). We determined the ancestral state and calculated divergence using one bobcat (*Lynx rufus*) and one Canada lynx (*Lynx canadensis*) genome as in Lucena-Perez et al. (2020).

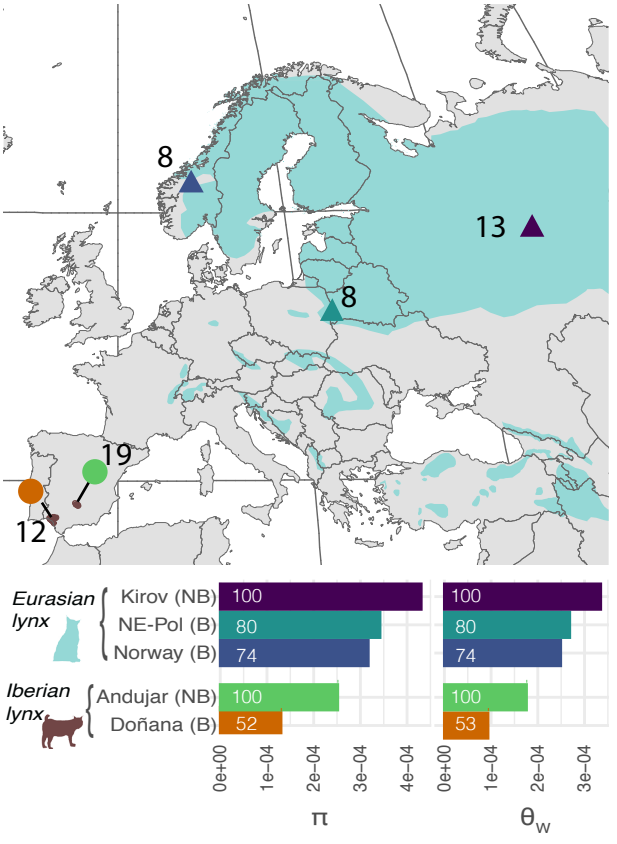


Fig. 1. Sampling and global diversity (π and θ_W) of the five studied populations.. Populations sampled for Eurasian lynx (triangle) and Iberian lynx (circle) are marked within the distribution of the species (Eurasian lynx, European part of the distribution in blue, and Iberian lynx in dark red). Diversity estimates for each population are represented below. For B populations the diversity percentage relative to the reference NB population is indicated in the bar plot.

Samples were digested overnight using proteinase K and DNA was extracted using NucleoMag B-beads (NucleoMag DNA from tissue kit) in LEM-EBD facilities (Seville, Spain). gDNA samples were used for preparing Illumina sequencing compatible paired-end libraries. The libraries were prepared, quantified and sequenced in Illumina HiSeq2000 flowcell v3 (Illumina Inc.), in 2x101bp paired-end mode, following standard Illumina procedures at Centro Nacional de Análisis Genómico (CNAG-CRG, Barcelona, Spain).

Quality control, trimming & mapping

Iberian and Eurasian lynx data was quality-controlled using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc), and adaptors were removed when necessary using SeqPrep (https://github.com/jstjohn/SeqPrep). All short-read sequence data were mapped to a 2.8Gb Iberian lynx LYPA1.0 genome assembly

(Abascal et al., 2016) using BWA-MEM (Li, 2013) with default parameters. We added read groups to each sample using picard-tools (https://broadinstitute.github.io/picard/) and merged the bam files from the same individual with SAMtools merge (Li et al., 2009). We marked PCR duplicates using picard-tools (https://broadinstitute.github.io/picard/) and performed a local realignment and a base quality recalibration of the data using GATK 3.4 (McKenna et al., 2010). We calculated overall mapping stats using SAMtools flagstat (Li et al., 2009) and average depth using SAMtools depth (Li et al., 2009). In order to standardized sample depth to avoid any bias, medium-high depth samples were subsampled to a depth within the range of the newly sequenced samples, using SAMtools view -s (Li et al., 2009).

Genome annotation

The Iberian lynx reference genome has been annotated for several genomic features, namely 3UTR, CDS, introns, and 5UTR (defined on the principal isoform of each gene), lncRNAs, and ncRNAs (Abascal et al., 2016). Here, we also included promoters of protein-coding genes and lncRNA, defined as sequences 1000 bp upstream of the gene or lncRNA, respectively, and Ultra Conserved Non-Coding Elements (UCNEs). UCNEs (Dimitrieva & Bucher, 2013) were annotated by translating human coordinates (https://ccg.vital-it.ch/UCNEbase/) into domestic cat Felis_catus_5.0 coordinates using LiftOver (http://rohsdb.cmb.usc.edu/GBshape/cgi-bin/hgLiftOver), and these to lynx LYPA1.0 coordinates using lynx to cat synteny (Abascal et al., 2016). To define intergenic regions we added a security buffer of +/-1000bp to any annotated region to avoid the influence of adjacent areas, using bedtools subtract and bedtools intersect (Quinlan & Hall, 2010). Each single 3UTR, CDS, intron, 5UTR, lncRNA, ncRNA, promoter (for genes and lncRNA), UCNE, and intergenic region – from now on, each single unit – was also assigned to chromosomes when possible using lynx to cat synteny, and excluding from the analysis units that overlapped more than one chromosome. We annotated as subtelomeric and pericentromeric those contigs containing more than 1000 bp syntenic to these regions in the domestic cat genome, i.e. 2Mb away from the telomeres and 10Mb around the centromere, respectively (Abascal et al., 2016). Then, we calculated the percentage of sites in each unit overlapping subtelomeric, pericentromeric or interstitial regions (defined as regions not pericentromeric or subtelomeric), and assigned units to regions where the overlap was $\geq 75\%$.

Genetic diversity per unit

For each population, genetic diversity per site was calculated using ANGSD (Korneliussen et al., 2013) with the following filters: −uniqueOnly 1 −remove_bads 1 −only_proper_pairs 1 −baq 1 −C 50 −minMapQ 30 −minQ 20 −doCounts 1 −minInd (number of individuals in the population/2) −setMaxDepth (average [AVR] depth for the population + [0.95 * stdev depth for the population]) −setMinDepth (AVR depth for the population − [0.95 * stdev depth for the population]) −skipTriallelic 1). First, we calculated the site allele frequency (SAF) and the site frequency spectrum (SFS) for each population using ANGSD and NGSTOOLS/realSFS (Korneliussen et al., 2013). This SFS plus its corresponding SAF were used to calculate nucleotide diversity (π), and Watterson theta estimator (θ_W) per site (Korneliussen et al., 2013). It must be noted that ANGSD reports values of diversity per site whenever a site passes the filters, without distinguishing between variable and invariant sites, and that site diversity is not zero even for invariant sites because it considers the likelihood of all genotypes in the calculation. Then, we calculated diversity per unit by averaging diversity across sites, considering only informative sites (defined as all sites remaining after filtering). To avoid biases due to very few informative sites, we only considered units with more than 50 informative sites and with information for at least 20% of the positions. Results are presented for autosomes (A), unless otherwise noted. For the X chromosome, θ_W was corrected for the sample size of this chromosome based on the number of males and females sampled in each population (Lucena-Perez et al., 2020). Additionally, we excluded the X-chromosome pseudoautosomal region (PAR) delimited here as 10Mb from the extreme of the chromosome, based on an estimate of 9 Mb for the domestic cat X-chromosome PAR (Li et al., 2016).

We averaged π, and θ_W for the different populations, chromosomes, regions (subtelomeric regions vs. pericentromeric regions vs. interstitial), and features (Intergenic, Gen promoter, 5UTR, CDS, Intron, 3UTR, lncRNA promoter, lncRNA exons, lncRNA intron, ncRNA, UCNE), weighting by number of informative sites. Then, we used π, and θ_W weighted means to calculate a measure of the SFS skewness (S) as $1-(\pi/\theta_W)$ (Becher, Jackson & Charlesworth, 2020). Dispersion in diversity estimates (sd) was calculated by bootstrapping units using the R package boot (Canty & Ripley, 2017) with 100 iterations. We then computed ratios of diversity in B vs. NB populations in different chromosomes, regions, and features.

We also computed π, and θ_W X/A diversity ratios for each population, and feature. Dispersion in X/A ratios was calculated from dispersion in X and dispersion in A, using a propagation of the uncertainties formula:

$$\sqrt{}(A\ stdv\text{^}2/A\text{^}2 + X\ stdv\text{^}2/X\text{^}2\ ) * (X/A)\text{^}2$$

Relative diversity differences between B and NB populations across features

We tested diversity differences across features between pairs of one B population and its corresponding NB population. By doing so, we take the latter as the closer representation of the ancestral population, and assume that observed differences in diversity are mostly due to population contraction in the latter. Comparisons included Norway vs. Kirov and NE-Poland vs. Kirov for Eurasian lynx, and Doñana vs. Andújar for Iberian lynx. We used fix sized adjacent windows on a concatenation of units of each feature to avoid distortions caused by differences in unit length across features. Additionally, we subsampled populations within species to the same number of individuals (i.e. eight for Eurasian, and 12 for Iberian lynx) to avoid biases due to differences in sample size.

We filtered concatenated sites for each population and feature within each species so that it contained the same positions in the two populations being compared, and computed diversity for each population and feature in non-overlapping 1kb windows. We calculated differences in diversity over these non-overlapping 1kb windows to be able to compute statistical tests without the bias introduced by different features having different average length. Differences in diversity (δ_S) were computed as:

$$(B\ diversity - NB\ diversity)/(B\ diversity + NB\ diversity).$$

A positive δ_S means that B population shows more diversity than NB populations and vice versa. Window diversity, as computed as the arithmetic mean of site diversities estimated by ANGSD, shows a clearly bimodal distribution, with the lowest mode (usually 10-5 to 10-7) corresponding to windows with probably no diversity (Fig. S1). As the average diversity estimated for these windows with no diversity might still differ in the two populations, rendering irrelevant and probably biased δ_S, we

transformed $\delta_S\theta_W$ and $\delta_S\pi$ to zero when diversity in both populations dropped below their respective empirically determined diversity threshold separating the two modes aforementioned (Fig. S1). We plotted δ_S, and their dispersion calculated by bootstrapping over windows as implemented in Hmisc package in R (Harrell Jr, 2019) with 1,000 iterations. Also, in order to directly compare the magnitude of the changes in the density of variable sites with changes in their allele frequencies, we plotted $\delta_S\theta_W$ against $\delta_S\pi$. Due to non-normality of the δ_S statistic revealed by Shapiro test, we compared δ_S distribution among different features using a Wilcoxon signed-rank test for paired samples and applied a strict Bonferroni correction for multiple tests to assess significance.

Genomic variables

We annotated first units for the following genomic variables: average recombination rate, divergence – as a proxy of mutation rate –, GC content, and two different scores related to selection strength (RVIS, and functional site percentage). Then, non-overlapping 1kb windows used for calculating diversity differences were also annotated or these genomics variables by obtaining their overlap with the previously annotated units and averaging across overlapping units weighted by the number of overlapping sites.

For calculating recombination rate we used the domestic cat linkage map (Li et al., 2016), and cat-lynx synteny (Abascal et al., 2016), averaging the recombination rate over 2 MB window. Divergence was computed as the number of observed substitutions between Iberian lynx or Eurasian lynx and bobcat (L. rufus), divided by the number of informative sites in each unit, based on a genus-wide VCF. To estimate GC content per unit we used the Iberian lynx LYPA1.0 genome assembly, or a consensus Eurasian lynx genome (Abascal et al., 2016). Using BEDTools nuc (Quinlan & Hall, 2010) we parsed a bed file with units coordinates to the reference file, and estimated the GC percentage of each unit. The level of tolerance to variation of each gene was estimated by RVIS, a gene-based score computed for human sequence data that reaches more negative values for intolerant genes (Petrovski et al., 2013). Lynx genes were annotated using a lynx-to-human orthologs database (Abascal et al., 2016). For functional sites percentage we annotated the genome based on 2Mb

windows. For each 2Mb window we summed up the number of sites belonging to 3UTR, CDS, 5UTR, lncRNAs, ncRNAs, promoters, or UCNEs, divided it by 2Mb and multiplied it by 100. Windows were then intersected with units using bedtools intersect (Quinlan & Hall, 2010). Testing differences in genomic variables between windows with contrasting behaviour

First we explored correlations among chromosome size, gene content, recombination, divergence, GC content, RVIS and functional sites percentage. Chromosome length, gene content, and recombination data at the chromosome level were obtained from domestic cat, downloaded from https://www.ncbi.nlm.nih.gov/genome/?term=txid9685[orgn]. We plotted, for each of the populations studied, the correlation between diversity and chromosome size, gene content, recombination, divergence, GC content, RVIS and functional sites percentage at the chromosome level.

We then focused on genic –CDS and introns– and intergenic regions to assess to which extent differences in diversity between NB and B populations across windows of the same feature were related to differences in the following genomic variables: NB diversity (π, θ_W, and S) as a proxy of the ancestral diversity, recombination rate, RVIS and functional sites percentage as a proxy of selection, divergence as proxy of mutation rate, and GC content. After rejecting normality of these variables of interest, we used a Wilcoxon signed-rank test for unpaired samples to test whether there were significant differences in those variables between windows that showed higher diversity in B than NB population ($\delta_S\theta_W>0.1$) compared to the rest of the windows. Significance was corrected by applying a strict Bonferroni correction to the p-values. To get an estimation of the effect of these variables we calculated effect size (r), as implemented in Mangiafico (2020). As we found a generalized effect of NB diversity, and all genomic variables have a known relationship with diversity, we also compared windows with no θ_W diversity in either population, i.e. non-diversity non-diversity windows (ND_{NB}-ND_B windows), versus windows with θ_W diversity only in the B population (ND_{NB}-D_B windows) (Fig. S1). Again, we used a Wilcoxon signed-rank test for unpaired samples and Bonferroni correction to assess significant differences. All statistics were calculated using R (R Core Team, 2019) and results were plot using ggplot2 R package (Wickham, 2009).

RESULTS

Bottlenecks reduce genetic species diversity

At the species level, Eurasian lynx shows about twice the genome-wide diversity of Iberian lynx (Eurasian lynx: π weighted mean (wm)=3.3*10-4; $\theta_W$wm=2.8*10-4); Iberian lynx πwm=1.9*10-4; $\theta_W$wm=1.4*10-4) (Fig. 1). Both species are among the mammals with the lowest overall genomic diversity reported so far (Abascal et al., 2016; Lucena-Perez et al., 2020). Within Eurasian lynx, whole genome diversity statistics are πwm=3.4*10-4 and $\theta_W$wm=3.3*10-4 in the NB Kirov population, and these are reduced to 80% and 74% in the B NE-Poland and Norway populations, respectively, for both π and θ_W (Fig. 1, Table S2). In Iberian lynx, the NB Andújar population is the most diverse (πwm=2.5*10-4; $\theta_W$wm=1.8*10-4) and diversity is ~52% of this in the B Doñana population for both π and θ_W (Fig. 1, Table S2). Based on the relative diversity with respect to their most diverse population counterpart, bottlenecks rank from less to more extreme (longer duration and/or smaller Ne) in the order: NE-Poland, Norway, and Doñana. All populations show a negative S, an indication of a generalized scarcity of rare alleles, being more severe in Iberian lynx (average across populations of S=-0.41 vs. S=-0.28 in Eurasian lynx) (Table S2).

Bottlenecks lead to a larger accumulation of low frequency variants in the X chromosome

Chromosomes differ in diversity within populations (Fig. S2; Table S4). Largest diversity differences occur between A and the X chromosome, with the latter showing lower diversity than A consistently across populations (Fig. S2; Table S3, S4). Global X/A ratios range across populations from 0.31 to 0.37 for πX/πA, and from 0.36 to 0.84 for θ_WX/θ_WA (Table S3). While πX/πA is similar to θ_WX/θ_WA in NB populations and NE-Poland, θ_WX/θ_WA is larger than πX/πA in the B Doñana and Norway populations (Fig. 2). This pattern is consistent across different genomic features, namely CDS, intergenic and introns (Fig. 2). Also, the relative diversity reduction in B relative to NB populations is globally similar for autosomes and X chromosome (πB_X/πNB_X is similar to πB_A/πNB_A); however, the larger (above 1) θ_WB_X/

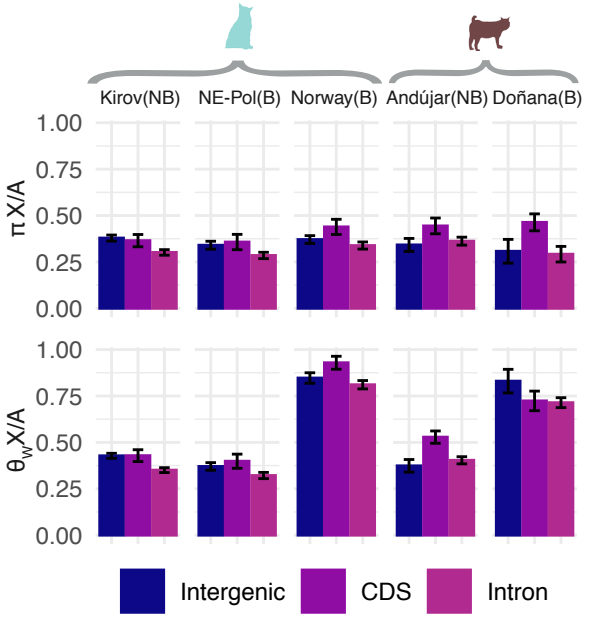


Fig. 2. X/A ratio for π and θ_W diversity for the five populations analysed, and three different genomic features (CDS, introns, and intergenic).

θ_WNB_X than θ_WB_A/θ_WNB_A for Norway and Doñana, when compared to Kirov and Andújar, respectively, indicates a larger accumulation of low frequency variants in the X chromosome in B relative to NB populations. Accordingly, S shows negative values in the NB Andújar and Kirov, but also in B NE-Poland populations, and positive in the B Doñana and Norway (Table S2).

Bottlenecks blur expected differences in diversity among and along autosomes

Among the autosomes, chromosome-wide genetic diversity is positively related to divergence, GC content, gene content, functional sites percentage, RVIS, and recombination, and negatively related to chromosome size in all populations of both species (Fig. S3, Table S4). However, these correlations show lower predictors and smaller fractions of explained variance (R^2) in the B populations, especially so for θ_W (Fig. S3).

Along autosomes, subtelomeric regions show the highest diversity, followed by the interstitial regions, and the pericentromeric regions (Fig. S4) in all populations. Subtelomeric regions show 1.52-1.66, and 1.29-1.62 of interstitial π and θ_W diversity,

respectively, with lower ratios of θ_W occurring in the B Norway and Doñana populations, indicating that θ_W diversity in B pops is still higher in subtelomeric regions than in other chromosomal regions, but not as high as π in B, or θ_W in NB populations (Table S5). Pericentromeric regions show 0.83-0.94 and 0.87-0.94 of interstitial π and θ_W diversity, respectively, both ratios tending to be higher in B populations that in their NB counterparts. Skewness of SFS is moderate and similar across chromosomal regions in NB pops, but the more extremely bottlenecked populations show more negative *S* in subtelomeric regions, indicating a relatively larger overall scarcity of low frequency alleles.

Diversity across chromosomal regions ranks according to recombination rates (subtelomeric=3.7cM/Mb, interstitial=3.5M/Mb, and pericentromeric=2.3cM/Mb), divergence (subtelomeric =7.9e-03, interstitial=5.5e-03, and pericentromeric=5.2e-03), and inversely to functional sites percentage (subtelomeric=3.3, interstitial=6.0, and pericentromeric=6.3). Also, subtelomeric regions show higher GC content (subtelomeric=0.58, interstitial=0.47 and pericentromeric=0.48), and higher RVIS (subtelomeric=0.2, interstitial=-0.31 and pericentromeric=-0.28) (Table S5).

Autosomal diversities in NB and B populations are generally correlated, but the more extreme the bottleneck, the weaker these correlations became both in terms of predictor and R^2 (Fig. S5), reflecting the random effects of genetic drift. In addition, these correlations are weaker in subtelomeric regions compared to pericentromeric or interstitial regions (Fig. S5).

Smaller relative diversity reductions in B populations in selectively constrained features

For all the features considered –namely: intergenic regions, coding gene promoters, 5'UTR, CDS, introns, 3'UTR, long-non-coding RNA promoters, long-non-coding RNA exons, long-non-coding RNA introns, non-coding RNA (mostly miRNAs, snRNAs and snoRNAs), and ultra-conserved-non-coding-elements (UCNE)) patterns of genetic diversity generally follows the anticipated selective pressure (Fig. 3, Table S6). Taking intergenic as a reference, the most selective constrained features present the lowest diversity, in particular UCNE (0.19-0.20 and 0.21-0.56 of intergenic π and θ_W, respectively), followed by CDS (π=0.58-0.6; θ_W=0.59-0.79; both relative to intergenic). Furthermore, regulatory regions such as gene promoters, non-coding

RNAs, but also 5'UTR and 3'UTR, and even introns, show smaller reductions of diversity in B populations with respect to NB than intergenic. *S* is negative for all populations and features, except for UCNE in Norway and Doñana, which shows a high positive value (S=0.53 and S=0.50, respectively), and for CDS in Norway (S=0.04), suggesting an accumulation of low frequency variants in these constrained features in the most bottlenecked populations (Table S6). Notably, despite the lower diversity in B populations than in NB populations in most features, diversity in UCNE in the B Doñana and Norway populations is actually higher than in their reference NB population (Andújar and Kirov, respectively) when measured as Watterson's theta (θ_W), but not as nucleotide diversity (π) (Fig. 3).

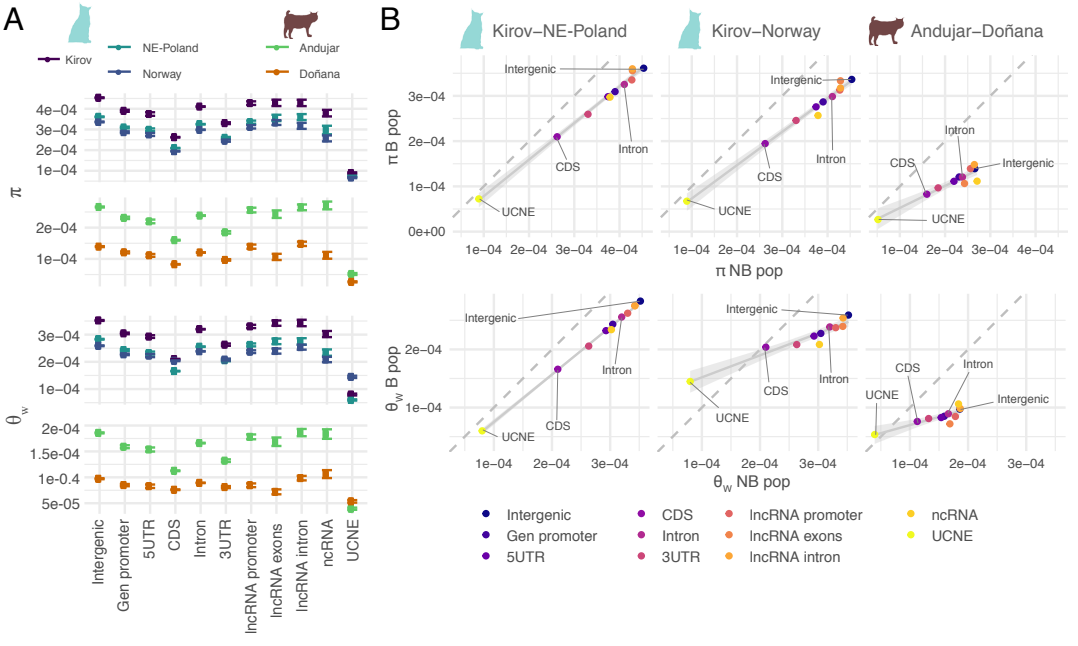


Fig. 3. A) Average θ_W and π diversity for different features. Error bars represent the standard deviation (stdev) obtained from bootstrapping. B) Relationship between diversity (θ_W or π) of the NB vs. B populations for different features.

Sliding window analysis of relative diversity differences between NB and B populations (δ_S) show the lower overall diversity in the B population as a negative global mean of δ_S for most features, both for π ($\delta_S\pi$) and θ_W ($\delta_S\theta_W$) (Fig. 4). For intergenic, Doñana-Andújar shows the more extreme diversity difference, followed by Norway-Kirov, and finally NE-Poland-Kirov. Regarding other features, and consistently among comparisons, UCNE's $\delta_S\theta_W$ and $\delta_S\pi$, are less negative, and $\delta_S\theta_W$ became even positive for the Doñana-Andújar and Norway-Kirov comparisons (Fig. 4). A similar trend of less negative δ_S values is also noticeable for CDS, and more subtly for 3'UTR, and 5'UTR, indicating a smaller diversity loss for selectively constrained genomic features in bottlenecked populations (Fig. 4).

Fig. 4. $\delta_S\theta_W$ and $\delta_S\pi$ across features for the three comparisons of B vs. NB populations. Positive values indicate higher diversity in the B population. The dashed lines represent the average δ_S value considering all features. Errors bars are the stdev obtained from bootstrapping the data. The coloured circle, triangles and squares represent significant p-values when compared to intergenic after Bonferroni correction (Table S7).

$\delta_S\theta_W$ and $\delta_S\pi$ are generally correlated, but $\delta_S\theta_W$ tends to be less negative than $\delta_S\pi$ for most features and in the most bottlenecked populations (Doñana and Norway). This pattern is particularly notable for UCNE followed by CDS, 3'UTR and 5'UTR. Although with lower δ_S than intergenic, ncRNA $\delta_S\theta_W$ tends to be larger than $\delta_S\pi$ in the Norway-Kirov comparison, and to a lesser extent in the Doñana-Andújar comparison (Fig. S6).

Mutational input and selection are related to changes in diversity in B populations

Windows with $\delta_S\theta_W > 0.1$, i. e. with larger θ_W in B than in NB populations, show significantly lower diversity in NB populations (π, θ_W) and larger *S* than the rest of windows (p-value < 3e-04; strict Bonferroni correction); this pattern is consistent across comparisons and features, except for CDS in NE-Poland-Kirov that shows the opposite pattern (Fig. S7, Table S8). The effect sizes of these variables measured as the correlation coefficient (r) are larger for Norway-Kirov (range: |0.15|-|0.22|), followed by Doñana-Andújar (range: |0.09|-|0.14|), and finally NE-Poland-Kirov comparison (|0.01|-|0.04|) (Table S9). Interestingly, the significant effects observed of other genomic variables like divergence, recombination, or selective constraints go in the opposite direction to that expected from their positive correlation with diversity (Fig. S7, Table S8). For instance, windows with $\delta_S\theta_W > 0.1$ tend to show lower NB diversity, but higher divergence, GC content, recombination, and RVIS. To further test the effect of other genomic variables while eliminating the influence of NB diversity, we considered only those windows with no θ_W diversity in NB populations and within this subset we compared windows that also lack θ_W diversity in the B population (ND_{NB}-ND_B windows) (Fig. 5) with those that do show some θ_W diversity in the B population (ND_{NB}-D_B windows). Some of the previous contrasts, which were only significant for NE-Poland-Kirov when using all windows, become now significant in all three comparisons, although with small effect sizes (Fig. 5, Table S10, S11). For instance, ND_{NB}-D_B windows are in regions with higher divergence and GC content than ND_{NB}-ND_B windows. Also, recombination rate is significantly higher in ND_{NB}-D_B windows, for non-selectively constrained features in the Norway-Kirov (introns), and Doñana-Andújar comparisons (intergenic and introns). Regarding genomic variables related to selection, RVIS is higher in ND_{NB}-D_B windows in NE-Poland-Kirov for both CDS and introns and also in introns for

the Norway-Kirov comparison, i.e. genes with diversity only in the B population tend to be less intolerant to change. This difference disappears in the most extreme Doñana-Andújar comparison, where genes in ND_{NB}-ND_B windows are on average less intolerant to change than in other comparisons, and not significantly different in tolerance to ND_{NB}-D_B genes. The percentage of functional sites is smaller in ND_{NB}-D_B windows for non-selectively constrained features in NE-Poland-Kirov (intergenic and introns) and Doñana-Andújar (introns). Differences in genomic variables between ND_{NB}-ND_B and ND_{NB}-D_B windows are smaller overall when the bottleneck is more intense (Norway and Doñana) and the features are selectively constrained (CDS), as expected.

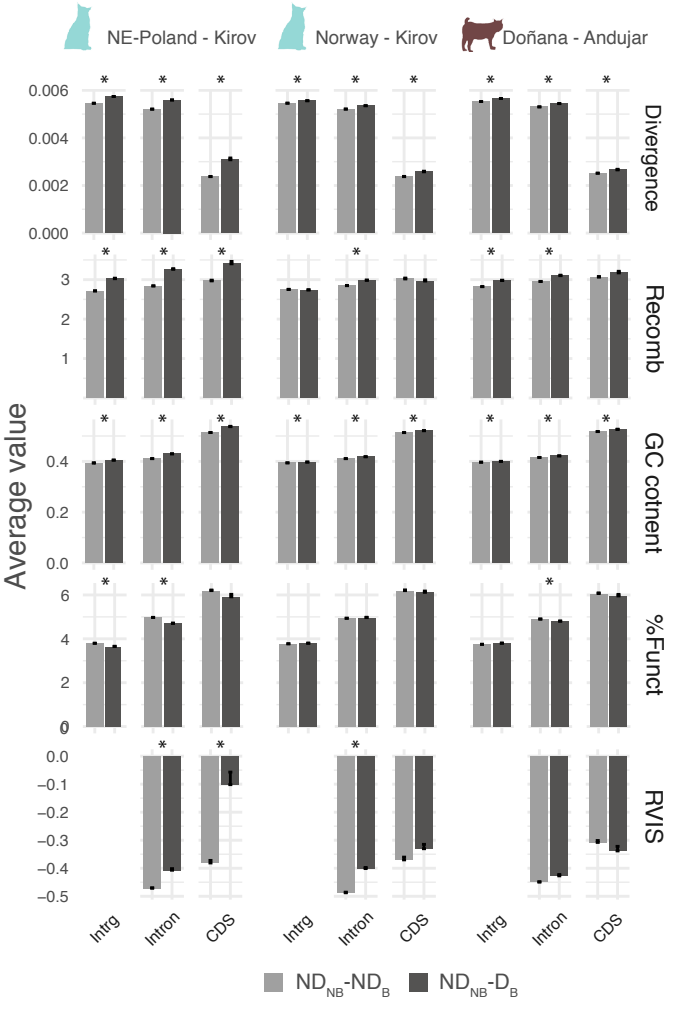


Fig. 5. Average values of different genomic variables for ND_{NB}-ND_B windows vs. ND_{NB}-D_B. The asterisk represents significant differences between both categories.

DISCUSSION

Here we present one of the most comprehensive empirical studies of the genomic consequences of bottlenecks in natural populations. By directly comparing genomic diversity in Eurasian and Iberian lynx populations with contrasting recent demography, we assessed how the impact of recent bottlenecks on population genetic diversity varies across chromosomes, chromosomal regions, and genomic features. Also, we explored how different genomic variables, such as mutation and selection, as well as the initial SFS, might be affecting the different outcomes of genetic drift and, in consequence, the behaviour of different genomic features and regions under a demographic decline.

As expected, recent bottlenecks globally reduce the genomic diversity in a magnitude following the intensity and duration of the bottleneck, as inferred from records and demographic reconstructions from genetic and genomic data (Abascal et al., 2016; Casas-Marce et al., 2017; Lucena-Perez et al., 2020). Skewness of the SFS is negative in all populations, indicating a general scarcity of low frequency variants, which is however more pronounced in the Iberian lynx populations, consistent with the expectation for its smaller effective size at mutation-drift equilibrium.

Diversity reduction in the B populations is conspicuous across the genome, but its magnitude varied across chromosomes, chromosomal regions and genomic features. First, overall correlations between the genetic diversities in NB and B populations across chromosomes and chromosomal regions weaken as bottleneck intensity increases, especially in subtelomeric regions. The correlations of genetic diversity with genomic variables like recombination, divergence, or gene content, which are expected at migration-mutation-selection-drift equilibrium and still observed in the NB populations, also weaken or disappear in the B populations, and are especially low in the Iberian lynx comparison. Both observations indicate that, in B populations, the reduction in mean diversity is accompanied by an increased variance across chromosomes and chromosomal regions, and that genetic diversity patterns progressively depart from those expected at equilibrium.

Regarding the X and A comparison, results suggest a larger accumulation of low frequency alleles in the X chromosome relative to autosomes (X/A ratios) in B populations. The X chromosome differ from the A in their response to a bottleneck

because: i) its effective size is ¾ of A, which predicts a ratio of diversity X/A of 0.75 at mutation-drift equilibrium; ii) it has a lower recombination rate; iii) it has a higher mutation rate, and iv) recessive deleterious variation is exposed to selection in hemizygote individuals (i.e. males in mammals) (Arbiza et al., 2014). Besides, under population size change scenarios, the dynamics of the two types of chromosomes might differ, with X diversity undergoing faster changes than autosomal diversity (Pool & Nielsen, 2007). It is unclear which of these distinctive characteristics is responsible for the observed marked accumulation of low frequency variants in X in B populations; however, we suggest a combination the relaxation of selection and a higher mutational input as the main drivers, as discussed below for selectively constrained features.

An accumulation of low frequency alleles is also observed in some selectively constrained features, such as CDS and UCNE in the autosomes. Features that are subject to unequal selective pressures behave very differently in populations that underwent different degrees of population size decline. Whereas some features, like intergenic regions, ncRNA and promoters show a similar relative reduction of θ_W and π diversity in B populations, UCNE and CDS and, to a lesser extent, 5' and 3' UTR, show a relatively smaller reduction in θ_W than in π, which in the extreme case of UCNE results in higher θ_W in the B than in the corresponding NB populations. The increase in θ_W in B populations in selectively constrained features is, in principle, unexpected from purely random processes, since random genetic drift in bottlenecked populations would result in the preferential loss of alleles in the low-frequency side of the spectrum, where most selectively constrained sites will be (Nei & Maruyama, 1975). The pattern observed for the X chromosome and some selectively constrained features could thus be the consequence of different, non-exclusive and interacting processes that include genetic drift, mutation and selection.

First, it must be noted that the estimates of genetic diversity, and especially θ_W diversity, are conditioned by both the sample sizes and the SFS. Given our sample sizes of 16 and 24 genomes, alleles segregating at frequencies lower than these in the population will be underrepresented in our limited size samples. One consequence is that their likely loss in bottlenecked populations will remain largely unnoticed. This is especially true for those regions and features with the highest skew in SFS, i.e. the most selectively constrained features like UCNE and CDS.

On the other hand, the possibility of observing an increase in θ_W diversity following a bottleneck will also depend on the SFS and thus on starting diversity. Indeed, we found that windows within features with higher diversity in B than in NB tend to show lower diversity in the NB population. It must be noted thought that the correlation of NB diversity with δ_S might be spurious, as NB diversity is included in the calculation of δ_S, and to some extent trivial, as it suggests that windows with higher diversity tend to lose diversity whereas those with low diversity are more likely to gain diversity in B populations. More importantly, as other genomic variables tested tend to be themselves correlated with equilibrium diversity, such correlations may create indirect correlations of δ_S with these when effects go in the same direction, or weaken the correlations when they go in opposite directions. However, results suggest that the effect of these variables on δ_S go in the opposite direction to that expected from their correlation with equilibrium diversity, and they are thus reinforced when the effect of diversity is neglected.

That divergence and GC content tend to be higher in windows with larger diversity in B relative to NB populations could be directly related to a higher mutational input increasing the density of mutations at low frequency in these windows, Increased mutational input in these regions could have originated a larger number of alleles segregating at low frequencies in the pre-bottleneck population and favour the accumulation of de novo mutations during the bottleneck. Given the relatively short separation of the populations (i.e. likely less than 60 generations in the case of the Eurasian lynx (Lucena-Perez et al., 2020), and around 40 generations in Iberian lynx (Casas-Marce et al., 2017), assuming a generation time of 5 years (Lucena-Perez et al., 2018), and their small size in the range of tens, it is unlikely that a large fraction of the observed accumulation of variants at low frequencies in the B populations is due to de novo mutations since the start of the bottleneck. Thus, most of these alleles increasing in frequency during and after the bottleneck would be pre-existing alleles that remain undetected in the NB populations and that are now recorded and result in increased θ_W in the B population.

Recombination could also contribute to this pattern through its direct mutagenic effect (Duret & Arndt, 2008; Duret & Galtier, 2009; Pratto et al., 2014; Williams et al., 2015; Terekhanova et al., 2017; Smith, Arndt & Eyre-Walker, 2018; Halldorsson et al., 2019), but also through its interaction with selection. Higher recombination will facilitate the increase in frequency of moderately deleterious variants that behave as neutral in the bottlenecked population by breaking possible interference with more selected sites (Hill & Robertson, 1966; Felsenstein, 1974).

The observed accumulation of θ_W in selectively constrained regions could be in part due to the accumulation of additive deleterious variants due to the relaxation of purifying selection (Kirkpatrick & Jarne, 2000; Balick et al., 2015). Empirical evidence for this has been found in domesticated species (Marsden et al., 2016; Makino et al., 2018), while evidence in humans remains controversial (Lohmueller, 2014; Simons et al., 2014; Do et al., 2015; Henn et al., 2016; Simons & Sella, 2016). The relaxation of purifying selection would affect those variants with $1/2Ne\_NB < s < 1/2Ne\_B$, i.e. alleles with s between these values would be efficiently selected in the NB but will behave as neutral in the B population; for example, in the case of Andújar ($Ne = 20$) and Doñana ($Ne = 10$) those would have $0.025 < s < 0.05$. However, it must be noted that while Doñana has remained at that low size for decades, Andújar was much larger a few generations ago, so that the range of s that behave as neutral in Doñana and not in Andújar may has a much lower inferior limit (Casas-Marce et al., 2017). The contribution of the relaxation of purifying selection is further supported by the observation that in shallower demographic declines the smaller relative loss occurs mainly in genes with high RVIS (i.e. more tolerant to changes). However, with more extreme bottlenecks the relationship with RVIS weakens and becomes non significant, suggesting that the relaxation of purifying selection results in the accumulation of potentially deleterious variants even in highly conserved genes, and alerting of possible fitness reductions.

Furthermore, empirical and theoretical studies showed that selectively constrained diversity approaches a new equilibrium faster than neutral diversity (Gordo & Dionisio, 2005; Song & Steinrücken, 2012; Pennings, Kryazhimskiy & Wakeley, 2014; Brandvain & Wright, 2016). Under this assumption, we might expect that the CDS and UCNE diversity reach their increased level expected for the new size faster than intergenic or intronic diversity reach their reduced new equilibrium level.

Possible consequences for the B populations

Our study provides empirical evidences of the conspicuous action of drift across the genome in replicated natural bottleneck scenarios. Comparing three bottlenecks of different intensity allow us to assess the consistency of the results across replicates, as well as to account for differences regarding drift intensity.

Our results suggest that overall diversity reductions in bottlenecked populations affect regulatory elements, such as promoters, or ncRNA, in a similar fashion as

intergenic regions. This loss of diversity is likely removing pre-existing variation in the regulation of gene expression. At the population level, the loss of variation in regulatory elements of transcription likely compromises these populations' potential for acclimation to the new environments through phenotypic plasticity.

For other selectively constrained features, our results suggest that overall genetic diversity reductions in bottlenecked populations are in part counterbalanced by the accumulation of possible deleterious alleles in functional regions, including genes highly intolerant to changes (in the most extreme bottlenecks), and extending beyond coding regions to include UCNE and to a lesser extent 5'UTR and 3'UTR. These features are essential for efficient transcription and for the post-transcriptional regulation of gene expression, and variants in both regions have been linked to several diseases (Hindorff et al., 2009). Regarding UCNE, an accumulation of deleterious variation in these regions is likely to affect their prominent function as regulatory cis elements, especially during development (Marcovitz, Jia & Bejerano, 2016; Polychronopoulos et al., 2017), leading to a reduction in fitness. CDS results show a similar scenario.

Still, we acknowledge that reduced fitness of the B populations cannot be directly inferred from our observations, as the increase in frequency of some deleterious variants might be compensated by the loss of many low frequency ones due to drift. Also, population contractions allow a more efficient purging of recessive deleterious alleles, which may compensate the fitness reductions due to the accumulation of additive alleles. Finally, our results call for extending the analyses of genetic variation in endangered species to include functional variation beyond coding sequences. Including regulatory elements will broaden our understanding of the effects of drift on the genome and allow for a more comprehensive assessment of the possible fitness reductions in short term and the loss of adaptive potential in the long term.

AUTHOR CONTRIBUTION

J.A.G. conceived the project and designed the study; A.P.S., J.A.G., K.S., provided samples; E.M., and M.L.P. performed the lab work; M.L.P. analysed data; E.M. provided help coding; D.K.R. helped calculating divergence rate and RVIS; J.A.G., and M.L.P. interpreted the results; M.L.P. drafted the manuscript with support from J.A.G. and critical input from D.K.R., A.P.S., K.S.; J.A.G. supervised the project. All authors approved the final version of the manuscript.

ACKNOWLEDGEMENTS

This project was funded by the National Science Center, Poland through project no 2014/15/B/ NZ8/00212, project "Biodiversity of East-European and Siberian large mammals on the level of genetic variation of populations" (BIOGEAST), contract No. PIRSESGA-2009-247652, and by the Dirección General de Investigación Científica y Técnica of the Spanish Ministerio de Ciencia e Innovacion (MICINN) through projects CGL2013-47755-P and CGL2017-84641-P, and by Russian Science Foundation (project No. 18-14-00093). MLP and DKR were supported by PhD contracts from the Programa Internacional de Becas "La Caixa-Severo Ochoa" from "La Caixa" Foundation (ID 100010434), under agreement LCF/BQ/SO14/52250035 and LCF/BQ/SO15/52260006. Also, DKR was supported by the Dirección General de Investigación Científica y Técnica of the Spanish Ministerio de Ciencia e Innovación by the PGC2018-095810-B-I00 project to Aurora Garcia-Dorado. Logistical support was provided by Laboratorio de Ecología Molecular (LEM-EBD) (certified to ISO9001:2015 and ISO14001:2015 quality and environmental management systems), and Laboratorio de SIG y Teledetección (LAST-EBD) in Estación Biológica de Doñana. Calculations were carried out in the Genomics servers of the Doñana's Singular Scientic-Technical Infrastructure (ICTS-RBD). Authors also acknowledge the computing and storage resources provided by Fundación Pública Galega Centro Tecnolóxico de Supercomputación de Galicia (CESGA). EBD-CSIC received support from the Spanish Ministry of Economy and Competitiveness under the 'Centro de Excelencia Severo Ochoa 2013-2017' program, grants SEV-2012-0262. We also acknowledged Jerez Zoo (Spain) for providing samples.

Lynx pardinus samples were collected following the licenses adhere to the Directive 2010/63/EU on the protection of animals used for scientific purpose. The license for Eurasian lynx live-trapping and blood sampling in Poland was obtained from the National Ethics Committee for Animal Experiments (no. DB/KKE/PL—110/2001) and the Local Ethics Committee for Animal Experiments at the Medical University of Białystok, Poland (no. 52/2007). Import of samples was licensed by CITES permissions no 12RU000512 and 15NO-046-EX, respectively. No animals were harmed during live-trapping and handling. We are grateful to Ingrid Reinkind for kind providing the Eurasian lynx samples from Norway. We also want to thank Enrico Bazzicalupo, Pedro Jordano, Rubén Bernardo-Madrid, Isabel García-Barón and Elena del Águila-Pérez for insightful discussion about the analysis.

Besides the free software mentioned in the text, this work was possible thanks to the free, open source packages: dplyr (Wickham et al., 2017), ggplot2 (Wickham, 2009), tidyr (Wickham y Henry, 2018), broom (Robinson, 2018), RColorBrewer (Neuwirth, 2014), stringr (Wickham, 2018), gridExtra (Auguie, 2017), knitr (Xie, 2014, 2015, 2017), reshape2 (Wickham, 2007), lemon (Edwards, 2017), data.table (Dowle y Srinivasan, 2017), rlang (Henry y Wickham, 2018), png (Urbanek, 2013), rgdal (Bivand, Keitt y Rowlingson, 2017).

04 APÉNDICE

BOTTLENECKED-ASSOCIATED CHANGES IN THE GENOMIC LANDSCAPE OF GENETIC DIVERSITY IN WILD LYNX POPULATIONS

SUPPLEMENTARY FIGURES

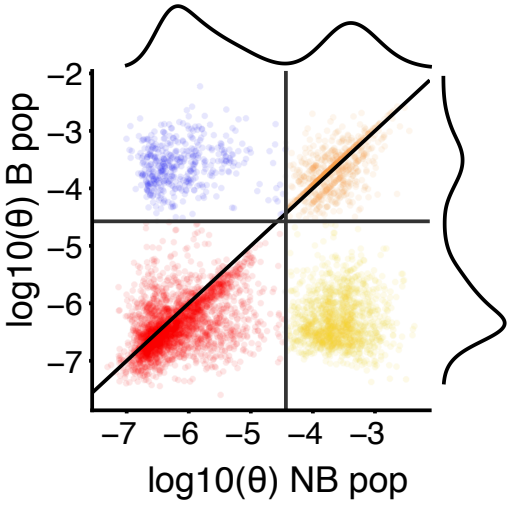


Fig. S1. B vs NB diversity plot and quadrant definition. The distribution of window diversity within populations, as estimated by ANGSD, is markedly bimodal with one mode corresponding to units that show almost no diversity (i.e. ~10-6), and a second mode, with units that do hold some diversity (i.e. ~10-4). Therefore, in a graph plotting diversity of B vs. NB we can define four quadrants: units with no diversity in both populations (ND-ND) vs. units with no diversity in the NB population but showing diversity in the B population (ND-D), units that show diversity in both populations (D-D), and units that show diversity in the NB population but no diversity in the B one (D-ND). For some of our analyses we compare the value of several genomic variables between windows in ND-ND quadrants versus windows in the ND-D quadrants.

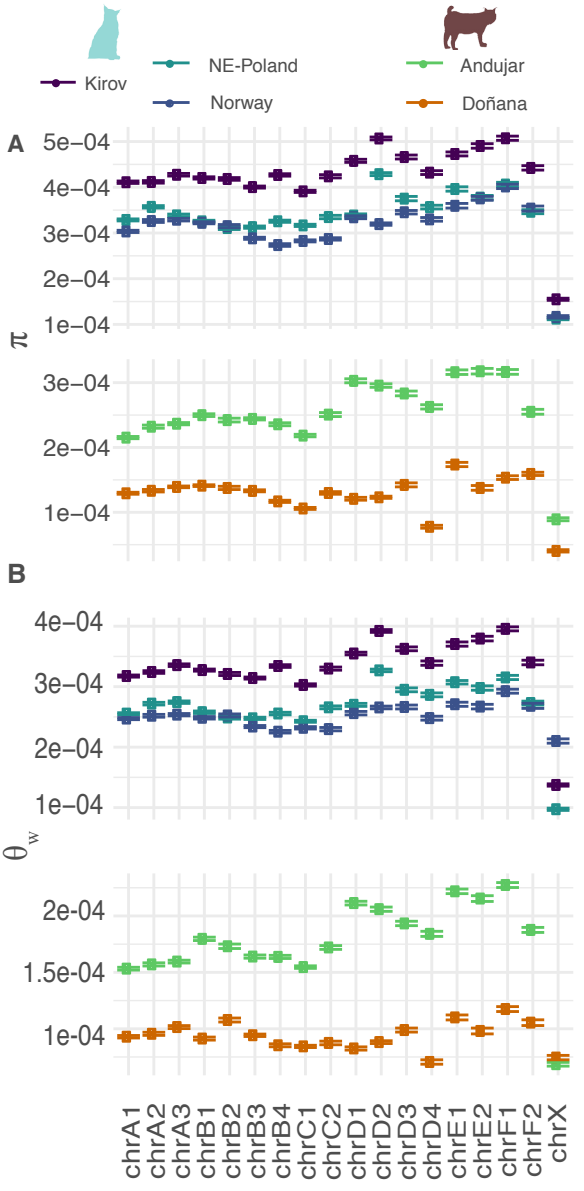


Fig. S2. Average θ_w and π diversity for different chromosomes. Error bars are standard deviation (stdev) obtained from bootstrapping.

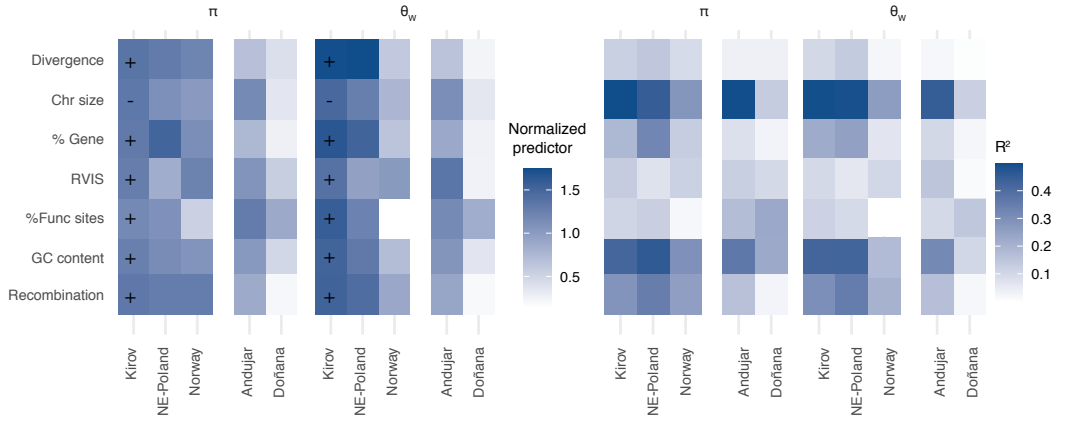


Fig. S3. Heatmap representing correlation between different genomic variables and θ_W and π diversity for the 5 populations studied. On the left the predictor normalized accross variables for comparison purposes across populations (normalized predictor=predictor value / mean predictor value). On the right R^2 values.

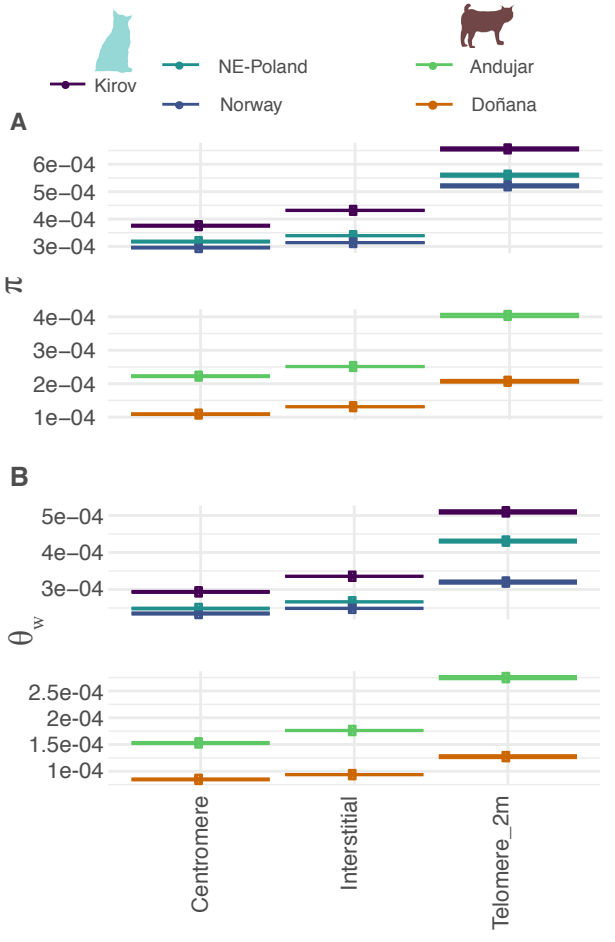


Fig. S4. Average θ_W and π diversity for different chromosomal regions. Error bars are standard deviation (stdev) obtained from bootstrapping.

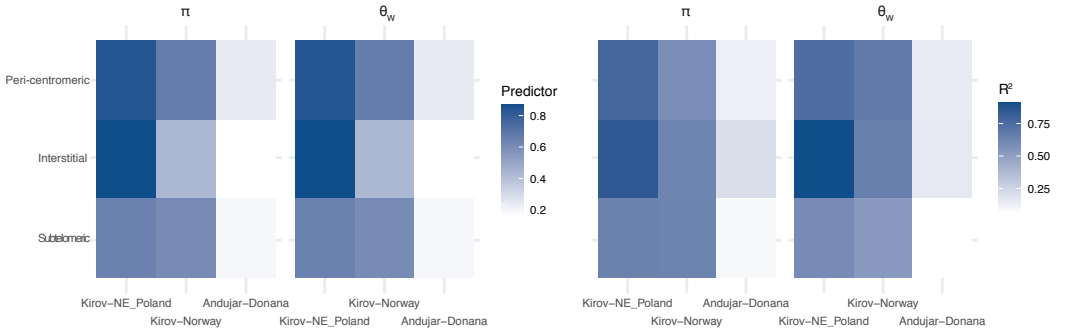


Fig. S5. Heatmap representing correlation between different genomic regions and θ_W and π diversity for the 5 populations studied. On the left the predictor normalized accross regions for comparison purposes across populations (normalized predictor=predictor value / mean predictor value). On the right R^2 values.

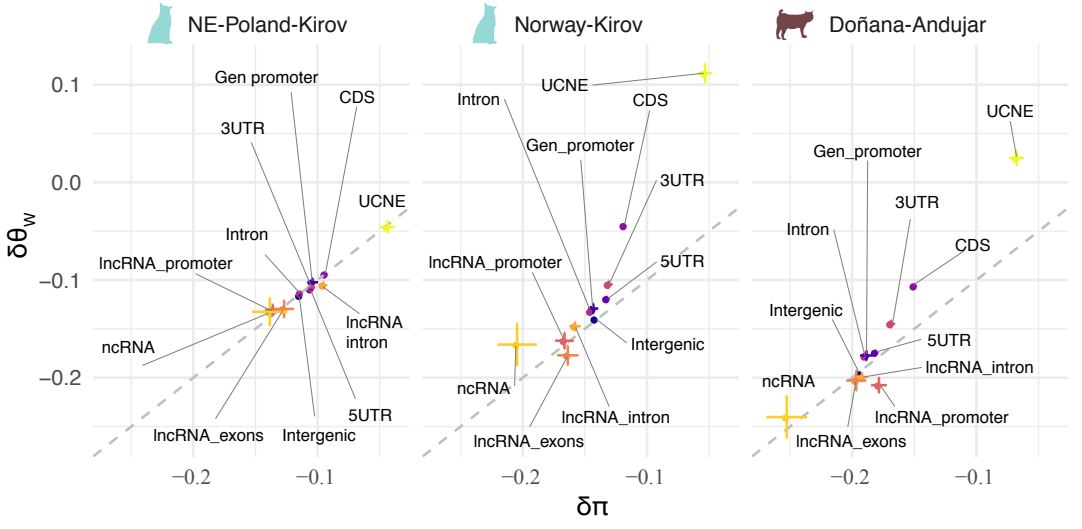


Fig. S6. Relationship between $\delta_s\theta_w$ vs. $\delta_s\pi$ for each feature for different B-NB comparisons.

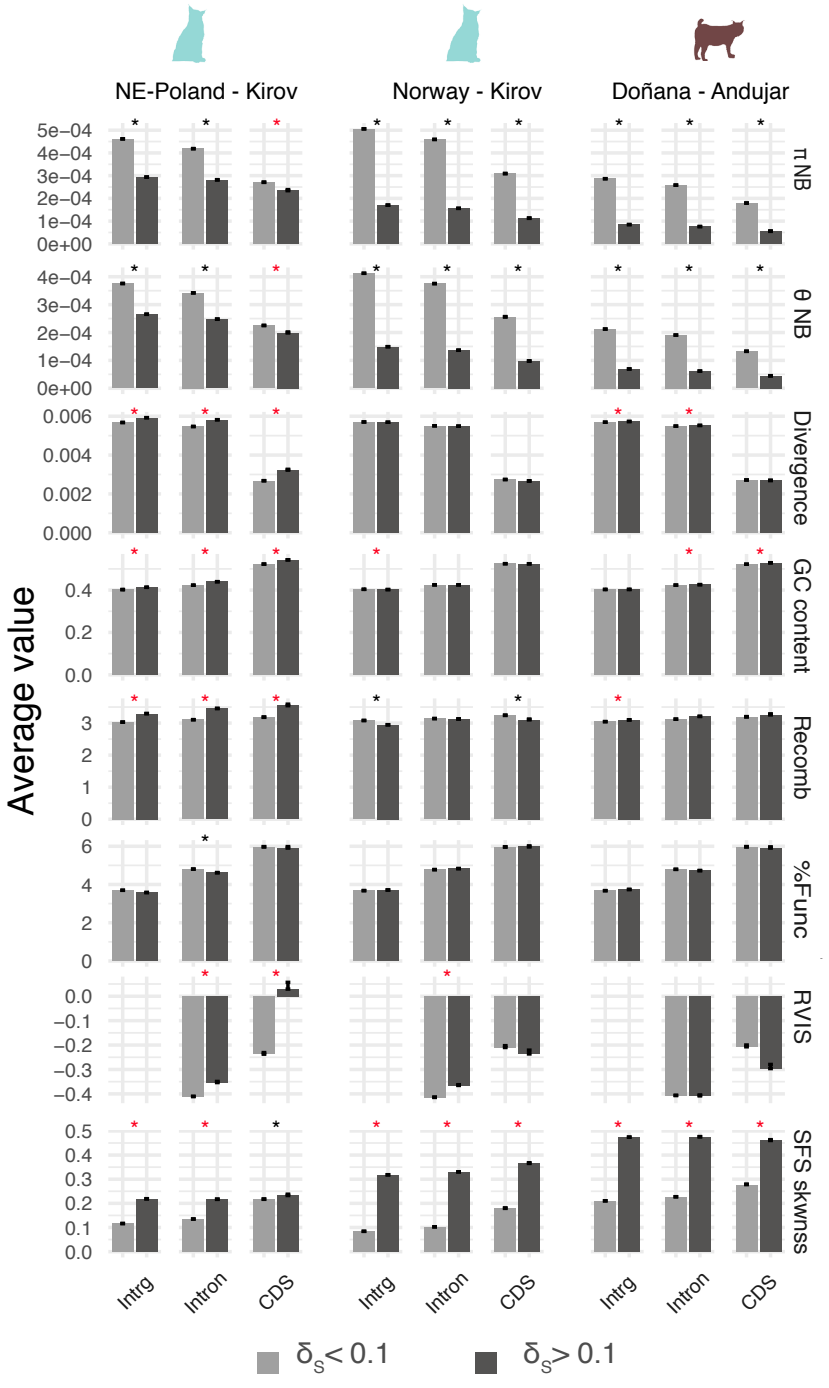


Fig. S7. Average values for different genomic variables for windows with $\delta_S\theta_W > 0.1$ vs. windows with $\delta_S\theta_W < 0.1$. Black asterisk represents significant differences between both categories where $\delta_S\theta_W < 0.1$ is significantly higher, while red asterisk represents the opposite. Note that even thought the mean value could be higher, sometimes the test is significant for the opposite direction.

SUPPLEMENTARY TABLES

Table S1. Detailed information about the Eurasian and Iberian lynx samples analysed in this study.

Sample	Specie	Population	WG depth used	Source	Sex	Sampling year (L.l.) or assinged birth year (L.p.)
c_ll_no_0075	Lynx lynx	Norway	5,43	Lucena-Perez et al. 2020	M	2008
c_ll_no_0076	Lynx lynx	Norway	5,80	Lucena-Perez et al. 2020	M	2008
c_ll_no_0077	Lynx lynx	Norway	5,14	Lucena-Perez et al. 2020	M	2008
c_ll_no_0078	Lynx lynx	Norway	5,25	Lucena-Perez et al. 2020	M	2008
c_ll_no_0079	Lynx lynx	Norway	5,09	Lucena-Perez et al. 2020	M	2008
c_ll_no_0080	Lynx lynx	Norway	5,22	Lucena-Perez et al. 2020	F	2008
c_ll_no_0081	Lynx lynx	Norway	5,54	Lucena-Perez et al. 2020	F	2008
c_ll_no_0082	Lynx lynx	Norway	5,36	Lucena-Perez et al. 2020	M	2008
c_ll_po_0105	Lynx lynx	NE-Poland	6,06	Lucena-Perez et al. 2020	M	2004
c_ll_po_0106	Lynx lynx	NE-Poland	6,11	Lucena-Perez et al. 2020	F	2006
c_ll_ki_0090	Lynx lynx	Kirov region	6,57	Lucena-Perez et al. 2020	F	2011
c_ll_ki_0091	Lynx lynx	Kirov region	6,11	Lucena-Perez et al. 2020	M	2011
c_ll_ki_0092	Lynx lynx	Kirov region	5,32	Lucena-Perez et al. 2020	F	2011
c_ll_ki_0093	Lynx lynx	Kirov region	5,29	Lucena-Perez et al. 2020	M	2011
c_ll_ki_0094	Lynx lynx	Kirov region	5,85	Lucena-Perez et al. 2020	M	2011
c_ll_ki_0095	Lynx lynx	Kirov region	5,29	Lucena-Perez et al. 2020	M	2011
c_ll_ki_0096	Lynx lynx	Kirov region	6,12	Lucena-Perez et al. 2020	M	2011
c_ll_ki_0097	Lynx lynx	Kirov region	6,03	Lucena-Perez et al. 2020	F	2011
c_ll_ki_0098	Lynx lynx	Kirov region	5,78	Lucena-Perez et al. 2020	F	2011
c_ll_ki_0099	Lynx lynx	Kirov region	6,01	Lucena-Perez et al. 2020	F	2011
c_ll_ki_0100	Lynx lynx	Kirov region	6,17	Lucena-Perez et al. 2020	M	2011
c_ll_ki_0101	Lynx lynx	Kirov region	6,30	Lucena-Perez et al. 2020	F	2011
c_ll_ki_0102	Lynx lynx	Kirov region	6,07	Lucena-Perez et al. 2020	M	2011
c_ll_po_0001	Lynx lynx	NE-Poland	6,52	Lucena-Perez et al. 2020	M	2001
c_ll_po_0002	Lynx lynx	NE-Poland	6,61	Lucena-Perez et al. 2020	F	1994
c_ll_po_0003	Lynx lynx	NE-Poland	6,44	Lucena-Perez et al. 2020	F	2002
c_ll_po_0011	Lynx lynx	NE-Poland	6,45	Lucena-Perez et al. 2020	M	2007
c_ll_po_0014	Lynx lynx	NE-Poland	6,39	Lucena-Perez et al. 2020	M	2008
c_ll_po_0019	Lynx lynx	NE-Poland	6,38	Lucena-Perez et al. 2020	M	2010
c_lp_do_0007	Lynx pardinus	Doñana	6,06	Abascal et al. 2016	M	1987
c_lp_do_0141	Lynx pardinus	Doñana	4,99	This article	M	2001
c_lp_do_0144	Lynx pardinus	Doñana	6,03	This article	M	2000

c_lp_do_0153	Lynx pardinus	Doñana	6,36	Abascal et al. 2016	M	2004
c_lp_do_0162	Lynx pardinus	Doñana	5,84	This article	F	2002
c_lp_do_0163	Lynx pardinus	Doñana	5,57	This article	F	2001
c_lp_do_0173	Lynx pardinus	Doñana	6,30	Abascal et al. 2016	M	2003
c_lp_do_0300	Lynx pardinus	Doñana	5,61	This article	M	2007
c_lp_do_0333	Lynx pardinus	Doñana	6,14	This article	F	2008
c_lp_do_0335	Lynx pardinus	Doñana	5,19	This article	F	2008
c_lp_do_0443	Lynx pardinus	Doñana	6,42	Abascal et al. 2016	M	2010
c_lp_do_0444	Lynx pardinus	Doñana	5,27	This article	F	2010
c_lp_sm_0134	Lynx pardinus	Andujar	5,89	This article	F	2004
c_lp_sm_0138	Lynx pardinus	Andujar	6,61	Abascal et al. 2016	M	2002
c_lp_sm_0140	Lynx pardinus	Andujar	6,24	Abascal et al. 2016	M	2000
c_lp_sm_0155	Lynx pardinus	Andujar	5,50	This article	F	2004
c_lp_sm_0156	Lynx pardinus	Andujar	6,02	This article	F	2004
c_lp_sm_0161	Lynx pardinus	Andujar	5,26	This article	F	2002
c_lp_sm_0185	Lynx pardinus	Andujar	6,30	Abascal et al. 2016	M	2005
c_lp_sm_0186	Lynx pardinus	Andujar	6,46	Abascal et al. 2016	M	2004
c_lp_sm_0206	Lynx pardinus	Andujar	5,66	This article	F	2006
c_lp_sm_0208	Lynx pardinus	Andujar	5,32	This article	F	2005
c_lp_sm_0213	Lynx pardinus	Andujar	5,46	This article	F	2005
c_lp_sm_0221	Lynx pardinus	Andujar	7,60	This article	M	2006
c_lp_sm_0226	Lynx pardinus	Andujar	5,59	This article	M	2006
c_lp_sm_0276	Lynx pardinus	Andujar	5,55	This article	F	2006
c_lp_sm_0298	Lynx pardinus	Andujar	6,88	Abascal et al. 2016	M	2007
c_lp_sm_0320	Lynx pardinus	Andujar	5,57	This article	F	2008
c_lp_sm_0325	Lynx pardinus	Andujar	5,40	This article	F	1999
c_lp_sm_0359	Lynx pardinus	Andujar	6,95	Abascal et al. 2016	M	2004
c_lp_sm_0450	Lynx pardinus	Andujar	5,48	This article	M	2009

Table S2. Weighted genomic diversity (watterson estimator and nucleotide divesity), Tajima's D and SFS skewness (S) for the different populations

Species	Population	Weighted mean Watterson estimator	Stdv boot Watterson estimator	Weighted mean nucleotide diversity	Stdv boot nucleotide diversity	Weighted Tajima's D	SFS Skewness	Type chr
L.lynx	Kirov	3,35E-04	6,23E-07	4,30E-04	8,38E-07	1,41E-02	-0,286	Autosomes
L.lynx	NE-Poland	2,68E-04	5,95E-07	3,42E-04	8,03E-07	1,41E-02	-0,275	Autosomes
L.lynx	Norway	2,49E-04	6,05E-07	3,17E-04	7,07E-07	7,25E-03	-0,273	Autosomes
L.pardinus	Andujar	1,76E-04	4,37E-07	2,52E-04	6,69E-07	1,39E-02	-0,432	Autosomes
L.pardinus	Doñana	9,34E-05	3,77E-07	1,30E-04	4,59E-07	5,42E-03	-0,393	Autosomes
L.lynx	Kirov	1,37E-04	2,05E-06	1,55E-04	2,74E-06	5,38E-03	-0,128	chrX
L.lynx	NE_Poland	9,72E-05	1,88E-06	1,13E-04	2,51E-06	9,70E-03	-0,160	chrX
L.lynx	Norway	2,10E-04	3,79E-06	1,17E-04	2,81E-06	-5,40E-03	0,445	chrX
L.pardinus	Andujar	6,86E-05	1,64E-06	8,92E-05	2,50E-06	8,71E-03	-0,301	chrX
L.pardinus	Doñana	7,45E-05	1,92E-06	4,05E-05	1,69E-06	-5,39E-03	0,457	chrX
	L.lynx sps.	2,84E-04		3,63E-04		1,18E-02	-0,278	Autosomes
	L.pardinus sps.	1,35E-04		1,91E-04		9,64E-03	-0,412	Autosomes

Table S3. X/A ratio for watterson estimator and nucleotide diversity, and associated error.

Populations	Ratio X/A Watterson estimator	Error ratio X/A Watterson	Ratio X/A Nucleotide diversity	Error ratio X/A Nucleotide diversity
Kirov	0,410	0,006	0,360	0,006
NE-Poland	0,362	0,007	0,330	0,007
Norway	0,844	0,015	0,368	0,009
Andujar	0,390	0,009	0,355	0,010
Doñana	0,797	0,021	0,311	0,013

Table S4. Weighted genomic diversity (watterson estimator and nucleotide divesity), Tajima's D, S, mean recombination rate, divergence, GC content, RVIS, and functional sites percentage for different chromosomes.

Species	Populations	Chr	Number of units	Weighted mean Watterson estimator	Weighted mean nucleotide diversity	Weighted Tajima's D	SFS Skewness
L.lynx	Kirov	chrA1	25888	3,18E-04	4,11E-04	1,40E-02	-0,295
L.lynx	Kirov	chrA2	30793	3,24E-04	4,12E-04	1,28E-02	-0,270
L.lynx	Kirov	chrA3	22859	3,36E-04	4,27E-04	1,39E-02	-0,274
L.lynx	Kirov	chrB1	20444	3,27E-04	4,20E-04	1,40E-02	-0,284
L.lynx	Kirov	chrB2	19676	3,21E-04	4,18E-04	1,49E-02	-0,303
L.lynx	Kirov	chrB3	25096	3,14E-04	4,00E-04	1,29E-02	-0,274
L.lynx	Kirov	chrB4	24526	3,34E-04	4,27E-04	1,37E-02	-0,278
L.lynx	Kirov	chrC1	37667	3,03E-04	3,91E-04	1,37E-02	-0,292
L.lynx	Kirov	chrC2	18312	3,30E-04	4,24E-04	1,36E-02	-0,285
L.lynx	Kirov	chrD1	21637	3,55E-04	4,58E-04	1,43E-02	-0,290
L.lynx	Kirov	chrD2	13624	3,92E-04	5,07E-04	1,56E-02	-0,292
L.lynx	Kirov	chrD3	14538	3,63E-04	4,66E-04	1,49E-02	-0,286
L.lynx	Kirov	chrD4	15194	3,39E-04	4,32E-04	1,33E-02	-0,275
L.lynx	Kirov	chrE1	20072	3,70E-04	4,73E-04	1,43E-02	-0,277
L.lynx	Kirov	chrE2	15971	3,79E-04	4,90E-04	1,57E-02	-0,292
L.lynx	Kirov	chrF1	12905	3,96E-04	5,07E-04	1,53E-02	-0,282
L.lynx	Kirov	chrF2	8764	3,40E-04	4,42E-04	1,58E-02	-0,301
L.lynx	Kirov	chrX	10898	1,37E-04	1,55E-04	5,38E-03	-0,128
L.lynx	NE_Poland	chrA1	25888	2,55E-04	3,29E-04	1,51E-02	-0,287
L.lynx	NE_Poland	chrA2	30793	2,72E-04	3,57E-04	1,63E-02	-0,314
L.lynx	NE_Poland	chrA3	22859	2,74E-04	3,39E-04	1,12E-02	-0,234
L.lynx	NE_Poland	chrB1	20444	2,58E-04	3,26E-04	1,28E-02	-0,263
L.lynx	NE_Poland	chrB2	19676	2,49E-04	3,11E-04	1,21E-02	-0,249
L.lynx	NE_Poland	chrB3	25096	2,48E-04	3,13E-04	1,34E-02	-0,263
L.lynx	NE_Poland	chrB4	24526	2,56E-04	3,26E-04	1,45E-02	-0,275
L.lynx	NE_Poland	chrC1	37667	2,43E-04	3,17E-04	1,52E-02	-0,304
L.lynx	NE_Poland	chrC2	18312	2,66E-04	3,35E-04	1,23E-02	-0,260
L.lynx	NE_Poland	chrD1	21637	2,70E-04	3,38E-04	1,36E-02	-0,253
L.lynx	NE_Poland	chrD2	13624	3,27E-04	4,29E-04	1,79E-02	-0,312
L.lynx	NE_Poland	chrD3	14538	2,95E-04	3,75E-04	1,52E-02	-0,273
L.lynx	NE_Poland	chrD4	15194	2,86E-04	3,57E-04	1,34E-02	-0,245
L.lynx	NE_Poland	chrE1	20072	3,07E-04	3,96E-04	1,71E-02	-0,290
L.lynx	NE_Poland	chrE2	15971	2,98E-04	3,77E-04	1,38E-02	-0,268
L.lynx	NE_Poland	chrF1	12905	3,15E-04	4,06E-04	1,53E-02	-0,288
L.lynx	NE_Poland	chrF2	8764	2,73E-04	3,47E-04	1,40E-02	-0,270
L.lynx	NE_Poland	chrX	10898	9,72E-05	1,13E-04	9,70E-03	-0,160
L.lynx	Norway	chrA1	25888	2,47E-04	3,03E-04	7,80E-03	-0,228
L.lynx	Norway	chrA2	30793	2,52E-04	3,26E-04	8,52E-03	-0,295
L.lynx	Norway	chrA3	22859	2,54E-04	3,29E-04	7,97E-03	-0,298
L.lynx	Norway	chrB1	20444	2,48E-04	3,22E-04	8,36E-03	-0,297
L.lynx	Norway	chrB2	19676	2,52E-04	3,15E-04	8,27E-03	-0,248
L.lynx	Norway	chrB3	25096	2,34E-04	2,88E-04	4,05E-03	-0,233
L.lynx	Norway	chrB4	24526	2,25E-04	2,74E-04	4,50E-03	-0,214
L.lynx	Norway	chrC1	37667	2,32E-04	2,83E-04	5,92E-03	-0,219
L.lynx	Norway	chrC2	18312	2,30E-04	2,87E-04	5,87E-03	-0,249
L.lynx	Norway	chrD1	21637	2,56E-04	3,34E-04	6,13E-03	-0,305
L.lynx	Norway	chrD2	13624	2,66E-04	3,19E-04	4,38E-03	-0,203
L.lynx	Norway	chrD3	14538	2,66E-04	3,45E-04	9,63E-03	-0,297
L.lynx	Norway	chrD4	15194	2,48E-04	3,30E-04	7,55E-03	-0,331

Mean recombination rate	Weighted mean divergence	Mean GC content	Mean RVIS	Mean percentage functional sites
3,737	5,66E-03	0,444	-0,288	4,245
2,149	5,79E-03	0,509	-0,199	7,206
2,824	5,80E-03	0,482	-0,380	6,316
2,370	5,40E-03	0,432	-0,210	4,281
2,741	5,25E-03	0,458	-0,263	4,955
3,251	5,23E-03	0,468	-0,478	6,287
2,906	5,22E-03	0,469	-0,325	6,596
3,132	5,14E-03	0,471	-0,222	6,537
2,531	5,26E-03	0,447	-0,267	6,973
4,489	5,36E-03	0,505	-0,201	5,128
3,685	5,69E-03	0,471	-0,297	6,127
2,922	5,77E-03	0,485	-0,239	5,769
3,903	5,60E-03	0,497	-0,287	4,613
4,362	5,51E-03	0,531	-0,257	8,150
4,240	6,03E-03	0,537	-0,198	9,271
4,680	5,53E-03	0,487	-0,311	5,999
2,767	5,45E-03	0,455	-0,285	3,448
2,394	4,89E-03	0,440	-0,417	2,025
3,737	5,66E-03	0,444	-0,288	4,245
2,149	5,79E-03	0,509	-0,199	7,206
2,824	5,80E-03	0,482	-0,380	6,316
2,370	5,40E-03	0,432	-0,210	4,281
2,741	5,25E-03	0,458	-0,263	4,955
3,251	5,23E-03	0,468	-0,478	6,287
2,906	5,22E-03	0,469	-0,325	6,596
3,132	5,14E-03	0,471	-0,222	6,537
2,531	5,26E-03	0,447	-0,267	6,973
4,489	5,36E-03	0,505	-0,201	5,128
3,685	5,69E-03	0,471	-0,297	6,127
2,922	5,77E-03	0,485	-0,239	5,769
3,903	5,60E-03	0,497	-0,287	4,613
4,362	5,51E-03	0,531	-0,257	8,150
4,240	6,03E-03	0,537	-0,198	9,271
4,680	5,53E-03	0,487	-0,311	5,999
2,767	5,45E-03	0,455	-0,285	3,448
2,394	4,89E-03	0,440	-0,417	2,025
3,737	5,66E-03	0,444	-0,288	4,245
2,149	5,79E-03	0,509	-0,199	7,206
2,824	5,80E-03	0,482	-0,380	6,316
2,370	5,40E-03	0,432	-0,210	4,281
2,741	5,25E-03	0,458	-0,263	4,955
3,251	5,23E-03	0,468	-0,478	6,287
2,906	5,22E-03	0,469	-0,325	6,596
3,132	5,14E-03	0,471	-0,222	6,537
2,531	5,26E-03	0,447	-0,267	6,973
4,489	5,36E-03	0,505	-0,201	5,128
3,685	5,69E-03	0,471	-0,297	6,127
2,922	5,77E-03	0,485	-0,239	5,769
3,903	5,60E-03	0,497	-0,287	4,613

L.lynx	Norway	chrE1	20072	2,71E-04	3,60E-04	7,75E-03	-0,328
L.lynx	Norway	chrE2	15971	2,67E-04	3,76E-04	1,12E-02	-0,407
L.lynx	Norway	chrF1	12905	2,92E-04	4,02E-04	1,09E-02	-0,374
L.lynx	Norway	chrF2	8764	2,68E-04	3,53E-04	8,12E-03	-0,317
L.lynx	Norway	chrX	10898	2,10E-04	1,17E-04	-5,40E-03	0,445
L.pardinus	Andujar	chrA1	25901	1,53E-04	2,15E-04	1,20E-02	-0,404
L.pardinus	Andujar	chrA2	30560	1,57E-04	2,32E-04	1,33E-02	-0,478
L.pardinus	Andujar	chrA3	22808	1,60E-04	2,37E-04	1,53E-02	-0,484
L.pardinus	Andujar	chrB1	20435	1,80E-04	2,50E-04	1,22E-02	-0,390
L.pardinus	Andujar	chrB2	19651	1,73E-04	2,42E-04	1,32E-02	-0,399
L.pardinus	Andujar	chrB3	25079	1,64E-04	2,44E-04	1,58E-02	-0,489
L.pardinus	Andujar	chrB4	24456	1,64E-04	2,36E-04	1,39E-02	-0,440
L.pardinus	Andujar	chrC1	37618	1,55E-04	2,18E-04	1,25E-02	-0,412
L.pardinus	Andujar	chrC2	18292	1,72E-04	2,51E-04	1,45E-02	-0,457
L.pardinus	Andujar	chrD1	21525	2,11E-04	3,03E-04	1,60E-02	-0,433
L.pardinus	Andujar	chrD2	13615	2,06E-04	2,96E-04	1,59E-02	-0,435
L.pardinus	Andujar	chrD3	14547	1,93E-04	2,83E-04	1,50E-02	-0,466
L.pardinus	Andujar	chrD4	15121	1,84E-04	2,63E-04	1,43E-02	-0,426
L.pardinus	Andujar	chrE1	19928	2,22E-04	3,16E-04	1,56E-02	-0,425
L.pardinus	Andujar	chrE2	15829	2,15E-04	3,18E-04	1,74E-02	-0,475
L.pardinus	Andujar	chrF1	12849	2,27E-04	3,17E-04	1,42E-02	-0,392
L.pardinus	Andujar	chrF2	8722	1,88E-04	2,55E-04	1,15E-02	-0,359
L.pardinus	Andujar	chrX	10948	6,86E-05	8,92E-05	8,71E-03	-0,301
L.pardinus	Doñana	chrA1	25901	9,30E-05	1,29E-04	5,83E-03	-0,391
L.pardinus	Doñana	chrA2	30560	9,55E-05	1,33E-04	6,49E-03	-0,395
L.pardinus	Doñana	chrA3	22808	1,02E-04	1,39E-04	6,47E-03	-0,371
L.pardinus	Doñana	chrB1	20435	9,13E-05	1,41E-04	6,64E-03	-0,545
L.pardinus	Doñana	chrB2	19651	1,08E-04	1,38E-04	5,09E-03	-0,276
L.pardinus	Doñana	chrB3	25079	9,42E-05	1,33E-04	6,31E-03	-0,412
L.pardinus	Doñana	chrB4	24456	8,54E-05	1,17E-04	4,76E-03	-0,369
L.pardinus	Doñana	chrC1	37618	8,44E-05	1,06E-04	2,40E-03	-0,255
L.pardinus	Doñana	chrC2	18292	8,74E-05	1,30E-04	6,00E-03	-0,485
L.pardinus	Doñana	chrD1	21525	8,25E-05	1,21E-04	3,26E-03	-0,464
L.pardinus	Doñana	chrD2	13615	8,82E-05	1,23E-04	5,59E-03	-0,397
L.pardinus	Doñana	chrD3	14547	9,89E-05	1,42E-04	7,61E-03	-0,439
L.pardinus	Doñana	chrD4	15121	7,06E-05	7,77E-05	9,60E-04	-0,101
L.pardinus	Doñana	chrE1	19928	1,10E-04	1,74E-04	9,91E-03	-0,578
L.pardinus	Doñana	chrE2	15829	9,80E-05	1,38E-04	4,95E-03	-0,404
L.pardinus	Doñana	chrF1	12849	1,17E-04	1,53E-04	3,88E-03	-0,306
L.pardinus	Doñana	chrF2	8722	1,05E-04	1,59E-04	6,67E-03	-0,511
L.pardinus	Doñana	chrX	10948	7,45E-05	4,05E-05	-5,39E-03	0,457

4,362	5,51E-03	0,531	-0,257	8,150
4,240	6,03E-03	0,537	-0,198	9,271
4,680	5,53E-03	0,487	-0,311	5,999
2,767	5,45E-03	0,455	-0,285	3,448
2,394	4,89E-03	0,440	-0,417	2,025
3,739	5,80E-03	0,444	-0,288	4,245
2,151	5,83E-03	0,508	-0,202	7,203
2,821	5,84E-03	0,481	-0,390	6,317
2,372	5,44E-03	0,431	-0,212	4,284
2,738	5,32E-03	0,458	-0,264	4,956
3,253	5,29E-03	0,468	-0,480	6,291
2,900	5,32E-03	0,469	-0,325	6,594
3,127	5,20E-03	0,470	-0,225	6,530
2,527	5,31E-03	0,446	-0,268	6,976
4,479	5,39E-03	0,504	-0,201	5,127
3,685	5,76E-03	0,471	-0,297	6,129
2,914	5,77E-03	0,484	-0,236	5,771
3,901	5,65E-03	0,497	-0,292	4,613
4,361	5,55E-03	0,530	-0,259	8,144
4,239	6,09E-03	0,536	-0,197	9,275
4,681	5,56E-03	0,486	-0,312	6,000
2,779	5,49E-03	0,454	-0,286	3,452
2,386	4,95E-03	0,440	-0,417	2,025
3,739	5,80E-03	0,444	-0,288	4,245
2,151	5,83E-03	0,508	-0,202	7,203
2,821	5,84E-03	0,481	-0,390	6,317
2,372	5,44E-03	0,431	-0,212	4,284
2,738	5,32E-03	0,458	-0,264	4,956
3,253	5,29E-03	0,468	-0,480	6,291
2,900	5,32E-03	0,469	-0,325	6,594
3,127	5,20E-03	0,470	-0,225	6,530
2,527	5,31E-03	0,446	-0,268	6,976
4,479	5,39E-03	0,504	-0,201	5,127
3,685	5,76E-03	0,471	-0,297	6,129
2,914	5,77E-03	0,484	-0,236	5,771
3,901	5,65E-03	0,497	-0,292	4,613
4,361	5,55E-03	0,530	-0,259	8,144
4,239	6,09E-03	0,536	-0,197	9,275
4,681	5,56E-03	0,486	-0,312	6,000
2,779	5,49E-03	0,454	-0,286	3,452
2,386	4,95E-03	0,440	-0,417	2,025

Table S5. Weighted genomic diversity (watterson estimator and nucleotide divesity), Tajima's D, mean recombination rate, divergence, GC content, RVIS, and functional sites percentage for different chromosome regions.

Species	Populations	Region	Number of units	Weighted mean Watterson estimator	Stdv boot Watterson estimator	Weighted mean nucleotide diversity	Stdv boot nucleotide diversity	Weighted Tajima's D
L.lynx	Kirov	Subtelomeric	16992	5,10E-04	3,43E-06	6,56E-04	4,37E-06	1,81E-02
L.lynx	Kirov	Interstitial	279984	3,36E-04	6,30E-07	4,32E-04	1,02E-06	1,41E-02
L.lynx	Kirov	Pericentromeric	50990	2,93E-04	1,59E-06	3,76E-04	2,42E-06	1,26E-02
L.lynx	NE_Poland	Subtelomeric	16992	4,31E-04	3,04E-06	5,60E-04	4,34E-06	2,00E-02
L.lynx	NE_Poland	Interstitial	279984	2,66E-04	5,80E-07	3,39E-04	7,83E-07	1,40E-02
L.lynx	NE_Poland	Pericentromeric	50990	2,48E-04	1,40E-06	3,18E-04	2,27E-06	1,39E-02
L.lynx	Norway	Subtelomeric	16992	3,20E-04	3,28E-06	5,21E-04	4,37E-06	1,39E-02
L.lynx	Norway	Interstitial	279984	2,49E-04	7,26E-07	3,14E-04	8,46E-07	7,09E-03
L.lynx	Norway	Pericentromeric	50990	2,35E-04	1,89E-06	2,96E-04	1,95E-06	7,02E-03
L.pardinus	Andujar	Subtelomeric	16651	2,75E-04	2,29E-06	4,04E-04	3,79E-06	2,06E-02
L.pardinus	Andujar	Interstitial	279327	1,76E-04	4,49E-07	2,51E-04	6,80E-07	1,37E-02
L.pardinus	Andujar	Pericentromeric	50958	1,53E-04	1,23E-06	2,22E-04	1,78E-06	1,37E-02
L.pardinus	Doñana	Subtelomeric	16651	1,27E-04	1,99E-06	2,07E-04	2,99E-06	8,64E-03
L.pardinus	Doñana	Interstitial	279327	9,37E-05	4,26E-07	1,31E-04	5,28E-07	5,55E-03
L.pardinus	Doñana	Pericentromeric	50958	8,48E-05	1,06E-06	1,09E-04	1,47E-06	3,85E-03

SFS Skewness	Mean recombination rate (cM/Mb)	Weighted mean divergence	Mean GC content	Mean RVIS	Mean functional sites percentage
-0,287	3,705	7,90E-03	0,584	0,290	3,301
-0,286	3,490	5,45E-03	0,471	-0,309	6,035
-0,281	2,261	5,19E-03	0,485	-0,284	6,331
-0,301	3,705	7,90E-03	0,584	0,290	3,301
-0,273	3,490	5,45E-03	0,471	-0,309	6,035
-0,279	2,261	5,19E-03	0,485	-0,284	6,331
-0,629	3,705	7,90E-03	0,584	0,290	3,301
-0,261	3,490	5,45E-03	0,471	-0,309	6,035
-0,259	2,261	5,19E-03	0,485	-0,284	6,331
-0,470	3,705	7,88E-03	0,582	0,279	3,311
-0,427	3,488	5,52E-03	0,471	-0,311	6,033
-0,456	2,255	5,29E-03	0,485	-0,283	6,330
-0,628	3,705	7,88E-03	0,582	0,279	3,311
-0,397	3,488	5,52E-03	0,471	-0,311	6,033
-0,286	2,255	5,29E-03	0,485	-0,283	6,330

Table S6. Weighted genomic diversity (watterson estimator and nucleotide divesity), Tajima's D and *S* for different features in the autosomes.

Species	Populations	Feature	Number of units	Wmean Watterson estimator	Stdv boot Watterson estimator	Wmean nt diversity	Stdv boot nucleotide diversity	WTajima's D	S
L.l.	Kirov	Intergenic	26674	3,52E-04	1,56E-06	4,54E-04	1,83E-06	1,49E-02	-0,288
L.l.	Kirov	Gen promoter	6174	3,05E-04	2,98E-06	3,90E-04	4,89E-06	1,28E-02	-0,282
L.l.	Kirov	5UTR	4012	2,93E-04	4,89E-06	3,75E-04	8,54E-06	1,23E-02	-0,279
L.l.	Kirov	CDS	146653	2,10E-04	1,00E-06	2,62E-04	1,40E-06	1,08E-02	-0,249
L.l.	Kirov	Intron	147927	3,20E-04	9,03E-07	4,11E-04	1,30E-06	1,29E-02	-0,284
L.l.	Kirov	3UTR	5057	2,63E-04	4,14E-06	3,30E-04	4,38E-06	9,67E-03	-0,256
L.l.	Kirov	lncRNA promoter	2550	3,30E-04	6,03E-06	4,27E-04	8,77E-06	1,51E-02	-0,294
L.l.	Kirov	lncRNA exons	1726	3,42E-04	9,72E-06	4,29E-04	1,54E-05	1,56E-02	-0,252
L.l.	Kirov	lncRNA intron	1473	3,43E-04	1,18E-05	4,28E-04	1,55E-05	1,12E-02	-0,250
L.l.	Kirov	ncRNA	1839	3,02E-04	1,13E-05	3,79E-04	1,59E-05	2,05E-02	-0,254
L.l.	Kirov	UCNE	3881	7,99E-05	3,38E-06	8,91E-05	3,37E-06	2,24E-03	-0,115
L.l.	NE_Poland	Intergenic	26674	2,83E-04	1,35E-06	3,61E-04	1,71E-06	1,45E-02	-0,276
L.l.	NE_Poland	Gen promoter	6174	2,43E-04	3,39E-06	3,09E-04	4,18E-06	1,66E-02	-0,272
L.l.	NE_Poland	5UTR	4012	2,32E-04	5,46E-06	2,98E-04	6,77E-06	1,35E-02	-0,284
L.l.	NE_Poland	CDS	146653	1,66E-04	1,08E-06	2,10E-04	1,20E-06	1,27E-02	-0,264
L.l.	NE_Poland	Intron	147927	2,56E-04	8,07E-07	3,25E-04	1,08E-06	1,35E-02	-0,272
L.l.	NE_Poland	3UTR	5057	2,06E-04	3,50E-06	2,59E-04	4,41E-06	1,24E-02	-0,261
L.l.	NE_Poland	lncRNA promoter	2550	2,62E-04	5,43E-06	3,35E-04	7,54E-06	1,81E-02	-0,278
L.l.	NE_Poland	lncRNA exons	1726	2,75E-04	1,02E-05	3,56E-04	1,53E-05	1,62E-02	-0,292
L.l.	NE_Poland	lncRNA intron	1473	2,75E-04	1,19E-05	3,60E-04	1,45E-05	1,58E-02	-0,309
L.l.	NE_Poland	ncRNA	1839	2,34E-04	1,16E-05	2,97E-04	2,07E-05	1,12E-02	-0,270
L.l.	NE_Poland	UCNE	3881	5,99E-05	2,31E-06	7,24E-05	3,62E-06	5,51E-04	-0,208
L.l.	Norway	Intergenic	26674	2,59E-04	1,43E-06	3,36E-04	2,06E-06	8,15E-03	-0,299
L.l.	Norway	Gen promoter	6174	2,27E-04	2,84E-06	2,87E-04	4,36E-06	4,49E-03	-0,261
L.l.	Norway	5UTR	4012	2,23E-04	5,65E-06	2,76E-04	7,20E-06	5,78E-03	-0,237
L.l.	Norway	CDS	146653	2,04E-04	1,19E-06	1,95E-04	1,31E-06	-6,38E-03	0,045
L.l.	Norway	Intron	147927	2,39E-04	7,47E-07	2,99E-04	1,04E-06	6,07E-03	-0,251
L.l.	Norway	3UTR	5057	2,08E-04	3,72E-06	2,46E-04	5,08E-06	4,34E-03	-0,180
L.l.	Norway	lncRNA promoter	2550	2,37E-04	5,09E-06	3,13E-04	8,29E-06	6,08E-03	-0,318
L.l.	Norway	lncRNA exons	1726	2,40E-04	9,34E-06	3,33E-04	1,16E-05	8,80E-03	-0,392
L.l.	Norway	lncRNA intron	1473	2,54E-04	9,13E-06	3,17E-04	1,45E-05	7,26E-03	-0,250
L.l.	Norway	ncRNA	1839	2,08E-04	9,87E-06	2,57E-04	1,44E-05	-3,20E-03	-0,231
L.l.	Norway	UCNE	3881	1,45E-04	4,18E-06	6,73E-05	3,05E-06	-2,78E-02	0,535

L.p.	Andujar	Intergenic	26755	1,86E-04	9,74E-07	2,66E-04	1,63E-06	1,41E-02	-0,430
L.p.	Andujar	Gen promoter	6158	1,59E-04	2,86E-06	2,31E-04	3,65E-06	1,78E-02	-0,453
L.p.	Andujar	5UTR	3984	1,54E-04	3,92E-06	2,20E-04	5,91E-06	1,26E-02	-0,429
L.p.	Andujar	CDS	145973	1,13E-04	7,68E-07	1,60E-04	1,32E-06	1,47E-02	-0,418
L.p.	Andujar	Intron	147559	1,66E-04	6,38E-07	2,38E-04	9,75E-07	1,34E-02	-0,435
L.p.	Andujar	3UTR	5051	1,32E-04	2,55E-06	1,85E-04	3,83E-06	1,15E-02	-0,397
L.p.	Andujar	lncRNA promoter	2538	1,78E-04	4,67E-06	2,56E-04	7,51E-06	1,77E-02	-0,437
L.p.	Andujar	lncRNA exons	1720	1,69E-04	8,19E-06	2,43E-04	1,21E-05	1,67E-02	-0,440
L.p.	Andujar	lncRNA intron	1460	1,86E-04	7,02E-06	2,65E-04	9,19E-06	1,47E-02	-0,420
L.p.	Andujar	ncRNA	1857	1,84E-04	9,17E-06	2,71E-04	1,25E-05	4,04E-02	-0,475
L.p.	Andujar	UCNE	3881	3,94E-05	2,09E-06	5,18E-05	3,03E-06	1,39E-02	-0,313
L.p.	Doñana	Intergenic	26755	9,73E-05	9,48E-07	1,39E-04	1,54E-06	5,84E-03	-0,431
L.p.	Doñana	Gen promoter	6158	8,50E-05	2,12E-06	1,21E-04	3,16E-06	6,54E-03	-0,422
L.p.	Doñana	5UTR	3984	8,28E-05	3,31E-06	1,11E-04	4,46E-06	4,67E-03	-0,344
L.p.	Doñana	CDS	145973	7,61E-05	6,85E-07	8,27E-05	8,53E-07	-9,82E-04	-0,087
L.p.	Doñana	Intron	147559	8,93E-05	4,42E-07	1,20E-04	6,91E-07	4,71E-03	-0,349
L.p.	Doñana	3UTR	5051	8,11E-05	1,98E-06	9,67E-05	2,40E-06	2,32E-03	-0,192
L.p.	Doñana	lncRNA promoter	2538	8,46E-05	3,62E-06	1,39E-04	6,65E-06	1,21E-02	-0,644
L.p.	Doñana	lncRNA exons	1720	7,20E-05	4,78E-06	1,06E-04	9,88E-06	6,05E-03	-0,477
L.p.	Doñana	lncRNA intron	1460	9,84E-05	4,34E-06	1,48E-04	6,88E-06	6,94E-03	-0,505
L.p.	Doñana	ncRNA	1857	1,06E-04	7,66E-06	1,12E-04	1,09E-05	-6,63E-03	-0,050
L.p.	Doñana	UCNE	3881	5,35E-05	2,44E-06	2,67E-05	1,70E-06	-2,31E-02	0,502

Table S7. Wilcoxon signed-rank test for paired samples comparing different features to intergenic for Delta based on watterson estimator and nucleotide diversity.

Comparison	Method	Feature1	Feature2	p-value	Diversity measure
Kirov-Poland	wilcox test	3UTR	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Poland	wilcox test	5UTR	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Poland	wilcox test	CDS	Intergenic	1,85E-09	Nucleotide diversity
Kirov-Poland	wilcox test	Gen promoter	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Poland	wilcox test	Intron	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Poland	wilcox test	lncRNA exons	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Poland	wilcox test	lncRNA intron	Intergenic	1,99E-01	Nucleotide diversity
Kirov-Poland	wilcox test	lncRNA promoter	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Poland	wilcox test	ncRNA	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Poland	wilcox test	UCNE	Intergenic	1,44E-05	Nucleotide diversity
Kirov-Poland	wilcox test	3UTR	Intergenic	1,00E+00	Watterson estimator
Kirov-Poland	wilcox test	5UTR	Intergenic	1,00E+00	Watterson estimator
Kirov-Poland	wilcox test	CDS	Intergenic	3,06E-04	Watterson estimator
Kirov-Poland	wilcox test	Gen promoter	Intergenic	5,80E-01	Watterson estimator
Kirov-Poland	wilcox test	Intron	Intergenic	2,23E-10	Watterson estimator
Kirov-Poland	wilcox test	lncRNA exons	Intergenic	1,00E+00	Watterson estimator
Kirov-Poland	wilcox test	lncRNA intron	Intergenic	1,00E+00	Watterson estimator
Kirov-Poland	wilcox test	lncRNA promoter	Intergenic	3,62E-03	Watterson estimator
Kirov-Poland	wilcox test	ncRNA	Intergenic	1,00E+00	Watterson estimator
Kirov-Poland	wilcox test	UCNE	Intergenic	8,56E-13	Watterson estimator
Kirov-Norway	wilcox test	3UTR	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Norway	wilcox test	5UTR	Intergenic	3,51E-02	Nucleotide diversity
Kirov-Norway	wilcox test	CDS	Intergenic	1,95E-18	Nucleotide diversity
Kirov-Norway	wilcox test	Gen promoter	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Norway	wilcox test	Intron	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Norway	wilcox test	lncRNA exons	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Norway	wilcox test	lncRNA intron	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Norway	wilcox test	lncRNA promoter	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Norway	wilcox test	ncRNA	Intergenic	1,00E+00	Nucleotide diversity
Kirov-Norway	wilcox test	UCNE	Intergenic	3,05E-09	Nucleotide diversity
Kirov-Norway	wilcox test	3UTR	Intergenic	3,25E-06	Watterson estimator
Kirov-Norway	wilcox test	5UTR	Intergenic	2,69E-04	Watterson estimator
Kirov-Norway	wilcox test	CDS	Intergenic	3,63E-124	Watterson estimator
Kirov-Norway	wilcox test	Gen promoter	Intergenic	1,00E+00	Watterson estimator
Kirov-Norway	wilcox test	Intron	Intergenic	5,73E-07	Watterson estimator
Kirov-Norway	wilcox test	lncRNA exons	Intergenic	1,00E+00	Watterson estimator
Kirov-Norway	wilcox test	lncRNA intron	Intergenic	1,00E+00	Watterson estimator
Kirov-Norway	wilcox test	lncRNA promoter	Intergenic	1,00E+00	Watterson estimator
Kirov-Norway	wilcox test	ncRNA	Intergenic	1,00E+00	Watterson estimator
Kirov-Norway	wilcox test	UCNE	Intergenic	5,26E-42	Watterson estimator

Andujar-Doñana	wilcox test	3UTR	Intergenic	2,21E-05	Nucleotide diversity
Andujar-Doñana	wilcox test	5UTR	Intergenic	1,40E-03	Nucleotide diversity
Andujar-Doñana	wilcox test	CDS	Intergenic	3,11E-64	Nucleotide diversity
Andujar-Doñana	wilcox test	Gen promoter	Intergenic	1,26E-02	Nucleotide diversity
Andujar-Doñana	wilcox test	Intron	Intergenic	1,21E-14	Nucleotide diversity
Andujar-Doñana	wilcox test	lncRNA exons	Intergenic	1,00E+00	Nucleotide diversity
Andujar-Doñana	wilcox test	lncRNA intron	Intergenic	1,00E+00	Nucleotide diversity
Andujar-Doñana	wilcox test	lncRNA promoter	Intergenic	1,44E-01	Nucleotide diversity
Andujar-Doñana	wilcox test	ncRNA	Intergenic	1,00E+00	Nucleotide diversity
Andujar-Doñana	wilcox test	UCNE	Intergenic	4,87E-14	Nucleotide diversity
Andujar-Doñana	wilcox test	3UTR	Intergenic	1,19E-21	Watterson estimator
Andujar-Doñana	wilcox test	5UTR	Intergenic	1,15E-08	Watterson estimator
Andujar-Doñana	wilcox test	CDS	Intergenic	3,70E-177	Watterson estimator
Andujar-Doñana	wilcox test	Gen promoter	Intergenic	1,34E-01	Watterson estimator
Andujar-Doñana	wilcox test	Intron	Intergenic	1,64E-72	Watterson estimator
Andujar-Doñana	wilcox test	lncRNA exons	Intergenic	1,00E+00	Watterson estimator
Andujar-Doñana	wilcox test	lncRNA intron	Intergenic	1,00E+00	Watterson estimator
Andujar-Doñana	wilcox test	lncRNA promoter	Intergenic	1,00E+00	Watterson estimator
Andujar-Doñana	wilcox test	ncRNA	Intergenic	1,00E+00	Watterson estimator
Andujar-Doñana	wilcox test	UCNE	Intergenic	1,09E-44	Watterson estimator

Table S8. Wilcoxon signed-rank test comparing $\delta_S\theta_W > 0.1$ windows vs. $\delta_S\theta_W < 0.1$ windows for NB θ_W, NB π, NB S, divergence, recombination, GC content, functional site percentage, and RVIS. Features considered for the analysis are intergenic, introns and CDS.

Comparison	Method	Feature	Genomic variable	Alternative hypothesis	Interpretation	Estimate	Statistic	p-value
Kirov-Poland	wilcox.test	CDS	divergence	greater	[d<0,1] < [d>0,1]	4,6E-04	2,4E+07	3,0E-25
Kirov-Poland	wilcox.test	CDS	divergence	less	[d<0,1] > [d>0,1]	4,6E-04	2,4E+07	1,0E+00
Kirov-Poland	wilcox.test	CDS	functional sites percentage	greater	[d<0,1] < [d>0,1]	-5,4E-02	3,5E+06	6,9E-01
Kirov-Poland	wilcox.test	CDS	functional sites percentage	less	[d<0,1] > [d>0,1]	-5,4E-02	3,5E+06	3,1E-01
Kirov-Poland	wilcox.test	CDS	GC content	greater	[d<0,1] < [d>0,1]	2,1E-02	2,4E+07	1,3E-22
Kirov-Poland	wilcox.test	CDS	GC content	less	[d<0,1] > [d>0,1]	2,1E-02	2,4E+07	1,0E+00
Kirov-Poland	wilcox.test	CDS	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-3,3E-06	2,3E+07	4,5E-07
Kirov-Poland	wilcox.test	CDS	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-3,3E-06	2,3E+07	1,0E+00
Kirov-Poland	wilcox.test	CDS	recombination	greater	[d<0,1] < [d>0,1]	2,5E-01	2,0E+07	1,3E-07
Kirov-Poland	wilcox.test	CDS	recombination	less	[d<0,1] > [d>0,1]	2,5E-01	2,0E+07	1,0E+00
Kirov-Poland	wilcox.test	CDS	RVIS	greater	[d<0,1] < [d>0,1]	1,8E-01	1,5E+07	2,9E-14
Kirov-Poland	wilcox.test	CDS	RVIS	less	[d<0,1] > [d>0,1]	1,8E-01	1,5E+07	1,0E+00
Kirov-Poland	wilcox.test	CDS	Skewness	greater	[d<0,1] < [d>0,1]	-4,8E-03	2,0E+07	1,0E+00
Kirov-Poland	wilcox.test	CDS	Skewness	less	[d<0,1] > [d>0,1]	-4,8E-03	2,0E+07	1,3E-04
Kirov-Poland	wilcox.test	CDS	Watterson estimator	greater	[d<0,1] < [d>0,1]	-2,8E-05	2,3E+07	8,9E-10
Kirov-Poland	wilcox.test	CDS	Watterson estimator	less	[d<0,1] > [d>0,1]	-2,8E-05	2,3E+07	1,0E+00
Kirov-Poland	wilcox.test	Intergenic	divergence	greater	[d<0,1] < [d>0,1]	2,1E-04	1,7E+10	0,0E+00
Kirov-Poland	wilcox.test	Intergenic	divergence	less	[d<0,1] > [d>0,1]	2,1E-04	1,7E+10	1,0E+00
Kirov-Poland	wilcox.test	Intergenic	GC content	greater	[d<0,1] < [d>0,1]	9,0E-03	1,7E+10	0,0E+00
Kirov-Poland	wilcox.test	Intergenic	GC content	less	[d<0,1] > [d>0,1]	9,0E-03	1,7E+10	1,0E+00
Kirov-Poland	wilcox.test	Intergenic	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-4,1E-05	1,4E+10	1,0E+00
Kirov-Poland	wilcox.test	Intergenic	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-4,1E-05	1,4E+10	9,2E-197
Kirov-Poland	wilcox.test	Intergenic	recombination	greater	[d<0,1] < [d>0,1]	1,7E-01	1,3E+10	1,5E-99

Kirov-Poland	wilcox.test	Intergenic	recombination	less	[d<0,1] > [d>0,1]	1,7E-01	1,3E+10	1,0E+00
Kirov-Poland	wilcox.test	Intergenic	Skewness	greater	[d<0,1] < [d>0,1]	8,4E-03	1,6E+10	1,6E-232
Kirov-Poland	wilcox.test	Intergenic	Skewness	less	[d<0,1] > [d>0,1]	8,4E-03	1,6E+10	1,0E+00
Kirov-Poland	wilcox.test	Intergenic	Watterson estimator	greater	[d<0,1] < [d>0,1]	-3,3E-05	1,4E+10	1,0E+00
Kirov-Poland	wilcox.test	Intergenic	Watterson estimator	less	[d<0,1] > [d>0,1]	-3,3E-05	1,4E+10	7,6E-71
Kirov-Poland	wilcox.test	Intron	divergence	greater	[d<0,1] < [d>0,1]	3,2E-04	6,1E+09	1,1E-231
Kirov-Poland	wilcox.test	Intron	divergence	less	[d<0,1] > [d>0,1]	3,2E-04	6,1E+09	1,0E+00
Kirov-Poland	wilcox.test	Intron	functional sites percentage	greater	[d<0,1] < [d>0,1]	-1,3E-01	1,2E+09	1,0E+00
Kirov-Poland	wilcox.test	Intron	functional sites percentage	less	[d<0,1] > [d>0,1]	-1,3E-01	1,2E+09	2,3E-09
Kirov-Poland	wilcox.test	Intron	GC content	greater	[d<0,1] < [d>0,1]	1,5E-02	6,2E+09	0,0E+00
Kirov-Poland	wilcox.test	Intron	GC content	less	[d<0,1] > [d>0,1]	1,5E-02	6,2E+09	1,0E+00
Kirov-Poland	wilcox.test	Intron	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-3,4E-05	5,2E+09	1,0E+00
Kirov-Poland	wilcox.test	Intron	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-3,4E-05	5,2E+09	9,9E-55
Kirov-Poland	wilcox.test	Intron	recombination	greater	[d<0,1] < [d>0,1]	3,0E-01	5,2E+09	3,0E-149
Kirov-Poland	wilcox.test	Intron	recombination	less	[d<0,1] > [d>0,1]	3,0E-01	5,2E+09	1,0E+00
Kirov-Poland	wilcox.test	Intron	RVIS	greater	[d<0,1] < [d>0,1]	4,7E-02	2,9E+09	1,0E-14
Kirov-Poland	wilcox.test	Intron	RVIS	less	[d<0,1] > [d>0,1]	4,7E-02	2,9E+09	1,0E+00
Kirov-Poland	wilcox.test	Intron	Skewness	greater	[d<0,1] < [d>0,1]	5,5E-03	5,8E+09	2,2E-64
Kirov-Poland	wilcox.test	Intron	Skewness	less	[d<0,1] > [d>0,1]	5,5E-03	5,8E+09	1,0E+00
Kirov-Poland	wilcox.test	Intron	Watterson estimator	greater	[d<0,1] < [d>0,1]	-6,4E-05	5,4E+09	1,0E+00
Kirov-Poland	wilcox.test	Intron	Watterson estimator	less	[d<0,1] > [d>0,1]	-6,4E-05	5,4E+09	1,9E-15
Kirov-Norway	wilcox.test	CDS	divergence	greater	[d<0,1] < [d>0,1]	-5,4E-05	4,6E+07	9,9E-01
Kirov-Norway	wilcox.test	CDS	divergence	less	[d<0,1] > [d>0,1]	-5,4E-05	4,6E+07	1,3E-02
Kirov-Norway	wilcox.test	CDS	functional sites percentage	greater	[d<0,1] < [d>0,1]	-2,2E-05	7,8E+06	5,7E-01
Kirov-Norway	wilcox.test	CDS	functional sites percentage	less	[d<0,1] > [d>0,1]	-2,2E-05	7,8E+06	4,3E-01

Kirov-Norway	wilcox.test	CDS	GC content	greater	[d<0,1] < [d>0,1]	-5,2E-04	4,7E+07	6,4E-01
Kirov-Norway	wilcox.test	CDS	GC content	less	[d<0,1] > [d>0,1]	-5,2E-04	4,7E+07	3,6E-01
Kirov-Norway	wilcox.test	CDS	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-1,1E-05	3,5E+07	1,0E+00
Kirov-Norway	wilcox.test	CDS	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-1,1E-05	3,5E+07	4,5E-160
Kirov-Norway	wilcox.test	CDS	recombination	greater	[d<0,1] < [d>0,1]	-1,1E-01	4,0E+07	1,0E+00
Kirov-Norway	wilcox.test	CDS	recombination	less	[d<0,1] > [d>0,1]	-1,1E-01	4,0E+07	8,3E-05
Kirov-Norway	wilcox.test	CDS	RVIS	greater	[d<0,1] < [d>0,1]	-4,3E-02	2,9E+07	1,0E+00
Kirov-Norway	wilcox.test	CDS	RVIS	less	[d<0,1] > [d>0,1]	-4,3E-02	2,9E+07	4,3E-03
Kirov-Norway	wilcox.test	CDS	Skewness	greater	[d<0,1] < [d>0,1]	1,9E-02	5,7E+07	1,2E-126
Kirov-Norway	wilcox.test	CDS	Skewness	less	[d<0,1] > [d>0,1]	1,9E-02	5,7E+07	1,0E+00
Kirov-Norway	wilcox.test	CDS	Watterson estimator	greater	[d<0,1] < [d>0,1]	-1,3E-05	3,5E+07	1,0E+00
Kirov-Norway	wilcox.test	CDS	Watterson estimator	less	[d<0,1] > [d>0,1]	-1,3E-05	3,5E+07	1,1E-160
Kirov-Norway	wilcox.test	Intergenic	divergence	greater	[d<0,1] < [d>0,1]	-6,9E-06	2,5E+10	9,8E-01
Kirov-Norway	wilcox.test	Intergenic	divergence	less	[d<0,1] > [d>0,1]	-6,9E-06	2,5E+10	2,4E-02
Kirov-Norway	wilcox.test	Intergenic	GC content	greater	[d<0,1] < [d>0,1]	-1,1E-03	2,5E+10	1,0E+00
Kirov-Norway	wilcox.test	Intergenic	GC content	less	[d<0,1] > [d>0,1]	-1,1E-03	2,5E+10	5,1E-13
Kirov-Norway	wilcox.test	Intergenic	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-1,7E-04	1,7E+10	1,0E+00
Kirov-Norway	wilcox.test	Intergenic	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-1,7E-04	1,7E+10	0,0E+00
Kirov-Norway	wilcox.test	Intergenic	recombination	greater	[d<0,1] < [d>0,1]	-8,1E-02	2,0E+10	1,0E+00
Kirov-Norway	wilcox.test	Intergenic	recombination	less	[d<0,1] > [d>0,1]	-8,1E-02	2,0E+10	1,3E-47
Kirov-Norway	wilcox.test	Intergenic	Skewness	greater	[d<0,1] < [d>0,1]	4,3E-02	3,2E+10	0,0E+00
Kirov-Norway	wilcox.test	Intergenic	Skewness	less	[d<0,1] > [d>0,1]	4,3E-02	3,2E+10	1,0E+00
Kirov-Norway	wilcox.test	Intergenic	Watterson estimator	greater	[d<0,1] < [d>0,1]	-2,0E-04	1,7E+10	1,0E+00
Kirov-Norway	wilcox.test	Intergenic	Watterson estimator	less	[d<0,1] > [d>0,1]	-2,0E-04	1,7E+10	0,0E+00
Kirov-Norway	wilcox.test	Intron	divergence	greater	[d<0,1] < [d>0,1]	-3,2E-06	9,4E+09	9,1E-01

Kirov-Norway	wilcox.test	Intron	divergence	less	[d<0,1] > [d>0,1]	-3,2E-06	9,4E+09	9,5E-02
Kirov-Norway	wilcox.test	Intron	functional sites percentage	greater	[d<0,1] < [d>0,1]	3,6E-02	2,1E+09	4,4E-03
Kirov-Norway	wilcox.test	Intron	functional sites percentage	less	[d<0,1] > [d>0,1]	3,6E-02	2,1E+09	1,0E+00
Kirov-Norway	wilcox.test	Intron	GC content	greater	[d<0,1] < [d>0,1]	2,2E-04	9,4E+09	2,2E-01
Kirov-Norway	wilcox.test	Intron	GC content	less	[d<0,1] > [d>0,1]	2,2E-04	9,4E+09	7,8E-01
Kirov-Norway	wilcox.test	Intron	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-1,1E-04	6,5E+09	1,0E+00
Kirov-Norway	wilcox.test	Intron	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-1,1E-04	6,5E+09	0,0E+00
Kirov-Norway	wilcox.test	Intron	recombination	greater	[d<0,1] < [d>0,1]	-9,1E-03	8,3E+09	9,5E-01
Kirov-Norway	wilcox.test	Intron	recombination	less	[d<0,1] > [d>0,1]	-9,1E-03	8,3E+09	5,4E-02
Kirov-Norway	wilcox.test	Intron	RVIS	greater	[d<0,1] < [d>0,1]	3,4E-02	5,0E+09	1,2E-13
Kirov-Norway	wilcox.test	Intron	RVIS	less	[d<0,1] > [d>0,1]	3,4E-02	5,0E+09	1,0E+00
Kirov-Norway	wilcox.test	Intron	Skewness	greater	[d<0,1] < [d>0,1]	3,5E-02	1,2E+10	0,0E+00
Kirov-Norway	wilcox.test	Intron	Skewness	less	[d<0,1] > [d>0,1]	3,5E-02	1,2E+10	1,0E+00
Kirov-Norway	wilcox.test	Intron	Watterson estimator	greater	[d<0,1] < [d>0,1]	-3,6E-05	6,5E+09	1,0E+00
Kirov-Norway	wilcox.test	Intron	Watterson estimator	less	[d<0,1] > [d>0,1]	-3,6E-05	6,5E+09	0,0E+00
Andujar-Doñana	wilcox.test	CDS	divergence	greater	[d<0,1] < [d>0,1]	-1,6E-05	3,0E+07	7,9E-01
Andujar-Doñana	wilcox.test	CDS	divergence	less	[d<0,1] > [d>0,1]	-1,6E-05	3,0E+07	2,1E-01
Andujar-Doñana	wilcox.test	CDS	functional sites percentage	greater	[d<0,1] < [d>0,1]	-1,5E-02	5,2E+06	6,2E-01
Andujar-Doñana	wilcox.test	CDS	functional sites percentage	less	[d<0,1] > [d>0,1]	-1,5E-02	5,2E+06	3,8E-01
Andujar-Doñana	wilcox.test	CDS	GC content	greater	[d<0,1] < [d>0,1]	6,6E-03	3,2E+07	1,0E-04
Andujar-Doñana	wilcox.test	CDS	GC content	less	[d<0,1] > [d>0,1]	6,6E-03	3,2E+07	1,0E+00
Andujar-Doñana	wilcox.test	CDS	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-1,1E-05	2,6E+07	1,0E+00
Andujar-Doñana	wilcox.test	CDS	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-1,1E-05	2,6E+07	3,9E-46
Andujar-Doñana	wilcox.test	CDS	recombination	greater	[d<0,1] < [d>0,1]	-1,0E-02	2,7E+07	6,6E-01
Andujar-Doñana	wilcox.test	CDS	recombination	less	[d<0,1] > [d>0,1]	-1,0E-02	2,7E+07	3,4E-01

Andujar-Doñana	wilcox.test	CDS	RVIS	greater	[d<0,1] < [d>0,1]	-6,2E-02	1,9E+07	1,0E+00
Andujar-Doñana	wilcox.test	CDS	RVIS	less	[d<0,1] > [d>0,1]	-6,2E-02	1,9E+07	9,7E-04
Andujar-Doñana	wilcox.test	CDS	Skewness	greater	[d<0,1] < [d>0,1]	2,7E-02	3,5E+07	9,9E-40
Andujar-Doñana	wilcox.test	CDS	Skewness	less	[d<0,1] > [d>0,1]	2,7E-02	3,5E+07	1,0E+00
Andujar-Doñana	wilcox.test	CDS	Watterson estimator	greater	[d<0,1] < [d>0,1]	-5,9E-05	2,6E+07	1,0E+00
Andujar-Doñana	wilcox.test	CDS	Watterson estimator	less	[d<0,1] > [d>0,1]	-5,9E-05	2,6E+07	1,4E-45
Andujar-Doñana	wilcox.test	Intergenic	divergence	greater	[d<0,1] < [d>0,1]	2,6E-05	1,8E+10	5,4E-10
Andujar-Doñana	wilcox.test	Intergenic	divergence	less	[d<0,1] > [d>0,1]	2,6E-05	1,8E+10	1,0E+00
Andujar-Doñana	wilcox.test	Intergenic	GC content	greater	[d<0,1] < [d>0,1]	-2,4E-05	1,8E+10	5,3E-01
Andujar-Doñana	wilcox.test	Intergenic	GC content	less	[d<0,1] > [d>0,1]	-2,4E-05	1,8E+10	4,7E-01
Andujar-Doñana	wilcox.test	Intergenic	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-2,2E-05	1,3E+10	1,0E+00
Andujar-Doñana	wilcox.test	Intergenic	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-2,2E-05	1,3E+10	0,0E+00
Andujar-Doñana	wilcox.test	Intergenic	recombination	greater	[d<0,1] < [d>0,1]	2,3E-02	1,4E+10	1,6E-04
Andujar-Doñana	wilcox.test	Intergenic	recombination	less	[d<0,1] > [d>0,1]	2,3E-02	1,4E+10	1,0E+00
Andujar-Doñana	wilcox.test	Intergenic	Skewness	greater	[d<0,1] < [d>0,1]	1,9E-02	2,2E+10	0,0E+00

Andujar-Doñana	wilcox.test	Intergenic	Skewness	less	[d<0,1] > [d>0,1]	1,9E-02	2,2E+10	1,0E+00
Andujar-Doñana	wilcox.test	Intergenic	Watterson estimator	greater	[d<0,1] < [d>0,1]	-6,9E-05	1,3E+10	1,0E+00
Andujar-Doñana	wilcox.test	Intergenic	Watterson estimator	less	[d<0,1] > [d>0,1]	-6,9E-05	1,3E+10	0,0E+00
Andujar-Doñana	wilcox.test	Intron	divergence	greater	[d<0,1] < [d>0,1]	3,3E-05	6,7E+09	9,8E-05
Andujar-Doñana	wilcox.test	Intron	divergence	less	[d<0,1] > [d>0,1]	3,3E-05	6,7E+09	1,0E+00
Andujar-Doñana	wilcox.test	Intron	functional sites percentage	greater	[d<0,1] < [d>0,1]	-6,2E-02	1,4E+09	1,0E+00
Andujar-Doñana	wilcox.test	Intron	functional sites percentage	less	[d<0,1] > [d>0,1]	-6,2E-02	1,4E+09	5,9E-04
Andujar-Doñana	wilcox.test	Intron	GC content	greater	[d<0,1] < [d>0,1]	1,5E-03	6,8E+09	1,1E-06
Andujar-Doñana	wilcox.test	Intron	GC content	less	[d<0,1] > [d>0,1]	1,5E-03	6,8E+09	1,0E+00
Andujar-Doñana	wilcox.test	Intron	Nucleotide diversity	greater	[d<0,1] < [d>0,1]	-4,8E-05	5,2E+09	1,0E+00
Andujar-Doñana	wilcox.test	Intron	Nucleotide diversity	less	[d<0,1] > [d>0,1]	-4,8E-05	5,2E+09	0,0E+00
Andujar-Doñana	wilcox.test	Intron	recombination	greater	[d<0,1] < [d>0,1]	2,5E-02	6,0E+09	2,0E-03
Andujar-Doñana	wilcox.test	Intron	recombination	less	[d<0,1] > [d>0,1]	2,5E-02	6,0E+09	1,0E+00
Andujar-Doñana	wilcox.test	Intron	RVIS	greater	[d<0,1] < [d>0,1]	-3,1E-03	3,4E+09	7,4E-01
Andujar-Doñana	wilcox.test	Intron	RVIS	less	[d<0,1] > [d>0,1]	-3,1E-03	3,4E+09	2,6E-01
Andujar-Doñana	wilcox.test	Intron	Skewness	greater	[d<0,1] < [d>0,1]	2,1E-02	8,1E+09	0,0E+00
Andujar-Doñana	wilcox.test	Intron	Skewness	less	[d<0,1] > [d>0,1]	2,1E-02	8,1E+09	1,0E+00
Andujar-Doñana	wilcox.test	Intron	Watterson estimator	greater	[d<0,1] < [d>0,1]	-4,5E-05	5,2E+09	1,0E+00
Andujar-Doñana	wilcox.test	Intron	Watterson estimator	less	[d<0,1] > [d>0,1]	-4,5E-05	5,2E+09	0,0E+00

Table S9. r effect size for Man-Whitney two sample rank sum test comparing $\delta_S\theta_W > 0.1$ windows vs. $\delta_S\theta_W < 0.1$ windows for NB θ_W, NB π, NB S, divergence, recombination, GC content, functional site percentage, and RVIS. Features considered for the analysis are intergenic, introns and CDS.

Comparison	Feature	Genomic variable	r
Kirov-Poland	CDS	Divergence	0,067
Kirov-Poland	CDS	Functional sites percentage	-0,003
Kirov-Poland	CDS	GC content	0,063
Kirov-Poland	CDS	NB Nucleotide diversity	0,032
Kirov-Poland	CDS	NB Skewness SFS	-0,024
Kirov-Poland	CDS	NB Watterson estimator	0,039
Kirov-Poland	CDS	Recombination rate	0,033
Kirov-Poland	CDS	RVIS	0,049
Kirov-Poland	Intergenic	Divergence	0,054
Kirov-Poland	Intergenic	GC content	0,055
Kirov-Poland	Intergenic	NB Nucleotide diversity	-0,039
Kirov-Poland	Intergenic	NB Skewness SFS	0,042
Kirov-Poland	Intergenic	NB Watterson estimator	-0,023
Kirov-Poland	Intergenic	Recombination rate	0,028
Kirov-Poland	Intron	Divergence	0,054
Kirov-Poland	Intron	Functional sites percentage	-0,010
Kirov-Poland	Intron	GC content	0,068
Kirov-Poland	Intron	NB Nucleotide diversity	-0,026
Kirov-Poland	Intron	NB Skewness SFS	0,028
Kirov-Poland	Intron	NB Watterson estimator	-0,013
Kirov-Poland	Intron	Recombination rate	0,043
Kirov-Poland	Intron	RVIS	0,013
Kirov-Norway	CDS	Divergence	-0,014
Kirov-Norway	CDS	Functional sites percentage	-0,001
Kirov-Norway	CDS	GC content	-0,002
Kirov-Norway	CDS	NB Nucleotide diversity	-0,174
Kirov-Norway	CDS	NB Skewness SFS	0,155
Kirov-Norway	CDS	NB Watterson estimator	-0,175
Kirov-Norway	CDS	Recombination rate	-0,024
Kirov-Norway	CDS	RVIS	-0,017
Kirov-Norway	Intergenic	Divergence	-0,003

Kirov-Norway	Intergenic	Divergence	-0,003
Kirov-Norway	Intergenic	GC content	-0,009
Kirov-Norway	Intergenic	NB Nucleotide diversity	-0,218
Kirov-Norway	Intergenic	NB Skewness SFS	0,184
Kirov-Norway	Intergenic	NB Watterson estimator	-0,215
Kirov-Norway	Intergenic	Recombination rate	-0,019
Kirov-Norway	Intron	Divergence	-0,002
Kirov-Norway	Intron	Functional sites percentage	0,004
Kirov-Norway	Intron	GC content	0,001
Kirov-Norway	Intron	NB Nucleotide diversity	-0,205
Kirov-Norway	Intron	NB Skewness SFS	0,175
Kirov-Norway	Intron	NB Watterson estimator	-0,202
Kirov-Norway	Intron	Recombination rate	-0,003
Kirov-Norway	Intron	RVIS	0,012
Andujar-Doñana	CDS	Divergence	-0,005
Andujar-Doñana	CDS	Functional sites percentage	-0,002
Andujar-Doñana	CDS	GC content	0,024
Andujar-Doñana	CDS	NB Nucleotide diversity	-0,093
Andujar-Doñana	CDS	NB Skewness SFS	0,085
Andujar-Doñana	CDS	NB Watterson estimator	-0,092
Andujar-Doñana	CDS	Recombination rate	-0,003
Andujar-Doñana	CDS	RVIS	-0,020
Andujar-Doñana	Intergenic	Divergence	0,008
Andujar-Doñana	Intergenic	GC content	0,000
Andujar-Doñana	Intergenic	NB Nucleotide diversity	-0,137
Andujar-Doñana	Intergenic	NB Skewness SFS	0,130
Andujar-Doñana	Intergenic	NB Watterson estimator	-0,133
Andujar-Doñana	Intergenic	Recombination rate	0,005
Andujar-Doñana	Intron	Divergence	0,006
Andujar-Doñana	Intron	Functional sites percentage	-0,005
Andujar-Doñana	Intron	GC content	0,008
Andujar-Doñana	Intron	NB Nucleotide diversity	-0,126
Andujar-Doñana	Intron	NB Skewness SFS	0,119
Andujar-Doñana	Intron	NB Watterson estimator	-0,123
Andujar-Doñana	Intron	Recombination rate	0,005
Andujar-Doñana	Intron	RVIS	-0,001

Table S10. Wilcoxon signed-rank test comparing ND_{NB}-ND_{B} windows vs. ND_{NB}-D_{B} windows for divergence, recombination, GC content, functional site percentage, and RVIS. Features considered for the analysis are intergenic, introns and CDS.

Comparison	Method	Feature	Genomic variable	Alternative hypothesis	Interpretation	Estimate	Statistic	p-value
Kirov-Poland	wilcox test	CDS	Divergence	greater	ND-ND>ND-D	-5,6E-04	5,3E+06	1,0E+00
Kirov-Poland	wilcox test	Intergenic	Divergence	greater	ND-ND>ND-D	-2,6E-04	2,8E+09	1,0E+00
Kirov-Poland	wilcox test	Intron	Divergence	greater	ND-ND>ND-D	-3,4E-04	1,1E+09	1,0E+00
Kirov-Poland	wilcox test	CDS	Divergence	less	ND-ND<ND-D	-5,6E-04	5,3E+06	6,5E-23
Kirov-Poland	wilcox test	Intergenic	Divergence	less	ND-ND<ND-D	-2,6E-04	2,8E+09	1,9E-257
Kirov-Poland	wilcox test	Intron	Divergence	less	ND-ND<ND-D	-3,4E-04	1,1E+09	4,9E-161
Kirov-Poland	wilcox test	CDS	Functional sites percentage	greater	ND-ND>ND-D	2,7E-01	1,2E+06	5,8E-02
Kirov-Poland	wilcox test	Intergenic	Functional sites percentage	greater	ND-ND>ND-D	1,2E-01	3,5E+08	3,7E-07
Kirov-Poland	wilcox test	Intron	Functional sites percentage	greater	ND-ND>ND-D	1,8E-01	2,9E+08	6,4E-08
Kirov-Poland	wilcox test	CDS	Functional sites percentage	less	ND-ND<ND-D	2,7E-01	1,2E+06	9,4E-01
Kirov-Poland	wilcox test	Intergenic	Functional sites percentage	less	ND-ND<ND-D	1,2E-01	3,5E+08	1,0E+00
Kirov-Poland	wilcox test	Intron	Functional sites percentage	less	ND-ND<ND-D	1,8E-01	2,9E+08	1,0E+00
Kirov-Poland	wilcox test	CDS	GC content	greater	ND-ND>ND-D	-2,5E-02	5,5E+06	1,0E+00
Kirov-Poland	wilcox test	Intergenic	GC content	greater	ND-ND>ND-D	-8,3E-03	2,8E+09	1,0E+00
Kirov-Poland	wilcox test	Intron	GC content	greater	ND-ND>ND-D	-1,6E-02	1,1E+09	1,0E+00
Kirov-Poland	wilcox test	CDS	GC content	less	ND-ND<ND-D	-2,5E-02	5,5E+06	1,3E-17
Kirov-Poland	wilcox test	Intergenic	GC content	less	ND-ND<ND-D	-8,3E-03	2,8E+09	1,5E-205
Kirov-Poland	wilcox test	Intron	GC content	less	ND-ND<ND-D	-1,6E-02	1,1E+09	6,0E-261
Kirov-Poland	wilcox test	CDS	Recombination	greater	ND-ND>ND-D	-3,2E-01	5,2E+06	1,0E+00
Kirov-Poland	wilcox test	Intergenic	Recombination	greater	ND-ND>ND-D	-1,8E-01	2,4E+09	1,0E+00
Kirov-Poland	wilcox test	Intron	Recombination	greater	ND-ND>ND-D	-3,1E-01	1,0E+09	1,0E+00
Kirov-Poland	wilcox test	CDS	Recombination	less	ND-ND<ND-D	-3,2E-01	5,2E+06	2,9E-07

Kirov-Poland	wilcox test	Intergenic	Recombination	less	ND-ND<ND-D	-1,8E-01	2,4E+09	9,0E-71
Kirov-Poland	wilcox test	Intron	Recombination	less	ND-ND<ND-D	-3,1E-01	1,0E+09	8,4E-94
Kirov-Poland	wilcox test	CDS	RVIS	greater	ND-ND>ND-D	-1,9E-01	3,7E+06	1,0E+00
Kirov-Poland	wilcox test	Intron	RVIS	greater	ND-ND>ND-D	-4,5E-02	6,3E+08	1,0E+00
Kirov-Poland	wilcox test	CDS	RVIS	less	ND-ND<ND-D	-1,9E-01	3,7E+06	5,4E-10
Kirov-Poland	wilcox test	Intron	RVIS	less	ND-ND<ND-D	-4,5E-02	6,3E+08	5,6E-08
Kirov-Norway	wilcox test	CDS	Divergence	greater	ND-ND>ND-D	-1,1E-04	1,7E+07	1,0E+00
Kirov-Norway	wilcox test	Intergenic	Divergence	greater	ND-ND>ND-D	-6,7E-05	6,2E+09	1,0E+00
Kirov-Norway	wilcox test	Intron	Divergence	greater	ND-ND>ND-D	-1,2E-04	2,5E+09	1,0E+00
Kirov-Norway	wilcox test	CDS	Divergence	less	ND-ND<ND-D	-1,1E-04	1,7E+07	3,9E-07
Kirov-Norway	wilcox test	Intergenic	Divergence	less	ND-ND<ND-D	-6,7E-05	6,2E+09	1,8E-67
Kirov-Norway	wilcox test	Intron	Divergence	less	ND-ND<ND-D	-1,2E-04	2,5E+09	7,9E-43
Kirov-Norway	wilcox test	CDS	Functional sites percentage	greater	ND-ND>ND-D	8,8E-02	3,1E+06	1,8E-01
Kirov-Norway	wilcox test	Intergenic	Functional sites percentage	greater	ND-ND>ND-D	-4,7E-02	6,6E+08	1,0E+00
Kirov-Norway	wilcox test	Intron	Functional sites percentage	greater	ND-ND>ND-D	-5,3E-02	5,6E+08	9,9E-01
Kirov-Norway	wilcox test	CDS	Functional sites percentage	less	ND-ND<ND-D	8,8E-02	3,1E+06	8,2E-01
Kirov-Norway	wilcox test	Intergenic	Functional sites percentage	less	ND-ND<ND-D	-4,7E-02	6,6E+08	1,8E-03
Kirov-Norway	wilcox test	Intron	Functional sites percentage	less	ND-ND<ND-D	-5,3E-02	5,6E+08	5,2E-03
Kirov-Norway	wilcox test	CDS	GC content	greater	ND-ND>ND-D	-7,8E-03	1,7E+07	1,0E+00
Kirov-Norway	wilcox test	Intergenic	GC content	greater	ND-ND>ND-D	-1,8E-03	6,3E+09	1,0E+00
Kirov-Norway	wilcox test	Intron	GC content	greater	ND-ND>ND-D	-6,0E-03	2,4E+09	1,0E+00
Kirov-Norway	wilcox test	CDS	GC content	less	ND-ND<ND-D	-7,8E-03	1,7E+07	3,8E-06
Kirov-Norway	wilcox test	Intergenic	GC content	less	ND-ND<ND-D	-1,8E-03	6,3E+09	1,1E-25

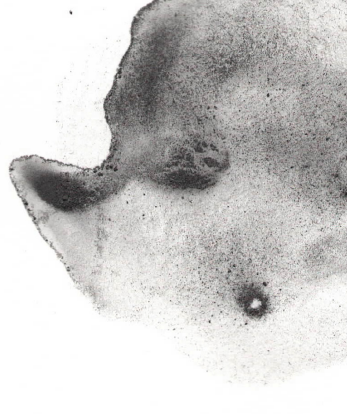
Kirov-Norway	wilcox test	Intron	GC content	less	ND-ND<ND-D	-6,0E-03	2,4E+09	6,2E-88
Kirov-Norway	wilcox test	CDS	Recombination	greater	ND-ND>ND-D	7,7E-02	1,6E+07	1,1E-02
Kirov-Norway	wilcox test	Intergenic	Recombination	greater	ND-ND>ND-D	1,8E-05	5,2E+09	3,3E-01
Kirov-Norway	wilcox test	Intron	Recombination	greater	ND-ND>ND-D	-6,3E-02	2,3E+09	1,0E+00
Kirov-Norway	wilcox test	CDS	Recombination	less	ND-ND<ND-D	7,7E-02	1,6E+07	9,9E-01
Kirov-Norway	wilcox test	Intergenic	Recombination	less	ND-ND<ND-D	1,8E-05	5,2E+09	6,7E-01
Kirov-Norway	wilcox test	Intron	Recombination	less	ND-ND<ND-D	-6,3E-02	2,3E+09	4,5E-12
Kirov-Norway	wilcox test	CDS	RVIS	greater	ND-ND>ND-D	-2,0E-02	1,1E+07	8,5E-01
Kirov-Norway	wilcox test	Intron	RVIS	greater	ND-ND>ND-D	-5,8E-02	1,3E+09	1,0E+00
Kirov-Norway	wilcox test	CDS	RVIS	less	ND-ND<ND-D	-2,0E-02	1,1E+07	1,5E-01
Kirov-Norway	wilcox test	Intron	RVIS	less	ND-ND<ND-D	-5,8E-02	1,3E+09	1,7E-23
Andujar-Doñana	wilcox test	CDS	Divergence	greater	ND-ND>ND-D	-6,4E-05	1,7E+07	9,8E-01
Andujar-Doñana	wilcox test	Intergenic	Divergence	greater	ND-ND>ND-D	-1,3E-04	7,6E+09	1,0E+00
Andujar-Doñana	wilcox test	Intron	Divergence	greater	ND-ND>ND-D	-1,2E-04	3,1E+09	1,0E+00
Andujar-Doñana	wilcox test	CDS	Divergence	less	ND-ND<ND-D	-6,4E-05	1,7E+07	2,3E-02
Andujar-Doñana	wilcox test	Intergenic	Divergence	less	ND-ND<ND-D	-1,3E-04	7,6E+09	8,9E-87
Andujar-Doñana	wilcox test	Intron	Divergence	less	ND-ND<ND-D	-1,2E-04	3,1E+09	1,0E-37
Andujar-Doñana	wilcox test	CDS	Functional sites percentage	greater	ND-ND>ND-D	7,8E-02	3,1E+06	2,3E-01
Andujar-Doñana	wilcox test	Intergenic	Functional sites percentage	greater	ND-ND>ND-D	-1,1E-02	8,6E+08	8,5E-01
Andujar-Doñana	wilcox test	Intron	Functional sites percentage	greater	ND-ND>ND-D	8,6E-02	7,2E+08	7,8E-05
Andujar-Doñana	wilcox test	CDS	Functional sites percentage	less	ND-ND<ND-D	7,8E-02	3,1E+06	7,7E-01
Andujar-Doñana	wilcox test	Intergenic	Functional sites percentage	less	ND-ND<ND-D	-1,1E-02	8,6E+08	1,5E-01
Andujar-Doñana	wilcox test	Intron	Functional sites percentage	less	ND-ND<ND-D	8,6E-02	7,2E+08	1,0E+00

Andujar-Doñana	wilcox test	CDS	GC content	greater	ND-ND>ND-D	-9,0E-03	1,7E+07	1,0E+00
Andujar-Doñana	wilcox test	Intergenic	GC content	greater	ND-ND>ND-D	-2,2E-03	7,7E+09	1,0E+00
Andujar-Doñana	wilcox test	Intron	GC content	greater	ND-ND>ND-D	-4,6E-03	3,0E+09	1,0E+00
Andujar-Doñana	wilcox test	CDS	GC content	less	ND-ND<ND-D	-9,0E-03	1,7E+07	1,7E-06
Andujar-Doñana	wilcox test	Intergenic	GC content	less	ND-ND<ND-D	-2,2E-03	7,7E+09	1,5E-30
Andujar-Doñana	wilcox test	Intron	GC content	less	ND-ND<ND-D	-4,6E-03	3,0E+09	1,8E-45
Andujar-Doñana	wilcox test	CDS	Recombination	greater	ND-ND>ND-D	-1,0E-02	1,6E+07	6,5E-01
Andujar-Doñana	wilcox test	Intergenic	Recombination	greater	ND-ND>ND-D	-7,7E-02	6,2E+09	1,0E+00
Andujar-Doñana	wilcox test	Intron	Recombination	greater	ND-ND>ND-D	-6,8E-02	2,8E+09	1,0E+00
Andujar-Doñana	wilcox test	CDS	Recombination	less	ND-ND<ND-D	-1,0E-02	1,6E+07	3,5E-01
Andujar-Doñana	wilcox test	Intergenic	Recombination	less	ND-ND<ND-D	-7,7E-02	6,2E+09	1,5E-26
Andujar-Doñana	wilcox test	Intron	Recombination	less	ND-ND<ND-D	-6,8E-02	2,8E+09	7,1E-12
Andujar-Doñana	wilcox test	CDS	RVIS	greater	ND-ND>ND-D	7,2E-03	1,2E+07	3,7E-01
Andujar-Doñana	wilcox test	Intron	RVIS	greater	ND-ND>ND-D	-1,5E-02	1,6E+09	9,9E-01
Andujar-Doñana	wilcox test	CDS	RVIS	less	ND-ND<ND-D	7,2E-03	1,2E+07	6,3E-01
Andujar-Doñana	wilcox test	Intron	RVIS	less	ND-ND<ND-D	-1,5E-02	1,6E+09	5,9E-03

Table S11. r effect size for Man-Whitney two sample rank sum test comparing ND_{NB}-ND_{B} windows vs. ND_{NB}-D_{B} windows for divergence, recombination, GC content, functional site percentage, and RVIS. Features considered for the analysis are intergenic, introns and CDS.

Comparison	Feature	Genomic variable	r
Kirov-Poland	CDS	Divergence	-0,085
Kirov-Poland	CDS	Functional sites percentage	0,014
Kirov-Poland	CDS	GC content	-0,073
Kirov-Poland	CDS	Recombination rate	-0,043
Kirov-Poland	CDS	RVIS	-0,053
Kirov-Poland	Intergenic	Divergence	-0,068
Kirov-Poland	Intergenic	Functional sites percentage	0,010
Kirov-Poland	Intergenic	GC content	-0,061
Kirov-Poland	Intergenic	Recombination rate	-0,035
Kirov-Poland	Intron	Divergence	-0,067
Kirov-Poland	Intron	Functional sites percentage	0,013
Kirov-Poland	Intron	GC content	-0,085
Kirov-Poland	Intron	Recombination rate	-0,051
Kirov-Poland	Intron	RVIS	-0,013
Kirov-Norway	CDS	Divergence	-0,043
Kirov-Norway	CDS	Functional sites percentage	0,008
Kirov-Norway	CDS	GC content	-0,039
Kirov-Norway	CDS	Recombination rate	0,020
Kirov-Norway	CDS	RVIS	-0,009
Kirov-Norway	Intergenic	Divergence	-0,034
Kirov-Norway	Intergenic	Functional sites percentage	-0,006
Kirov-Norway	Intergenic	GC content	-0,021
Kirov-Norway	Intergenic	Recombination rate	0,001
Kirov-Norway	Intron	Divergence	-0,034
Kirov-Norway	Intron	Functional sites percentage	-0,006
Kirov-Norway	Intron	GC content	-0,049
Kirov-Norway	Intron	Recombination rate	-0,017
Kirov-Norway	Intron	RVIS	-0,025
Andujar-Doñana	CDS	Divergence	-0,016
Andujar-Doñana	CDS	Functional sites percentage	0,006
Andujar-Doñana	CDS	GC content	-0,036
Andujar-Doñana	CDS	Recombination rate	-0,003
Andujar-Doñana	CDS	RVIS	0,003
Andujar-Doñana	Intergenic	Divergence	-0,034
Andujar-Doñana	Intergenic	Functional sites percentage	-0,002
Andujar-Doñana	Intergenic	GC content	-0,019
Andujar-Doñana	Intergenic	Recombination rate	-0,018
Andujar-Doñana	Intron	Divergence	-0,027
Andujar-Doñana	Intron	Functional sites percentage	0,008
Andujar-Doñana	Intron	GC content	-0,030
Andujar-Doñana	Intron	Recombination rate	-0,014
Andujar-Doñana	Intron	RVIS	-0,005

DISCUSIÓN



En esta tesis se ahonda en el conocimiento de la historia evolutiva de lince boreal y de lince ibérico y del rol que han podido tener distintos procesos, naturales y antrópicos, en el modelado de los genomas de las especies estudiadas, con un enfoque interpoblacional (Capítulo 1, 2, 3 y 4) y/o intragenómico (Capítulo 3, 4), y desde una escala más macroevolutiva (Capítulo 1 y 2) hasta una más microevolutiva (Capítulo 3 y 4).

Específicamente, hemos ampliado el conocimiento sobre la dinámica temporal del tamaño poblacional, las relaciones y el estatus genético de la mayoría de poblaciones contemporáneas de lince boreal (Capítulo 1). También hemos profundizado sobre la dinámica temporal de los patrones genéticos de lince boreal en Europa durante el Pleistoceno Tardío y el Holoceno, haciendo especial hincapié en la población extinta que habitó la Península Ibérica hace 2 kya (Capítulo 2). Hemos aportado nuevo conocimiento a la historia evolutiva de los últimos miles de años de lince ibérico gracias al análisis de muestras antiguas secuenciadas a genoma completo y su comparación con muestras contemporáneas (Capítulo 3). Por último, hemos investigado los cambios en diversidad a corto plazo que producen los cuellos de botella a lo largo del genoma usando poblaciones con un declive muy marcado frente a otras poblaciones demográficamente más estables de lince ibérico y boreal, y hemos discutido cómo esa heterogeneidad genómica está ligada a la actuación presente y pasada de distintas fuerzas evolutivas y la interacción entre estas (Capítulo 4). El análisis de las poblaciones usando datos de genoma completo, y cuando ha sido posible incluyendo individuos antiguos, ha permitido contestar las preguntas planteadas en los objetivos iniciales, y ha abierto además la puerta a nuevas cuestiones por resolver.

Los patrones de diversidad y estructura genética del lince boreal y del lince ibérico han venido determinados por cambios en su área de distribución y el tamaño efectivo de las distintas poblaciones durante el Pleistoceno y el Holoceno debido a distintas causas naturales y antrópicas.

Las variaciones climáticas han dado lugar a cambios en el área de distribución del lince boreal y han propiciado, por un lado, la aparición de distintos linajes a nivel mitocondrial que tienen a día de hoy una distribución más o menos localizada en el espacio y, por otro, señales de mezcla a nivel nuclear entre poblaciones geográficamente cercanas. Concretamente, el comienzo de la edad de hielo hace aproximadamente 110 mil años coincide aproximadamente con la divergencia de un linaje a nivel mitocondrial que a día de hoy ocupa Balcanes y Cáucaso. De igual modo, otros ciclos fríos, por ejemplo los acaecidos hace alrededor de 40 mil años (Kindler et al., 2014) o durante el Último Máximo Glacial hace en torno a 20 mil años (Mix, 2001; Clark et al., 2009) parecen estar detrás de la emergencia de otros linajes mitocondriales en lince boreal. Posiblemente, estos linajes evolucionaron tras el confinamiento de las poblaciones en refugios glaciales, en los que el hábitat era el adecuado. Al igual que los ciclos fríos parecen estar detrás de aislamiento y divergencia de distintas poblaciones, los ciclos interglaciares y ya más recientemente el calentamiento durante el Holoceno y los cambios de vegetación asociados, habrían impulsado una expansión del rango del lince boreal en Europa occidental y posibilitado el contacto secundario y la mezcla entre linajes divergentes. Específicamente, durante el final del Pleistoceno llegó al oeste de Europa un linaje similar a los que ocupan hoy Balcanes y Cáucaso. Posteriormente, y ya durante el interfaz entre el Holoceno y el Pleistoceno el lince boreal penetró en la península ibérica ocupando la cornisa cantábrica (Rodríguez-Varela et al., 2016). A partir de esta fecha y durante el Holoceno, el lince boreal que se encuentra en Europa occidental parece estar más relacionado mitocondrialmente con poblaciones centro-europeas y asiáticas que eventualmente se mezclaron con los linajes más sureños.

Además de las variaciones climáticas, la acción humana ha afectado a buena parte de las poblaciones de lince boreal, sobre todo en Europa, ya sea mediante acciones directas como la persecución y la caza (Linnell et al., 2010; Clavero & Delibes, 2013), como mediante acciones indirectas que han podido provocar el declive de sus presas y el deterioro, fragmentación y contracción de su hábitat (Valbuena-Carabaña et al., 2010; Marín-Arroyo, 2013; Chen et al., 2019; García-Ruiz et al., 2020; Melovski et al., 2020). Esta acciones, posiblemente empezaron durante el Pleistoceno tardío pero, sin duda, se intensificaron de manera notable durante el Holoceno, coincidiendo con la expansión humana. Especialmente la invención de la agricultura, y la llegada de la

ganadería en el Neolítico propició una expansión demográfica y una serie de cambios de uso del suelo asociados a esta que fueron haciéndose más intensos con el paso del tiempo (Peña-Chocarro et al., 2005; Gignoux, Henn & Mountain, 2011; Cubas et al., 2016, 2020; Nielsen et al., 2017; Pérez-Díaz et al., 2018). Por ejemplo, en la península ibérica debido a la actividad marítima con el uso de madera para barcos, y la ganadera con la creación de zonas de pastizal asociadas a la trashumancia, el lince boreal vio mermado su hábitat, el bosque, durante la edad media y la edad moderna (Valbuena-Carabaña et al., 2010; García-Ruiz et al., 2020). Además, en muchos casos, la expansión humana provocaba conflictos ganado-carnívoro (Clavero & Delibes, 2013), lo que sin duda fue un aliciente para la puesta en marcha de políticas de exterminio de alimañas que se llevaron a cabo en los últimos siglos (Breitenmoser, 1998; Linnell et al., 2010). Todos estos factores, sumados a la industrialización, urbanización y crecimiento exponencial de la población humana reciente, sin duda están detrás de la reducción del tamaño poblacional de lince boreal en todo su rango de distribución y de la extinción de muchas de sus poblaciones en Europa (Kratochvil, 1968; Hetherington, Lord & Jacobi, 2006; Clavero & Delibes, 2013; Rodríguez-Varela et al., 2016).

Todos estos procesos, que han afectado de forma heterogénea a las poblaciones, han modificado las fuerzas evolutivas y modelado los patrones genómicos que vemos hoy día. Por ejemplo, antes de su extinción, la disminución del tamaño poblacional en muchas poblaciones conllevó un aumento de la acción de la deriva, lo que provocó posiblemente una pérdida de diversidad y un aumento de la divergencia. Es el caso, por ejemplo, de lo ocurrido a la población de lince boreal residente en la cornisa cantábrica descrito en esta tesis (Capítulo 2). Este patrón también es observable hoy día en poblaciones europeas como Noruega y NE-Polonia, por ejemplo, que habiendo visto disminuido sus tamaños poblacionales, no llegaron a extinguirse. En contraposición, otras poblaciones como las residentes en Asia, o en el este de Europa, como la población de Kirov en Rusia, aunque con signos de declive durante el Holoceno, sostuvieron tamaños poblacionales mayores, que han hecho de estas poblaciones las más genéticamente diversas y sanas de la especie a día de hoy.

Otra fuerza evolutiva que ha sido determinante en los patrones que observamos hoy día en las poblaciones de lince es el flujo génico acaecido tras cambios en el área de

distribución y la conectividad entre distintas poblaciones. El primer ejemplo es la mezcla que se observa a nivel nuclear de los linajes mitocondriales típicos del sur con linajes centro-europeos. Este mezcla, como se ha mencionado anteriormente, tuvo lugar probablemente en el centro y el oeste de Europa durante el Holoceno. Por otro lado, encontramos un contacto secundario de las poblaciones europeas y asiáticas en la población de Tuva en Asia. Estos contactos secundarios de linajes divergentes provocan un aumento de la diversidad en la población mezclada, y por ejemplo, Tuva presenta los niveles más altos de diversidad en la especie.

En esta tesis, hemos detectado, además, un contacto secundario a nivel de especies entre lince boreal y el ibérico. Este contacto secundario ha sido concomitante a una ganancia de diversidad en la especie ibérica, percibido gracias a la caracterización de la población de lince ibérico que habitó la Península Ibérica entre 4000 y 2000 años atrás. Esta población antigua, que era genéticamente homogénea y similar a la actual Andújar, presentaba una diversidad genética más baja que la de las poblaciones remanentes en el 2000, Doñana y Andújar. Aparentemente, el proceso de mezcla que pudo aumentar esa diversidad fue progresivo, de modo que nuestra muestra más antigua (4300 años) presenta menos alelos compartidos con lince boreal que las muestras de hace 2000-2500 años, y estas a su vez menos que las dos poblaciones contemporáneas, las cuales presentan valores similares indicando que el proceso de mezcla tuvo lugar antes de la separación de estas dos poblaciones contemporáneas hace 200 años. ¿Qué poblaciones de lince boreal estuvieron involucradas en la hibridación? A pesar de que como hemos visto, el lince boreal ocupaba en ese momento la cornisa cantábrica, nuestros análisis detectan más alelos compartidos con las poblaciones contemporáneas de lince boreal, que con el individuo antiguo ibérico. Este patrón apoyaría que la población fuente fuese una población híbrida y ancestral a las poblaciones contemporáneas, mientras que la población de lince boreal que llegó al comienzo del Holoceno a la península ibérica, pudo ser anterior a esa mezcla de linajes y quedar aislada en zonas de bosque boreal en la cornisa cantábrica hasta su desaparición en el siglo XX (Clavero & Delibes, 2013). En línea con esta hipótesis encontramos que el lince ibérico estuvo presente en zonas del este de Francia hasta hace, al menos 3000 años (Rodríguez-Varela et al., 2015), por lo que no es descabellado pensar que el contacto tuviese lugar en zonas próximas a la costa mediterránea de la península ibérica y Francia. Este intercambio de alelos, perceptible a día de hoy en

todos los individuos ibéricos contemporáneos analizados, supuso muy probablemente para la especie ibérica una ganancia de diversidad al incorporar nuevos alelos de boreal, sin embargo sólo podemos especular con respecto a las consecuencias de la entrada de nuevos alelos con respecto a la eficacia biológica. Por un lado, las especies divergieron en alopatría y mantienen diferencias morfológicas y comportamentales (Sunquist & Sunquist, 2002; Nowak, 2005), lo que podría derivar en depresión por exogamia. Por otro lado, el número de cromosomas en ambas especies es idéntico y no esperamos encontrar grandes reorganizaciones cromosómicas (Abascal et al., 2016). Además, cada vez parece más claro que las especies no son compartimentos estancos, y que especialmente los felinos, y particularmente los linces, han tenido a lo largo de su historia evolutiva eventos recurrentes de mezcla entre las especies (Abascal et al., 2016; Li et al., 2016). Sin embargo, el hecho de que el periodo de híbridación y aumento de la diversidad coincida con un periodo de fragmentación y declive de las poblaciones de lince ibérico por causas antrópicas, hacen imposible valorar las consecuencias que estos procesos de mezcla han tenido para la eficacia biológica y la viabilidad de las poblaciones. Y es que, a pesar de esa historia común de mezcla, las poblaciones remanentes de lince ibérico sufrieron durante los últimos siglos un declive generalizado que acabó con gran parte de la especie y que sin embargo fue desigual entre poblaciones (Casas-Marce et al., 2017).

Esta heterogeneidad poblacional perceptible en las dos especies, lince ibérico y boreal, nos ha permitido investigar cómo distintos procesos evolutivos interaccionan entre sí a lo largo del genoma. Es decir, valiéndonos de la heterogeneidad poblacional, que nos da escenarios de estudio distintos, investigamos la heterogeneidad a lo largo del genoma. Para ello, hemos usado poblaciones sometidas a una intensa deriva genética, como son Doñana para lince ibérico, y NE-Polonia y Noruega para lince boreal, y las hemos comparado con poblaciones que no han sufrido esos procesos de deriva tan dramáticos y que se asemejan a los linajes ancestrales de las poblaciones en estudio, como son Andújar para lince ibérico y Kirov para boreal.

Encontramos que, efectivamente, aunque la acción de la deriva haya sido generalizada, sus efectos sobre la diversidad genética han sido muy variables en distintos compartimentos del genoma . Concretamente observamos que mientras que las zonas neutrales tienen una tendencia a la pérdida de diversidad, como esperamos por la teoría (Nei & Maruyama, 1975), esta pérdida es menor en regiones que están

sometidas a selección. Incluso, en algunos casos, regiones muy conservadas como elementos ultraconservados que no se transcriben (UCNE en sus siglas en inglés) presentan más diversidad en las poblaciones sometidas a deriva que en las que no. Este efecto es consistente entre las comparaciones que incluyen poblaciones muy erosionadas genéticamente, como Noruega y Doñana.

En aquellas zonas donde la diversidad es generalmente baja por efecto de la selección, cualquier ganancia de diversidad, por muy pequeña que sea, y aunque se produzca por efecto de la oscilación aleatoria que provoca la deriva, va a ser evidente, mientras que en zonas de alta diversidad, se va a percibir principalmente la pérdida de la misma. En nuestro trabajo, hemos detectado que estas zonas que ganan diversidad son además zonas de alta mutación, por lo que los alelos que se mantenían a baja frecuencia, más comunes en estas zonas de alta mutación, pueden aumentar ahora su frecuencia y pasar a ser perceptibles globalmente como ganancia de diversidad, o al menos como una menor pérdida de diversidad que en regiones con menor tasa de mutación.

¿Por qué aumentan en frecuencia estos alelos, que en poblaciones no sometidas a deriva pasaban desapercibidos, hasta ser ahora perceptibles en las poblaciones más erosionadas? La acumulación de estos alelos podría deberse a la relajación de la selección purificadora que se produce como consecuencia de la acción de la deriva (Kirkpatrick & Jarne, 2000). El hecho de que en poblaciones con un nivel de deriva más alto, encontramos que la acumulación de alelos se produce hasta en genes muy intolerantes a los cambios, mientras que cuando la deriva es más leve, son genes más tolerantes los que acumulan diversidad, parece que apoya el hecho que la relajación de la selección purificadora podría estar detrás del patrón observado. Además, el hecho de que las regiones sometidas a selección se acercan al equilibrio más rápido que las zonas neutrales, haría que este aumento en frecuencia de estos alelos ocurriese más rápido, y fuese más notable (Gordo & Dionisio, 2005; Song & Steinrücken, 2012; Pennings, Kryazhimskiy & Wakeley, 2014; Brandvain & Wright, 2016).

Implicaciones en conservación

Aunque esta tesis parte de un punto de vista evolutivo, los resultados obtenidos tienen implicaciones en la conservación de las dos especies estudiadas.

Por un lado, hemos caracterizado genéticamente las poblaciones asiáticas de lince boreal, que hasta la fecha eran relativamente desconocidas y hemos ayudado a clarificar la historia evolutiva y la taxonomía de la especie. Además hemos descubierto que la especie se encuentra en un declive generalizado incluso de las poblaciones que se creían más demográficamente sanas. Estos resultados, unidos al hecho de que la especie tiene una diversidad excepcionalmente baja a pesar de su amplia distribución global, y de que está sometida a un manejo cinegético en gran parte de su área de distribución, llaman a la cautela con respecto a la viabilidad a medio-largo plazo de la misma. Adicionalmente, en las poblaciones europeas más erosionadas, entre las que destaca la población de Balcanes, sería recomendable tomar medidas proactivas que restablecieran el flujo entre las poblaciones y disminuyeran su riesgo inmediato de depresión por endogamia, y restablecieran tamaños efectivos suficientemente altos como para permitir la acción de la selección purificadora. Estas recomendaciones van en la línea de las actuaciones de conservación que se están llevando a cabo actualmente en Europa. Por un lado, el manejo cinegético de las poblaciones con menor tamaño poblacional está sujeto a vedas y cuotas (Linnell et al., 2010). Por otro, hay en marcha programas de reintroducción de la especie en Europa que pretenden recuperar, al menos parcialmente, su distribución histórica en el continente (Linnell et al., 2009; Mueller et al., 2020).

En la misma línea, esta tesis apoya las medidas que se están tomando para la conservación de la especie ibérica, que pretenden restablecer el flujo entre poblaciones y aumentar su tamaño efectivo. Estas medidas de conservación ex-situ e in-situ puestas en marcha durante los últimos 15 años han conseguido aliviar la situación demográfica de la especie, que ha pasado a tener más de 600 individuos en libertad en 2018 (http://www.iberlince.eu/images/docs/3_InformesLIFE/Informe_Censo_2018.pdf), además de una población ex situ que es gestionada demográfica y genéticamente y que está sirviendo de fuente de animales para reintroducción (Kleinman-Ruiz et al., 2019).

Por último, esta tesis abre la puerta a plantear el rescate genético de la especie ibérica por parte del lince boreal, pero llama también a la cautela con respecto al mismo. Este rescate podría aumentar el potencial adaptativo de la especie al aumentar su diversidad, y con él su viabilidad a largo plazo. Sin embargo, debido al desconocimiento de los riesgos de depresión por exogamia que podría conllevar a corto plazo, aconsejamos su valoración en un entorno controlado y aplicando la máxima cautela. Si ocurriese en el entorno natural fruto del contacto entre las dos especies, sugerimos una monitorización del resultado de estos cruces.

Perspectivas futuras

En primer lugar me gustaría mencionar que esta tesis partía desde una posición ciertamente aventajada. El lince ibérico cuenta con un genoma anotado (Abascal et al., 2016), y un extenso conocimiento de la especie en muchos aspectos como su ecología (e.g. Rodríguez & Delibes, 1992, 2003) o su historia evolutiva, y los retos genéticos y demográficos a los que se enfrenta (Palomares et al., 2012; Casas-Marce et al., 2013, 2017; Abascal et al., 2016). Además, el grupo cuenta con un banco importante de muestras, tanto de lince ibérico como de lince boreal gracias a una larga trayectoria de trabajo con las especies, que ha permitido durante todos estos años tejer una red de colaboradores tanto dentro, como fuera de las fronteras.

Partiendo de esta base, durante el desarrollo de esta tesis, además del conocimiento, se han generado nuevos recursos y se han establecido nuevas colaboraciones que suponen un punto de partida sólido para poder ahondar en nuevas cuestiones relevantes relacionadas con la evolución de los linces. A nivel de recursos, se han generado datos de genoma completo de numerosas muestras de boreal y se ha aumentado sustancialmente los previamente disponibles de lince ibérico. Estas muestras han sido, están siendo, y serán usadas en el futuro para contestar otras preguntas que han quedado abiertas. Además se ha logrado secuenciar por primera vez en nuestro grupo genomas completos de muestras antiguas, tanto de lince ibérico como boreal. Con esta tesis se ha consolidado la transición efectiva de la Genética a la Genómica, que ya habían iniciado trabajos anteriores, y que en el caso de esta tesis se concretó en los análisis de genomas completos a escala poblacional. Por el camino, como no podía ser de otro modo, se han establecido nuevas colaboraciones y se han afianzado las ya existentes.

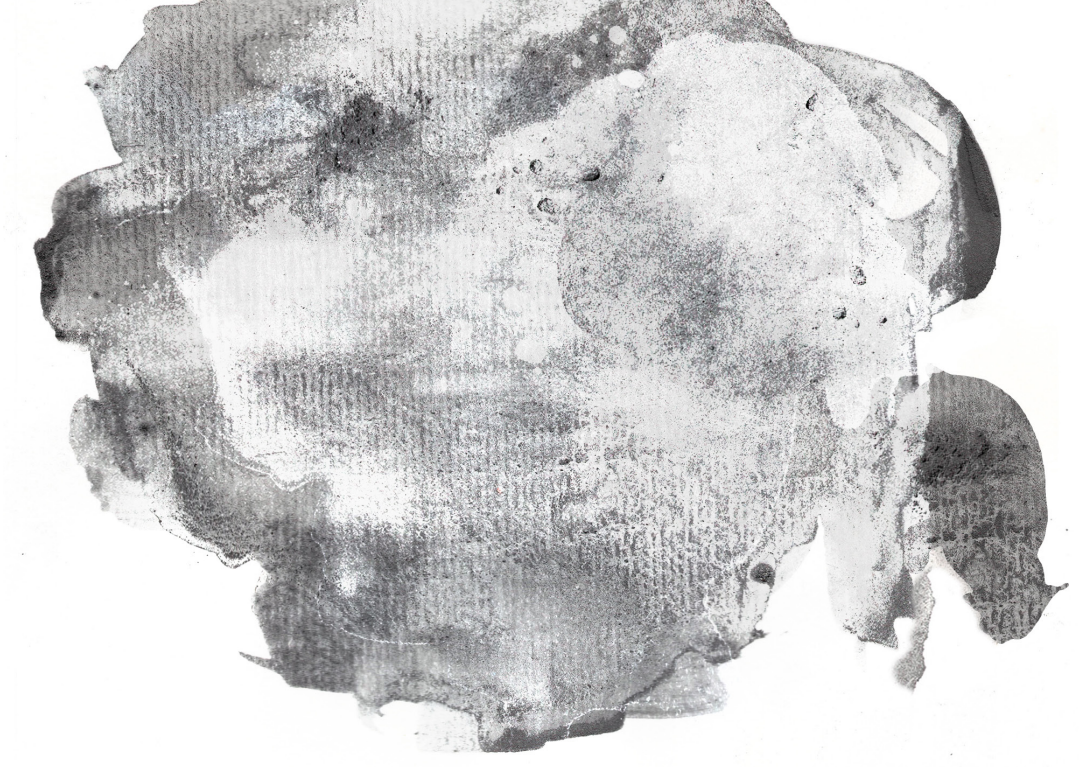
Muchas son las cuestiones que deja abierta esta tesis, algunas de las cuales ya están siendo contestadas en el marco de otros proyectos en el grupo. A nivel intraespecífico queda por resolver hasta qué punto la purga ha sido capaz de eliminar alelos deletéreos en las poblaciones sometidas a deriva, tanto de lince ibérico como de boreal. También quedan en el aire preguntas relacionadas con la existencia de adaptaciones locales en lince boreal. A un nivel de género, esta tesis abre la puerta a caracterizar los patrones de mezcla entre las especies en el tiempo y en el espacio, y a valorar sus posibles consecuencias en forma de introgresión adaptativa o de pérdida de eficacia biológica.

El continuo abaratamiento y mejora de técnicas secuenciación, junto con el desarrollo de metodologías de análisis cada vez más eficientes y curadas van a contribuir, además, a que en el futuro se aborden preguntas que hasta ahora quedaban fuera de alcance. Por ejemplo, estudios de variación epigenómica, o de variación estructural empiezan a ser una realidad para especies no modelo, incluso cuando estas involucran muestras antiguas o de museo (Hanghøj & Orlando, 2019; Resendez et al., 2019; Herrel, Joly & Danchin, 2020; Rubi, Knowles & Dantzer, 2020).

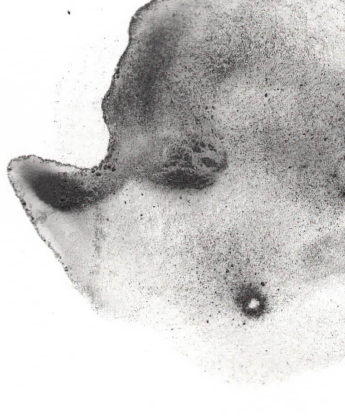
Todas estas herramientas vienen a informar y a arrojar luz sobre la evolución de las especies y su respuesta ante este escenario global de pérdida de biodiversidad en el que nos encontramos (Dirzo et al., 2014; Turvey & Crees, 2019). El conocimiento que gracias a ellas se genera, aunque se presente vacío de connotaciones éticas o morales, es una forma de acercarnos a los problemas que nos acucian y de hacer que tomemos conciencia de ellos. Solo se aprecia lo que se conoce. En cualquier caso, de nada sirve este conocimiento si no impulsa políticas que promuevan un desarrollo respetuoso con el planeta que nos ayuden a conservar el patrimonio biológico del que formamos parte.

“Only if we understand, can we care.
Only if we care, we will help.
Only if we help, we shall be saved.”

Jane Goodall



CONCLUSIONES



1. Las poblaciones de lince boreal han tenido una historia común hasta hace aproximadamente 100 mil años, cuando empezaron a divergir y entraron en un periodo de declive mantenido y generalizado en el que las poblaciones del oeste, exceptuando Kirov, mantuvieron tamaños efectivos menores que las poblaciones del este.

Eurasian lynx populations shared a common history until 100 kya, when Asian and European populations started to diverge and both entered a period of continuous and widespread decline, with western populations, except Kirov, maintaining lower effective sizes than eastern populations.

2. El declive poblacional y el aumento del aislamiento en tiempos recientes posiblemente condujo a la diferenciación genética entre poblaciones de lince boreal geográfica y ecológicamente cercanas del oeste de Europa. Por el contrario, y a pesar del amplio rango de hábitats cubierto, las poblaciones en Asia son bastante homogéneas genéticamente y muestran un patrón de aislamiento por distancia, lo que aporta poco soporte a las distintas subespecies propuestas en la literatura.

Population decline and increased isolation in more recent times likely drove the genetic differentiation between geographically and ecologically close westernmost Eurasian lynx populations. By contrast, and despite the wide range of habitats covered, populations are quite homogeneous genetically across the Asian range, showing a pattern of isolation by distance and providing little genetic support for the several proposed subspecies.

3. La divergencia nuclear y mitocondrial y los declives poblacionales del lince boreal ocurridos durante el Pleistoceno Tardío, se pueden atribuir en su mayoría a fluctuaciones climáticas e influencia humana temprana. Sin embargo, el declive mantenido y generalizado desde el Holoceno es más probablemente consecuencia de los impactos humanos que se han ido intensificando a lo largo del tiempo, especialmente en el oeste de Europa.

Mitogenomic and nuclear divergences and population declines from the Eurasian lynx starting during the Late Pleistocene can be mostly attributed to climatic fluctuations and early human influence, but the widespread and sustained decline since the Holocene is more probably the consequence of anthropogenic impacts which intensified during the last centuries, especially in western Europe.

4. La distribución de linajes mitocondriales en el tiempo y el espacio apoya la coexistencia de varios linajes de lince boreal en el oeste de Europa. Estos linajes probablemente fluctuaron en su distribución, de forma que el linaje más relacionado con Balcanes y Cáucaso fue predominante durante el Pleistoceno, mientras que el más relacionado con los que hoy día se encuentran en Europa central y Asia predomino durante el Holoceno.

Mitogenomic lineages distribution in space and time supports the long-term coexistence several lineages of Eurasian lynx in Western Europe. These two lineages probably fluctuated in distribution, with a mitogenomic lineage related to the one currently found in Balkans and Caucasus being predominant during the Pleistocene, and another related to those found in Central-Europe and Asia during the beginning of the Holocene.

5. La muestra de lince boreal que habitó la península ibérica hace 2 mil años formaba parte de un linaje con extremadamente baja diversidad que está actualmente extinto y que compartía cierto grado de ancestría con el linaje Asiático y el del Cáucaso, y que pudo permanecer aislada durante milenios hasta su extinción.

The Eurasian lynx sample from the Iberian Peninsula dated 2 kya was part of an extremely low-diverse population of Eurasian lynx that shares partial ancestry with current Asian and Caucasian population. This population may have remained isolated during millennia until its extinction.

6. Una combinación de factores históricos como el posible efecto fundador durante la colonización de la península ibérica por parte del lince boreal hace 11 mil años, junto con una actividad humana cada vez más intensa durante el Holoceno en la cornisa cantábrica podrían haber llevado a un empobrecimiento genético de la población que habitaba la península ibérica que precipitase su extinción.

A combination of historical factors, such as the colonization of the Iberian Peninsula by the Eurasian lynx 11 kya by relatively few individuals, together with intensified human impacts during the Holocene in the Cantabrian strip could have led to a genetic impoverishment of the Eurasian lynx population in the north of Iberian Peninsula and precipitated its extinction.

7. La población de lince ibérico que habitó la península hace 4 a 2 mil años es genéticamente homogénea y próxima a la población de Andújar. Esta población tiene la diversidad más baja reportada hasta la fecha para ninguna especie de mamífero, incluidas las poblaciones contemporáneas de lince ibérico.

The ancient Iberian lynx population inhabiting the Iberian Peninsula 4-2 kya was relatively homogeneous genetically and shows a closer relationship to the current Andújar than to current Doñana population. This population harbored the lowest diversity ever reported for a mammal, being lower than the already low diversity of the contemporary Iberian lynx populations.

8. El aumento en diversidad genética de lince ibérico a lo largo del tiempo es concomitante con un evento de flujo génico detectado entre las poblaciones contemporáneas de lince boreal e ibérico. El intercambio de alelos se produjo muy probablemente entre lince ibérico y una población mezcla de distintos linajes de boreal y ancestral a las poblaciones contemporáneas.

The recent increase in genetic diversity in the Iberian lynx species in the last two millennia was concomitant with an introgression event from Eurasian lynx. The source of this introgression was, very likely, a population inhabiting Europe, which was the result of admixture between different Eurasian lynx mitogenomic lineages and the ancestor of contemporary populations.

9. La señal de introgresión aumenta a lo largo del tiempo, de modo que la muestra más antigua de 4 mil años, presenta menos alelos compartidos con boreal que las dos muestras de aproximadamente hace 2 mil años, y esas a su vez menos que las contemporáneas. Además la hibridación fue anterior a la divergencia de las dos poblaciones remanentes de lince ibérico, Andújar y Doñana, ya que ambas muestras niveles similares de introgresión.

The signal of introgression from Eurasian lynx increases throughout time, with the oldest Iberian sample (4.2 kya) showing less shared alleles with Eurasian lynx than more recent, but also ancient, samples, and these samples showing less shared alleles than contemporary ones. Also, this admixture process was prior to the divergence of the two Iberian lynx remnant populations, Andújar and Doñana, as both show similar levels of introgression.

10. Poblaciones que han pasado por un cuello de botella de lince ibérico y boreal, concretamente Doñana y Noruega, presentan en zonas del genoma sometidas a selección y en el cromosoma X una acumulación de alelos a baja frecuencia que acaban traduciéndose, incluso, en una θW más alta que en poblaciones no sometidas a cuello de botella.

Bottlenecked populations from Iberian and Eurasian lynx, particularly Doñana and Norway, show a large accumulation of low frequency alleles in the more selectively constrained features and the X chromosome, resulting even in higher θW diversity than in closely-related non-bottlenecked populations.

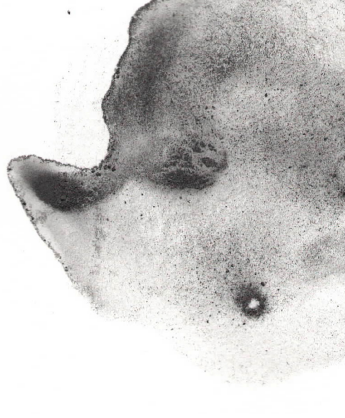
11. La acumulación de variación en las poblaciones de lince boreal e ibérico que han sufrido un cuello de botella está concentrada en zonas de alta mutación. Además encontramos evidencia parcial de relajación de la selección purificadora que afecta no solo a la estructura y función de las proteínas, sino también a la expresión génica.

The accumulation of low frequency alleles in bottleneck populations of Iberian and Eurasian lynx seems to concentrate in genomic regions with higher mutational input. Also, we found partial evidence of relaxation of purifying selection, which could affect not only protein structure and function but also the regulation of gene expression.

12. En resumen, el uso de datos genómicos poblacionales y cuando ha sido posible genomas antiguos ha demostrado ser una herramienta útil para entender la historia evolutiva de las especies y sus linajes, así como para revelar las consecuencias de los cuellos de botella recientes a lo largo del genoma.

In summary, population whole genome data, including when possible ancient genomes, has proven to be a useful tool for unraveling the evolutionary and demographic history of the species, as well as for assessing the genomic and fitness consequences of recent bottlenecks.

PUBLICACIONES



CAPÍTULO 1

Lucena-Perez, M., Marmesat, E., Kleinman-Ruiz, D., Martínez-Cruz, B., Węcek, K., Saveljev, A., ..., Godoy, J.A. (2020). Genomic patterns in the widespread Eurasian lynx shaped by Late Quaternary climatic fluctuations and anthropogenic impacts. Molecular Ecology, 29(4), 812-828. doi: 10.1111/mec.15366.

CAPÍTULO 2

Lucena-Pérez, M., Bazzicalupo, E., Paijmans, J., Kleinman, D., Dalen, L., Hofreiter, M., ... , José A. Godoy. Ancient genome provides insights into the history of Eurasian lynx in Iberia and Western Europe (*in prep*).

CAPÍTULO 3

Lucena-Pérez, M., Paijmans, J., Nocete, F., Nadal, J., Detry, C., Dalen, L., ..., Godoy, J.A. Genetic restoration after interspecific introgression in an endangered felid (*in prep*).

CAPÍTULO 4

Lucena-Pérez, M., Marmesat, E., Kleinman-Ruiz, D., Saveljev, A., Schmidt, K., Godoy, J.A. Bottlenecked-associated changes in the genomic landscape of genetic diversity in wild lynx populations (*submitted*).

OTRAS PUBLICACIONES

En este apartado se incluyen publicaciones que no forman parte de esta tesis, pero en las que la doctoranda ha trabajado durante el desarrollo de la misma, y por tanto, es autora. Sólo se incluyen trabajos de investigación originales en revisión o publicados.

Casas-Marce, M., Marmesat, E., Soriano, L., Martínez-Cruz, B., Lucena-Perez, M., Nocete, F., … Godoy, J. A. (2017). Spatiotemporal Dynamics of Genetic Variation in the Iberian Lynx along Its Path to Extinction Reconstructed with Ancient DNA. Molecular Biology and Evolution, 34(11), 2893-2907. doi: 10.1093/molbev/msx222

Centeno-Cuadros, A., Román, J., Sánchez-Recuero, A., Lucena-Pérez, M., Delibes, M., Godoy, J.A. Mating system, breeding success and pup mortality of a habitat specialist rodent: a field and molecular-based approach. Journal of Mammalian Evolution (*under review*)

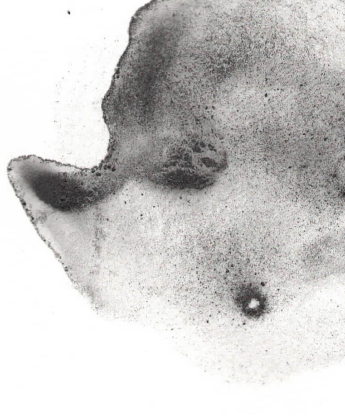
Kleinman-Ruiz, D., Martínez-Cruz, B., Soriano, L., Lucena-Perez, M., Cruz, F., Villanueva, B., … Godoy, J. A. (2017). Novel efficient genome-wide SNP panels for the conservation of the highly endangered Iberian lynx. BMC Genomics, 18(1), 556. doi: 10.1186/s12864-017-3946-5

Kleinman-Ruiz, D., Lucena-Perez, M., Villanueva, B., Fernández, J., Saveljev, A., Ratkiewicz, M., ... Godoy, J. A. The Iberian lynx unloaded: a comparative analysis of genomic load and purging. Molecular Biology and Evolution (*under review*)

Lucena-Perez, M., Soriano, L., López-Bao, J. V., Marmesat, E., Fernández, L., Palomares, F., & Godoy, J. A. (2018). Reproductive biology and genealogy in the endangered Iberian lynx: Implications for conservation. Mammalian Biology, 89, 7-13. doi: 10.1016/j.mambio.2017.11.006

Palomares, F., Lucena-Pérez, M., López-Bao, J. V., & Godoy, J. A. (2017). Territoriality ensures paternity in a solitary carnivore mammal. Scientific Reports. doi: 10.1038/s41598-017-04820-4

REFERENCIAS



Abascal, F., Corvelo, A., Cruz, F., Villanueva-Cañas, J. L., Vlasova, A., Marcet-Houben, M., … Godoy, J. A. (2016). Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. Genome Biology, 17(1), 251. doi: 10.1186/s13059-016-1090-1

Adamec, M., Álvares, F., Anders, O., Andrén, H., Balciauskas, L., Balys, V., … Zlatanova, D. (2012). Status, management and distribution of large carnivores – bear, lynx, wolf & wolverine – in Europe. European Commission, (December), 1-200.

Allendorf, F., Luikart, G., & Aitken, S. (2013). Conservation and the genetics of populations (2nd ed.). Wiley-Blackwell.

Altuna, J. (1980). Hallazgo de un lince nórdico (Lynx lynx L. Mammalia) en la sima de Pagolusieta, Gorbea (Vizcaya). Munibe Antropologia - Arkeologia, (32), 317-324.

Anijalg, P., Ho, S. Y. W., Davison, J., Keis, M., Tammeleht, E., Bobowik, K., … Saarma, U. (2018). Large-scale migrations of brown bears in Eurasia and to North America during the Late Pleistocene. Journal of Biogeography, 45(2), 394-405. doi: 10.1111/jbi.13126

Arbiza, L., Gottipati, S., Siepel, A., & Keinan, A. (2014). Contrasting X-linked and autosomal diversity across 14 human populations. American Journal of Human Genetics, 94(6), 827-844. doi: 10.1016/j.ajhg.2014.04.011

Auguie, B. (2017). gridExtra: Miscellaneous Functions for «Grid» Graphics.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., … Saunders, N. C. (1987). Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annual Review of Ecology and Systematics, 18(1), 489-522. doi: 10.1146/annurev.es.18.110187.002421

Axnér, E., Uhlhorn, H., Ågren, E., & Mörner, T. (2009). Reproductive maturation in the male eurasian Lynx (Lynx lynx): A study on 55 reproductive organs collected from carcasses during 2002-2005. Reproduction in Domestic Animals, 44(3), 467-473. doi: 10.1111/j.1439-0531.2008.01130.x

Balick, D. J., Do, R., Cassa, C. A., Reich, D., & Sunyaev, S. R. (2015). Dominance of Deleterious Alleles Controls the Response to a Population Bottleneck. PLoS Genetics, 11(8), 1-23. doi: 10.1371/journal.pgen.1005436

Ballenghien, M., Faivre, N., & Galtier, N. (2017). Patterns of cross-contamination in a multispecies population genomic project: detection, quantification, impact, and solutions. BMC Biology, 15(1), 25. doi: 10.1186/s12915-017-0366-6

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution, 16(1), 37-48. doi: 10.1093/oxfordjournals.molbev.a026036

Barbato, M., Orozco-terWengel, P., Tapio, M., & Bruford, M. W. (2015). SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. Frontiers in Genetics, 6(MAR), 1-6. doi: 10.3389/fgene.2015.00109

Barlow, A., Cahill, J. A., Hartmann, S., Theunert, C., Xenikoudakis, G., Fortes, G. G., … Hofreiter, M. (2018). Partial genomic survival of cave bears in living brown bears. Nature Ecology and Evolution, 2(10), 1563-1570. doi: 10.1038/s41559-018-0654-8

Barlow, A., Hartmann, S., Gonzalez, J., Hofreiter, M., & Paijmans, J. L. A. (2020). Consensify: A Method for Generating Pseudohaploid Genome Sequences from Palaeogenomic Datasets with Reduced Error Rates. Genes, 11(1), 50. doi: 10.3390/genes11010050

Becher, H., Jackson, B. C., & Charlesworth, B. (2020). Patterns of Genetic Variability in Genomic Regions with Low Rates of Recombination. Current Biology, 30(1), 94-100.e3. doi: 10.1016/j.cub.2019.10.047

Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using Genomic Data to Infer Historic Population Dynamics of Nonmodel Organisms. Annual Review of Ecology, Evolution, and Systematics, 49(1), 433-456. doi: 10.1146/annurev-ecolsys-110617-062431

Benazzo, A., Trucchi, E., Cahill, J. A., Maisano Delser, P., Mona, S., Fumagalli, M., … Bertorelle, G. (2017). Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers. Proceedings of the National Academy of Sciences, 114(45), E9589-E9597. doi: 10.1073/pnas.1707279114

Bieniek, M., Wolsan, M., & Okarma, H. (1998). Historical biogeography of the lynx in Poland. Acta Zoologica Cracoviensia, 41, 143-167.

Bilton, D. T., Mirol, P. M., Mascheretti, S., Fredga, K., Zima, J., & Searle, J. B. (1998). Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265(1402), 1219-1226. doi: 10.1098/rspb.1998.0423

Bivand, R., Keitt, T., & Rowlingson, B. (2017). rgdal: Bindings for the «Geospatial» Data Abstraction Library.

Boitani, L., Alvarez, F., Anders, O., Andren, H., Avanzinelli, E., Balys, V., … Zlatanova, D. (2015). Key actions for Large Carnivore populations in Europe. Report to DG Environment, European Commission, Bruxelles. Contract no. 07.0307/2013/654446/SER/B3. Rome, Italy, Italy.

Borg, I., & Groenen, P. (1997). Modern multidimensional scaling: Theory and applications. New York: Springer.

Boscaini, A., Alba, D. M., Beltrán, J. F., Moyà-Solà, S., & Madurell-Malapeira, J. (2016). Latest Early Pleistocene remains of Lynx pardinus (Carnivora, Felidae) from the Iberian Peninsula: Taxonomy and evolutionary implications. Quaternary Science Reviews, 143, 96-106. doi: 10.1016/j.quascirev.2016.05.015

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Computational Biology, 10(4), e1003537. doi: 10.1371/journal.pcbi.1003537

Braje, T. J., & Erlandson, J. M. (2013). Human acceleration of animal and plant extinctions: A Late Pleistocene, Holocene, and Anthropocene continuum. Anthropocene, 4, 14-23. doi: 10.1016/j.ancene.2013.08.003

Brandvain, Y., & Wright, S. I. (2016). The Limits of Natural Selection in a Nonequilibrium World. Trends in Genetics, 32(4), 201-210. doi: 10.1016/j.tig.2016.01.004

Brass, E. (1911). Aus dem Reiche der Pelze (I. & L. GmbH, Ed.). Berlin, Germany.

Breitenmoser, U., Breitenmoser-Würsten, C., Lanz, T., von Arx, M., Antonevich, A., Bao, W., & Avgan, B. (2015). Lynx lynx (errata version published in 2017). The IUCN Red List of Threatened Species.

Breitenmoser, Urs. (1998). Large predators in the Alps: The fall and rise of man's competitors. Biological Conservation, 83(3), 279-289. doi: 10.1016/S0006-3207(97)00084-0

Brunhoff, C., Galbreath, K. E., Fedorov, V. B., Cook, J. A., & Jaarola, M. (2003). Holarctic phylogeography of the root vole (Microtus oeconomus): implications for late Quaternary biogeography of high latitudes. Molecular Ecology, 12(4), 957-968. doi: 10.1046/j.1365-294X.2003.01796.x

Buckley, R. M., Davis, B. W., Brashear, W. A., Farias, F. H. G., Kuroki, K., Graves, T., … Warren, W. C. (2020). A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. PLOS Genetics, 16(10), e1008926. doi: 10.1371/journal.pgen.1008926

Cahill, J. A., Soares, A. E. R., Green, R. E., & Shapiro, B. (2016). Inferring species divergence times using pairwise sequential Markovian coalescent modelling and low-coverage genomic data. Philosophical Transactions of the Royal Society B: Biological Sciences, 371(1699), 20150138. doi: 10.1098/rstb.2015.0138

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. BMC bioinformatics, 10, 421. doi: 10.1186/1471-2105-10-421

Canty, A., & Ripley, B. D. (2017). boot: Bootstrap R (S-Plus) Functions.

Carrión, J. S., Fernández, S., González-Sampériz, P., Gil-Romera, G., Badal, E., Carrión-Marco, Y., … Burjachs, F. (2010). Expected trends and surprises in the Lateglacial and Holocene vegetation history of the Iberian Peninsula and Balearic Islands. Review of Palaeobotany and Palynology, 162(3), 458-475. doi: 10.1016/j.revpalbo.2009.12.007

Casas-Marce, M, Soriano, L., López-Bao, J. V, & Godoy, J. A. (2013). Genetics at the verge of extinction: insights from the Iberian lynx. Molecular Ecology, 22(22), 5503-5515. doi: 10.1111/mec.12498

Casas-Marce, Mireia, Marmesat, E., Soriano, L., Martínez-Cruz, B., Lucena-Perez, M., Nocete, F., … Godoy, J. A. (2017). Spatiotemporal Dynamics of Genetic Variation in the Iberian Lynx along Its Path to Extinction Reconstructed with Ancient DNA. Molecular Biology and Evolution, 34(11), 2893-2907. doi: 10.1093/molbev/msx222

Casas-Marce, Mireia, Soriano, L., López-Bao, J. V., & Godoy, J. A. (2013). Genetics at the verge of extinction: insights from the Iberian lynx. Molecular ecology, 22(22), 5503-5515. doi: 10.1111/mec.12498

Cassoli, P. F., & Tagliacozzo, A. (1994). I macromammiferi dei livelli tardopleistoceni delle Arene Candide (Savona, Italia): considerazioni paleontologiche e archeologiche. Quaternaria Nova, 4, 101-262.

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience, 4(1), 7. doi: 10.1186/s13742-015-0047-8

Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. Genetics, 134(4), 1289-1303.

Chen, L., Qiu, Q., Jiang, Y., Wang, K., Lin, Z., Li, Z., … Wang, W. (2019). Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science, 364(6446), eaav6202. doi: 10.1126/science.aav6202

Chen, L., Qiu, Q., Jiang, Y., Wang, K., Lin, Z., Li, Z., … Wang, W. (2019). Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science, 364(6446). doi: 10.1126/science.aav6202

Chikhi, L., Rodríguez, W., Grusea, S., Santos, P., Boitard, S., & Mazet, O. (2018). The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: insights into demographic inference and model choice. Heredity, 120(1), 13-24. doi: 10.1038/s41437-017-0005-6

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., … McCabe, A. M. (2009). The Last Glacial Maximum. Science, 325(5941), 710-714. doi: 10.1126/science.1172873

Clavero, M., & Delibes, M. (2013). Using historical accounts to set conservation baselines: The case of Lynx species in Spain. Biodiversity and Conservation, 22(8), 1691-1702. doi: 10.1007/s10531-013-0506-4

Corcoran, P., Gossmann, T. I., Barton, H. J., Great Tit HapMap Consortium, Slate, J., & Zeng, K. (2017). Determinants of the Efficacy of Natural Selection on Coding and Noncoding Variability in Two Passerine Species. Genome biology and evolution, 9(11), 2987-3007. doi: 10.1093/gbe/evx213

Creighton, H. B., & McClintock, B. (1931). A Correlation of Cytological and Genetical Crossing-Over in Zea Mays. Proceedings of the National Academy of Sciences, 17(8), 492-497. doi: 10.1073/pnas.17.8.492

Cubas, M., Altuna, J., Álvarez-Fernández, E., Armendariz, A., Fano, M. Á., López-Dóriga, I. L., … Arias, P. (2016). Re-evaluating the Neolithic: The Impact and the Consolidation of Farming Practices in the Cantabrian Region (Northern Spain). Journal of World Prehistory, 29(1), 79-116. doi: 10.1007/s10963-016-9091-2

Cubas, M., Lucquin, A., Robson, H. K., Colonese, A. C., Arias, P., Aubry, B., … Craig, O. E. (2020). Latitudinal gradient in dairy production with the introduction of farming in Atlantic Europe. Nature Communications, 11(1), 2036. doi: 10.1038/s41467-020-15907-4

Cutter, A. D., & Payseur, B. A. (2013). Genomic signatures of selection at linked sites: Unifying the disparity among species. Nature Reviews Genetics, 14(4), 262-274. doi: 10.1007/s10450-013-9579-3

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., … Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences of the United States of America, 110(39), 15758-15763. doi: 10.1073/pnas.1314445110

DePristo, M., Banks, E., Poplin, R., Garimella, K., Maguire, J., Hartl, C., … Daly, M. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics, 43(5), 491-498. doi: citeulike-article-id:9134853\rdoi: 10.1038/ng.806

Díez-del-Molino, D., Sánchez-Barreiro, F., Barnes, I., Gilbert, M. T. P., & Dalén, L. (2017). Quantifying Temporal Genomic Erosion in Endangered Species. Trends in Ecology and Evolution, 33(3), 176-185. doi: 10.1016/j.tree.2017.12.002

Dimitrieva, S., & Bucher, P. (2013). UCNEbase—a database of ultraconserved non-coding elements and genomic regulatory blocks. Nucleic Acids Research, 41(D1), D101-D109. doi: 10.1093/nar/gks1092

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., & Collen, B. (2014). Defaunation in the Anthropocene. Science, 345(6195), 401-406. doi: 10.1126/science.1251817

Do, R., Balick, D., Li, H., Adzhubei, I., Sunyaev, S., & Reich, D. (2015). No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. Nature Genetics, 47(2), 126-131. doi: 10.1038/ng.3186

Dowle, M., & Srinivasan, A. (2017). data.table: Extension of `data.frame`.

Dobzhansky, Theodosius (1937). Genetics and the Origin of Species. Columbia University Biological Series. New York: Columbia University Press.

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient admixture between closely related populations. Molecular Biology and Evolution, 28(8), 2239-2252. doi: 10.1093/molbev/msr048

Duret, L., & Arndt, P. F. (2008). The Impact of Recombination on Nucleotide Substitutions in the Human Genome. PLoS Genetics, 4(5), e1000071. doi: 10.1371/journal.pgen.1000071

Duret, L., & Galtier, N. (2009). Biased Gene Conversion and the Evolution of Mammalian Genomic Landscapes. Annual Review of Genomics and Human Genetics, 10(1), 285-311. doi: 10.1146/annurev-genom-082908-150001

Dutoit, L., Burri, R., Nater, A., Mugal, C. F., & Ellegren, H. (2017). Genomic distribution and estimation of nucleotide diversity in natural populations: perspectives from the collared flycatcher (Ficedula albicollis) genome. Molecular Ecology Resources, 17(4), 586-597. doi: 10.1111/1755-0998.12602

Edwards, S. M. (2017). lemon: Freshing Up your «ggplot2» Plots.

Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. Trends in Ecology and Evolution, 29(1), 51-63. doi: 10.1016/j.tree.2013.09.008

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. Nature Reviews Genetics, 17(7), 422-433. doi: 10.1038/nrg.2016.58

Endler, J. A. (1977). Geographic Variation, Speciation and Clines. Princenton, N-J.: Princeton University Press.

Endler, J. A. (1986). Natural selection in the wild. Princenton, N.J.: Princeton University Press.

Ersmark, E., Baryshnikov, G., Higham, T., Argant, A., Castaños, P., Döppes, D., … Dalén, L. (2019). Genetic turnovers and northern survival during the last glacial maximum in European brown bears. Ecology and Evolution, 9(10), 5891-5905. doi: 10.1002/ece3.5172

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology, 14(8), 2611-2620. doi: 10.1111/j.1365-294X.2005.02553.x

Fages, A., Hanghøj, K., Khan, N., Gaunitz, C., Seguin-Orlando, A., Leonardi, M., … Orlando, L. (2019). Tracking Five Millennia of Horse Management with Extensive Ancient Genome Time Series. Cell, 177(6), 1419-1435.e31. doi: 10.1016/j.cell.2019.03.049

Felsenstein, J. (1974). The evolutionary advantage of recombination. Genetics, 78(2), 737-756.

Feng, S., Fang, Q., Barnett, R., Li, C., Han, S., Kuhlwilm, M., … Zhang, G. (2019). The Genomic Footprints of the Fall and Recovery of the Crested Ibis. Current Biology, 29(2), 340-349.e7. doi: 10.1016/j.cub.2018.12.008

Ferreras, P., Rodríguez, A., Palomares, F., & Delibes, M. (2010). Iberian lynx: the uncertain future of a critically endangered cat. En D. Macdonald & A. J. Loveridge (Eds.), Biology and conservation of wild felids (pp. 507-520). New York: Oxford University Press.

Förster, D. W., Bull, J. K., Lenz, D., Autenrieth, M., Paijmans, J. L. A., Kraus, R. H. S., … Fickel, J. (2018). Targeted resequencing of coding DNA sequences for SNP discovery in nonmodel species. Molecular Ecology Resources, 18(6), 1356-1373. doi: 10.1111/1755-0998.12924

Frankham, R. (2015). Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. Molecular Ecology, 24(11), 2610-2618. doi: 10.1111/mec.13139

Frankham, R., Ballou, J. D., Briscoe, D. A., & McInnes, K. H. (2002). Introduction to Conservation Genetics (2nd ed.). doi: 10.1017/CBO9780511808999

Frankham, R., Ballou, Jo. D., Eldridge, M. D. B., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the Probability of Outbreeding Depression. Conservation Biology, 25(3), 465-475. doi: 10.1111/j.1523-1739.2011.01662.x

Fisher, R. A. (1930). The genetical theory of natural selection. Oxford: Oxford Univesity Press.

Fritts, S. H., & Sealander, J. A. (1978). Reproductive Biology and Population Characteristics of Bobcats (Lynx rufus) in Arkansas. Journal of Mammalogy, 59(2), 347-353. doi: 10.2307/1379919

Fulton, T. L. (2012). Setting Up an Ancient DNA Laboratory. En B. Shapiro & M. Hofreiter (Eds.), Ancient DNA: Methods and Protocols (pp. 1-11). New York: Springer.

Fumagalli, M. (2013). Assessing the effect of sequencing depth and sample size in population genetics inferences. PLoS ONE, 8(11), e79667. doi: 10.1371/journal.pone.0079667

Fumagalli, M., Vieira, F. G., Korneliussen, T. S., Linderoth, T., Huerta-Sánchez, E., Albrechtsen, A., & Nielsen, R. (2013). Quantifying population genetic differentiation from next-generation sequencing data. Genetics, 195(3), 979-992. doi: 10.1534/genetics.113.154740

Fumagalli, M., Vieira, F. G., Linderoth, T., & Nielsen, R. (2014). ngsTools: methods for population genetics analyses from next-generation sequencing data. Bioinformatics (Oxford, England), 30(10), 1486-1487. doi: 10.1093/bioinformatics/btu041

Galtier, N., Piganeau, G., Mouchiroud, D., & Duret, L. (2001). GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. Genetics, 159(2), 907-911.

Gansauge, M.-T., & Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nature protocols, 8(4), 737-748. doi: 10.1038/nprot.2013.038

Garcia-Dorado, A. (2012). Understanding and Predicting the Fitness Decline of Shrunk Populations: Inbreeding, Purging, Mutation, and Standard Selection. Genetics, 190(4), 1461-1476. doi: 10.1534/genetics.111.135541

García-Ruiz, J. M., Tomás-Faci, G., Diarte-Blasco, P., Montes, L., Domingo, R., Sebastián, M., … Beguería, S. (2020). Transhumance and long-term deforestation in the subalpine belt of the central Spanish Pyrenees: An interdisciplinary approach. CATENA, 195, 104744. doi: 10.1016/j.catena.2020.104744

García-Vázquez, A., Pinto Llona, A. C., & Grandal-d'Anglade, A. (2019). Post-glacial colonization of Western Europe brown bears from a cryptic Atlantic refugium out of the Iberian Peninsula. Historical Biology, 31(5), 618-630. doi: 10.1080/08912963.2017.1384473

Garnier, S. (2018). viridis: Default Color Maps from «matplotlib».

Garrison, E. P., & Marth, G. T. (2012). Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907 [q-bio.GN].

Garza, J. C., & Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. Molecular Ecology, 10(2), 305-318. doi: 10.1046/j.1365-294x.2001.01190.x

Gaunitz, C., Fages, A., Hanghøj, K., Albrechtsen, A., Khan, N., Schubert, M., … Orlando, L. (2018). Ancient genomes revisit the ancestry of domestic and Przewalski's horses. Science, 360(6384), 111-114. doi: 10.1126/science.aao3297

Gignoux, C. R., Henn, B. M., & Mountain, J. L. (2011). Rapid, global demographic expansions after the origins of agriculture. Proceedings of the National Academy of Sciences of the United States of America, 108(15), 6044-6049. doi: 10.1073/pnas.0914274108

Gonzalez-Perez, A., Sabarinathan, R., & Lopez-Bigas, N. (2019). Local Determinants of the Mutational Landscape of the Human Genome. Cell. doi: 10.1016/j.cell.2019.02.051

Gordo, I., & Dionisio, F. (2005). Nonequilibrium model for estimating parameters of deleterious mutations. Physical Review E, 71(3), 031907. doi: 10.1103/PhysRevE.71.031907

Göritz, F., Dehnhard, M., Hildebrandt, T. B., Naidenko, S. V., Vargas, A., Martinez, F., … Jewgenow, K. (2009). Non cat-like ovarian cycle in the Eurasian and the Iberian lynx - ultrasonographical and endocrinological analysis. Reproduction in domestic animals = Zuchthygiene, 44 Suppl 2(SUPPL. 2), 87-91. doi: 10.1111/j.1439-0531.2009.01380.x

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., … Pääbo, S. (2010). A draft sequence of the neandertal genome. Science, 328(5979), 710-722. doi: 10.1126/science.1188021

Gretzinger, J., Molak, M., Reiter, E., Pfrengle, S., Urban, C., Neukamm, J., … Schuenemann, V. J. (2019). Large-scale mitogenomic analysis of the phylogeography of the Late Pleistocene cave bear. Scientific Reports, 9(1), 1-11. doi: 10.1038/s41598-019-47073-z

Groves, T., & Godlee, F. (2012). Open science and reproducible research. BMJ, 344(jun26 1), e4383-e4383. doi: 10.1136/bmj.e4383

Gugolz, D., Bernasconi, M. V., Breitenmoser-Würsten, C., & Wandeler, P. (2008). Historical DNA reveals the phylogenetic position of the extinct Alpine lynx. Journal of Zoology, 275(2), 201-208. doi: 10.1111/j.1469-7998.2008.00428.x

Günther, T., & Nettelblad, C. (2019). The presence and impact of reference bias on population genomic studies of prehistoric human populations. PLOS Genetics, 15(7), e1008302. doi: 10.1371/journal.pgen.1008302

Hall, A. B., Qi, Y., Timoshevskiy, V., Sharakhova, M. V, Sharakhov, I. V, & Tu, Z. (2013). Six novel Y chromosome genes in Anopheles mosquitoes discovered by independently sequencing males and females. BMC genomics, 14, 273. doi: 10.1186/1471-2164-14-273

Halldorsson, B. V. ., Palsson, G., Stefansson, O. A., Jonsson, H., Hardarson, M. T., Eggertsson, H. P., … Stefansson, K. (2019). Characterizing mutagenic effects of recombination through a sequence-level genetic map. Science, 363(6425), eaau1043. doi: 10.1126/science.aau1043

Hanghøj, K., & Orlando, L. (2019). Ancient Epigenomics BT - Paleogenomics: Genome-Scale Analysis of Ancient DNA (C. Lindqvist & O. P. Rajora, Eds.). doi: 10.1007/13836_2018_18

Hansen, A. J., Willerslev, E., Wiuf, C., Mourier, T., & Arctander, P. (2001). Statistical Evidence for Miscoding Lesions in Ancient DNA Templates. Molecular Biology and Evolution, 18(2), 262-265. doi: 10.1093/oxfordjournals.molbev.a003800

Harrell Jr, F. E., Dupont, with contributions from C., & many others. (2019). Hmisc: Harrell Miscellaneous.

Hedrick, P. W. (1994). Purging inbreeding depression and the probability of extinction: full-sib mating. Heredity, 73(4), 363-372. doi: 10.1038/hdy.1994.183

Hedrick, P. W., & Fredrickson, R. (2010). Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. Conservation Genetics, 11(2), 615-626. doi: 10.1007/s10592-009-9999-5

Hedrick, P. W., & Garcia-Dorado, A. (2016). Understanding Inbreeding Depression, Purging, and Genetic Rescue. Trends in Ecology and Evolution, 31(12), 940-952. doi: 10.1016/j.tree.2016.09.005

Hell, P., & Slamečka, J. (1996). Current status of the lynx (Lynx lynx) in Slovakia. Acta Sc. Nat. Brno, 30(3), 64-78.

Hellborg, L., & Ellegren, H. (2004). Low Levels of Nucleotide Diversity in Mammalian Y Chromosomes. Molecular Biology and Evolution, 21(1), 158-163. doi: 10.1093/molbev/msh008

Hellborg, L., Walker, C. W., Rueness, E. K., Stacy, J. E., Kojola, I., Valdmann, H., … Ellegren, H. (2002). Differentiation and levels of genetic variation in northern European lynx (Lynx lynx) populations revealed by microsatellites and mitochondrial DNA analysis. Conservation Genetics, 3(2), 97-111. doi: 10.1023/A:1015217723287

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., … Nielsen, E. E. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources, 11, 123-136. doi: 10.1111/j.1755-0998.2010.02943.x

Henn, B. M., Botigué, L. R., Peischl, S., Dupanloup, I., Lipatov, M., Maples, B. K., … Bustamante, C. D. (2016). Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. Proceedings of the National Academy of Sciences of the United States of America, 113(4), E440-E449. doi: 10.1073/pnas.1510805112

Henry, L., & Wickham, H. (2018). rlang: Functions for Base Types and Core R and «Tidyverse» Features.

Herrel, A., Joly, D., & Danchin, E. (2020). Epigenetics in ecology and evolution. Functional Ecology, 34(2), 381-384. doi: 10.1111/1365-2435.13494

Hetherington, D. A., Lord, T. C., & Jacobi, R. M. (2006). New evidence for the occurence of Eurasian lynx (Lynx lynx) in medieval Britain. Journal of Quaternary Science, 21, 3-8.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. Nature, 405(6789), 907-913. doi: 10.1038/35016000

Hill, W. G. (1981). Estimation of effective population size from data on linkage disequilibrium. Genetical Research, 38(3), 209-216. doi: 10.1017/S0016672300020553

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. Genetical research, 8(3), 269-294. doi: 10.1017/S001667230800949X

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., & Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proceedings of the National Academy of Sciences of the United States of America, 106(23), 9362-9367. doi: 10.1073/pnas.0903103106

Hindrikson, M., Remm, J., Pilot, M., Godinho, R., Stronen, A. V., Baltrūnaité, L., … Saarma, U. (2017). Wolf population genetics in Europe: a systematic review, meta-analysis and suggestions for conservation and management. Biological Reviews, 92(3), 1601-1629. doi: 10.1111/brv.12298

Holmala, K., Herrero, A., Kopatz, A., Schregel, J., Eiken, H. G., & Hagen, S. B. (2018). Genetic evidence of female kin clusters in a continuous population of a solitary carnivore, the Eurasian lynx. Ecology and Evolution, 8(22), 10964-10975. doi: 10.1002/ece3.4562

Horn, S., Prost, S., Stiller, M., Makowiecki, D., Kuznetsova, T., Benecke, N., … Hofreiter, M. (2014). Ancient mitochondrial DNA and the genetic history of Eurasian beaver (Castor fiber) in Europe. Molecular Ecology, 23(7), 1717-1729. doi: 10.1111/mec.12691

Jambrina-Enríquez, M., Rico, M., Moreno, A., Leira, M., Bernárdez, P., Prego, R., … Valero-Garcés, B. L. (2014). Timing of deglaciation and postglacial environmental dynamics in NW Iberia: the Sanabria Lake record. Quaternary Science Reviews, 94, 136-158. doi: 10.1016/j.quascirev.2014.04.018

Jędrzejewski, W., Jędrzejewska, B., Okarma, H., Schmidt, K., Bunevich, A. N., & Miłkowski, L. (1996). Population dynamics (1869-1994), demography and home ranges of the lynx in Bialowieza Primeval Forest (Poland) and Belarus. Ecography, 19, 122-138.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P. L. F., & Orlando, L. (2013). mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics, 29(13), 1682-1684. doi: 10.1093/bioinformatics/btt193

Kaal, J., Carrión Marco, Y., Asouti, E., Martín Seijo, M., Martínez Cortizas, A., Costa Casáis, M., & Criado Boado, F. (2011). Long-term deforestation in NW Spain: linking the Holocene fire history to vegetation change and human activities. Quaternary Science Reviews, 30(1-2), 161-175. doi: 10.1016/j.quascirev.2010.10.006

Kahlke, R. D. (1999). The History of the Origin, Evolution and Dispersal of the Late Pleistocene Mammuthus-Coelodonta Faunal Complex in Eurasia (Large Mammals). South Dakota: Mammoth Site of Hot Springs.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics (Oxford, England), 28(12), 1647-1649. doi: 10.1093/bioinformatics/bts199

Kelly, M., Morin, D., & López-González, C. A. (2016). Lynx rufus. The IUCN Red List of Threatened Species 2016, 8235, e.T12521A50655874. doi: https://dx.doi.org/10.2305/IUCN.UK.2016-1.RLTS.T12521A50655874.en

Kim, S. Y., Lohmueller, K. E., Albrechtsen, A., Li, Y., Korneliussen, T., Tian, G., … Nielsen, R. (2011). Estimation of allele frequency and association mapping using next-generation sequencing data. BMC Bioinformatics, 12(1), 231. doi: 10.1186/1471-2105-12-231

Kimura, M. (1962). On the probability of fixation of mutant genes in a population. Genetics, 47(6), 713-719.

Kimura, M. (1968). Evolutionary Rate at the Molecular Level. Nature, 217(5129), 624-626. doi: 10.1038/217624a0

Kimura, M. (1983). The Neutral Theory of Molecular Evolution. Cambridge: University Press Cambridge.

Kindler, P., Guillevic, M., Baumgartner, M., Schwander, J., Landais, A., & Leuenberger, M. (2014). Temperature reconstruction from 10 to 120 kyr b2k from the NGRIP ice core. Climate of the Past, 10(2), 887-902. doi: 10.5194/cp-10-887-2014

Kirkpatrick, M., & Jarne, P. (2000). The effects of a bottleneck on inbreeding depression and the genetic load. American Naturalist, 155(2), 154-167. doi: 10.1086/303312

Kitchener, A. C., Breitenmoser-Würsten, C., Eizirik, E., Gentry, A., Werdelin, L., Wilting, A., … Tobe, S. (2018). A revised taxonomy of the Felidae. The final report of the Cat Classification Task Force of the IUCN/SSC Cat Specialist Group. Cat News Special Issue 11, 80.

Kleinman-Ruiz, D., Soriano, L., Casas-Marce, M., Szychta, C., Sánchez, I., Fernández, J., & Godoy, J. A. (2019). Genetic evaluation of the Iberian lynx ex situ conservation programme. Heredity, 123(5), 647-661. doi: 10.1038/s41437-019-0217-z

Kohli, B. A., Fedorov, V. B., Waltari, E., & Cook, J. A. (2015). Phylogeography of a Holarctic rodent (Myodes rutilus): testing high-latitude biogeographical hypotheses and the dynamics of range shifts. Journal of Biogeography, 42(2), 377-389. doi: 10.1111/jbi.12433

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Molecular ecology resources, 15(5), 1179-1191. doi: 10.1111/1755-0998.12387

Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., & Meyer, M. (2015). Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. BioTechniques, 59(2). doi: 10.2144/000114320

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. BMC bioinformatics, 15(1), 356. doi: 10.1186/s12859-014-0356-4

Korneliussen, T. S., Moltke, I., Albrechtsen, A., & Nielsen, R. (2013). Calculation of Tajima's D and other neutrality test statistics from low depth next-generation sequencing data. BMC Bioinformatics, 14(1), 289. doi: 10.1186/1471-2105-14-289

Kratochvil, J. (1968). Survey of the distribution of populations of the genus lynx in Europe. Acta scientiarum naturalium Academiae scientiarum bohemoslovacae - Brno, 2, 5-12.

Kratochvil, J. (1968a). Survey of the distribution of populations of the genus lynx in Europe. Acta scientiarum naturalium Academiae scientiarum bohemoslovacae - Brno, 2, 5-12.

Kratochvil, J. (1968b). The lynx population in Rumania. Acta Sc. Nat. Brno, 2(5/6), 65-70.

Krojerová-Prokešová, J., Turbaková, B., Jelenčič, M., Bojda, M., Kutal, M., Skrbinšek, T., … Bryja, J. (2019). Genetic constraints of population expansion of the Carpathian lynx at the western edge of its native distribution range in Central Europe. Heredity, 122(6), 785-799. doi: 10.1038/s41437-018-0167-x

Kubala, J., Smolko, P., Zimmermann, F., Rigg, R., Tám, B., Ilko, T., … Breitenmoser, U. (2019). Robust monitoring of the Eurasian lynx Lynx lynx in the Slovak Carpathians reveals lower numbers than officially reported. Oryx, 53(3), 548-556. doi: 10.1017/S003060531700076X

Kumar, S., & Dudley, J. (2007). Bioinformatics software for biologists in the genomics era. Bioinformatics, 23(14), 1713-1717. doi: 10.1093/bioinformatics/btm239

Kurten, B., & Werdelin, L. (1984). The relationships of Lynx shansius Teilhard. Annales Zoologici Fennici, 21, 129-133.

Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Molecular biology and evolution, 34(3), 772-773. doi: 10.1093/molbev/msw260

Leigh, J. W., & Bryant, D. (2015). Popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution, 6(9), 1110-1116. doi: 10.1111/2041-210X.12410

Leonard, J. A. (2008). Ancient DNA applications for wildlife conservation. Molecular Ecology, 17(19), 4186-4196. doi: 10.1111/j.1365-294X.2008.03891.x

Leonardi, M., Librado, P., Der Sarkissian, C., Schubert, M., Alfarhan, A. H., Alquraishi, S. A., … Orlando, L. (2017). Evolutionary patterns and processes: Lessons from ancient DNA. Systematic Biology, 66(1), e1-e29. doi: 10.1093/sysbio/syw059

Lercher, M. J., & Hurst, L. D. (2002). Human SNP variability and mutation rate are higher in regions of high recombination. Trends in Genetics, 18(7), 337-340. doi: 10.1016/S0168-9525(02)02669-0

Lewis, B. W. (2017). threejs: Interactive 3D Scatter Plots, Networks and Globes.

Li, G., Davis, B. W., Eizirik, E., & Murphy, W. J. (2016). Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). Genome Research, 26(1), 1-11. doi: 10.1101/gr.186668.114

Li, G., Figueiró, H. V., Eizirik, E., & Murphy, W. J. (2019). Recombination-Aware Phylogenomics Reveals the Structured Genomic Landscape of Hybridizing Cat Species. Molecular Biology and Evolution, 36(10), 2111-2126. doi: 10.1093/molbev/msz139

Li, G., Hillier, L. W., Grahn, R. A., Zimin, A. V, David, V. A., Menotti-Raymond, M., … Murphy, W. J. (2016). A High-Resolution SNP Array-Based Linkage Map Anchors a New Domestic Cat Draft Genome Assembly and Provides Detailed Patterns of Recombination. G3 (Bethesda, Md.), 6(6), 1607-1616. doi: 10.1534/g3.116.028746

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, 27(21), 2987-2993. doi: 10.1093/bioinformatics/btr509

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v1 [q-bio.GN].

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25(14), 1754-1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25(16), 2078-2079. doi: 10.1093/bioinformatics/btp352

Li, Heng, & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25(14), 1754-1760. doi: 10.1093/bioinformatics/btp324

Li, Heng, & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature, 475(7357), 493-496. doi: 10.1038/nature10231

Li, Heng, Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics, 25(16), 2078-2079. doi: 10.1093/bioinformatics/btp352

Li, Heng. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics, 27(21), 2987-2993. doi: 10.1093/bioinformatics/btr509

Li, Heng. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v1 [q-bio.GN].

Li, S., Li, B., Cheng, C., Xiong, Z., Liu, Q., Lai, J., … Yan, J. (2014). Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. Genome biology, 15(12), 557. doi: 10.1186/s13059-014-0557-1

Librado, P., Der Sarkissian, C., Ermini, L., Schubert, M., Jónsson, H., Albrechtsen, A., … Orlando, L. (2015). Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. Proceedings of the National Academy of Sciences, 112(50), 201513696. doi: 10.1073/pnas.1513696112

Linnell, J. D. C., Breitenmoser, U., Breitenmoser-Wrsten, C., Odden, J., & von Arx, M. (2009). Recovery of Eurasian Lynx in Europe: What Part has Reintroduction Played? En Reintroduction of Top-Order Predators (pp. 72-91). doi: 10.1002/9781444312034.ch4

Linnell, J. D. C., Broseth, H., Odden, J., & Nilsen, E. B. (2010). Sustainably Harvesting a Large Carnivore? Development of Eurasian Lynx Populations in Norway During 160 Years of Shifting Policy. Environmental Management, 45(5), 1142-1154. doi: 10.1007/s00267-010-9455-9

Liu, X., & Fu, Y.-X. (2015). Exploring population size changes using SNP frequency spectra. Nature genetics, 47(5), 555-559. doi: 10.1038/ng.3254

Lohmueller, K. E. (2014). The distribution of deleterious genetic variation in human populations. Current Opinion in Genetics and Development, 29, 139-146. doi: 10.1016/j.gde.2014.09.005

Lorenzen, E. D., Nogués-Bravo, D., Orlando, L., Weinstock, J., Binladen, J., Marske, K. A., … Willerslev, E. (2011). Species-specific responses of Late Quaternary megafauna to climate and humans. Nature, 479(7373), 359-364. doi: 10.1038/nature10574

Lucena-Perez, M., Marmesat, E., Kleinman-Ruiz, D., Martínez-Cruz, B., Węcek, K., Saveljev, A. P., … Godoy, J. A. (2020). Genomic patterns in the widespread Eurasian lynx shaped by Late Quaternary climatic fluctuations and anthropogenic impacts. Molecular Ecology, 29(4), 812-828. doi: 10.1111/mec.15366

Lucena-Perez, M., Soriano, L., López-Bao, J. V., Marmesat, E., Fernández, L., Palomares, F., & Godoy, J. A. (2018). Reproductive biology and genealogy in the endangered Iberian lynx: Implications for conservation. Mammalian Biology, 89, 7-13. doi: 10.1016/j.mambio.2017.11.006

Makino, T., Rubin, C.-J., Carneiro, M., Axelsson, E., Andersson, L., & Webster, M. T. (2018). Elevated Proportions of Deleterious Genetic Variation in Domestic Animals and Plants. Genome Biology and Evolution, 10(1), 276-290. doi: 10.1093/gbe/evy004

Mangiafico, S. (2020). rcompanion: Functions to Support Extension Education Program Evaluation.

Mangul, S., Martin, L. S., Eskin, E., & Blekhman, R. (2019). Improving the usability and archival stability of bioinformatics software. Genome Biology, 20(1), 47. doi: 10.1186/s13059-019-1649-8

Marcovitz, A., Jia, R., & Bejerano, G. (2016). «reverse Genomics» Predicts Function of Human Conserved Noncoding Elements. Molecular Biology and Evolution, 33(5), 1358-1369. doi: 10.1093/molbev/msw001

Marín-Arroyo, A. B. (2013). Human response to Holocene warming on the Cantabrian Coast (northern Spain): an unexpected outcome. Quaternary Science Reviews, 81, 1-11. doi: 10.1016/j.quascirev.2013.09.006

Marsden, C. D., Vecchyo, D. O. Del, O'Brien, D. P., Taylor, J. F., Ramirez, O., Vilà, C., … Lohmueller, K. E. (2016). Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. Proceedings of the National Academy of Sciences of the United States of America, 113(1), 152-157. doi: 10.1073/pnas.1512501113

Martin, A. D., Quinn, K. M., & Park, J. H. (2011). MCMCpack: Markov Chain Monte Carlo in R. Journal of Statistical Software, 42(9), 22.

Matyushkin, E. N., & Vaisfeld, M. A. (2003). The lynx. Regional features of ecology, use and protection. Moscow: Nauka.

Mazet, O., Rodríguez, W., Grusea, S., Boitard, S., & Chikhi, L. (2016). On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? Heredity, 116(4), 362-371. doi: 10.1038/hdy.2015.104

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 20(9), 1297-1303. doi: 10.1101/gr.107524.110

Mellars, P., & French, J. C. (2011). Tenfold Population Increase in Western Europe at the Neandertal-to-Modern Human Transition. Science, 333(6042), 623-627. doi: 10.1126/science.1206930

Melovski, D., von Arx, M., Avukatov, V., Breitenmoser-Würsten, C., Đurović, M., Elezi, R., … Breitenmoser, U. (2020). Using questionnaire surveys and occupancy modelling to identify conservation priorities for the Critically Endangered Balkan lynx Lynx lynx balcanicus. Oryx, 54(5), 706-714. doi: 10.1017/S0030605318000492

Meröndun, J., Murray, D. L., & Shafer, A. B. A. (2019). Genome-scale sampling suggests cryptic epigenetic structuring and insular divergence in Canada lynx. Molecular Ecology, mec.15131. doi: 10.1111/mec.15131

Meyer, M., & Kircher, M. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. Cold Spring Harbor Protocols, 6. doi: 10.1101/pdb.prot5448

Mix, A. (2001). Environmental processes of the ice age: land, oceans, glaciers (EPILOG). Quaternary Science Reviews, 20(4), 627-657. doi: 10.1016/S0277-3791(00)00145-1

Mueller, S. A., Reiners, T. E., Middelhoff, T. L., Anders, O., Kasperkiewicz, A., & Nowak, C. (2020). The rise of a large carnivore population in Central Europe: genetic evaluation of lynx reintroduction in the Harz Mountains. Conservation Genetics. doi: 10.1007/s10592-020-01270-w

Murchison, E. P., Schulz-Trieglaff, O. B., Ning, Z., Alexandrov, L. B., Bauer, M. J., Fu, B., … Stratton, M. R. (2012). Genome Sequencing and Analysis of the Tasmanian Devil and Its Transmissible Cancer. Cell, 148(4), 780-791. doi: 10.1016/j.cell.2011.11.065

Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G., & Ellegren, H. (2015). Temporal Dynamics of Avian Populations during Pleistocene Revealed by Whole-Genome Sequences. Current Biology, 25(10), 1375-1380. doi: 10.1016/j.cub.2015.03.047

Navascués, M., Depaulis, F., & Emerson, B. C. (2010). Combining contemporary and ancient DNA in population genetic and phylogeographical studies. Molecular Ecology Resources, 10(5), 760-772. doi: 10.1111/j.1755-0998.2010.02895.x

Nei, M., & Maruyama, T. (1975). Lewontin-Krakauer test for neutral genes. Genetics.

Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes.

Nielsen, R., Akey, J. M., Jakobsson, M., Pritchard, J. K., Tishkoff, S., & Willerslev, E. (2017). Tracing the peopling of the world through genomics. Nature, 541(7637), 302-310. doi: 10.1038/nature21347

Nilsen, E. B., Brøseth, H., Odden, J., & Linnell, J. D. C. (2010). The cost of maturing early in a solitary carnivore. Oecologia, 164(4), 943-948. doi: 10.1007/s00442-010-1713-2

Nowak, R. M. (2005). Walker's Carnivores of the World. Baltimore: The Johns Hopkins University Press.

Nowell, K., & Jackson, P. (1996). Wild cats: A status survey and conservation action plan. Gland, Switzerland: International Union for the Conservation of Nature and Natural Resources (IUCN).

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2018). vegan: Community Ecology Package.

Oppenheimer, S. (2012). Out-of-Africa, the peopling of continents and islands: tracing uniparental gene trees across the map. Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1590), 770-784. doi: 10.1098/rstb.2011.0306

Ohta, T. (1992). The Nearly Neutral Theory of Molecular Evolution. Annual Review of Ecology and Systematics, 23(1), 263-286. doi: 10.1146/annurev.es.23.110192.001403

Otto, S. P. (2018). Adaptation, speciation and extinction in the Anthropocene. Proceedings of the Royal Society B: Biological Sciences, 285(1891), 20182047. doi: 10.1098/rspb.2018.2047

Paijmans, J. L. A., Barnett, R., Gilbert, M. T. P., Zepeda-Mendoza, M. L., Reumer, J. W. F., de Vos, J., … Hofreiter, M. (2017). Evolutionary History of Saber-Toothed Cats Based on Ancient Mitogenomics. Current Biology, 27(21), 3330-3336.e5. doi: 10.1016/j.cub.2017.09.033

Palkopoulou, E., Lipson, M., Mallick, S., Nielsen, S., Rohland, N., Baleka, S., … Reich, D. (2018). A comprehensive genomic history of extinct and living elephants. Proceedings of the National Academy of Sciences of the United States of America, 115(11), E2566-E2574. doi: 10.1073/pnas.1720554115

Palkopoulou, E., Mallick, S., Skoglund, P., Enk, J., Rohland, N., Li, H., … Dalén, L. (2015). Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. Current Biology, 25(10), 1395-1400. doi: 10.1016/j.cub.2015.04.007

Palomares, F., Godoy, J. A., López-Bao, J. V, Rodríguez, A., Roques, S., Casas-Marce, M., … Delibes, M. (2012). Possible extinction vortex for a population of Iberian lynx on the verge of extirpation. Conservation Biology, 26(4), 689-697. doi: 10.1111/J.1523-1739.2012.01870.X

Palomares, F., Godoy, J. A., López-Bao, J. V., Rodríguez, A., Roques, S., Casas-Marce, M., … Delibes, M. (2012). Possible Extinction Vortex for a Population of Iberian Lynx on the Verge of Extirpation. Conservation Biology, 26(4), 689-697. doi: 10.1111/j.1523-1739.2012.01870.x

Palomares, F., Rodríguez, A., Revilla, E., López-Bao, J. V., & Calzada, J. (2011). Assessment of the conservation efforts to prevent extinction of the Iberian lynx. Conservation biology, 25(1), 4-8. doi: 10.1111/j.1523-1739.2010.01607.x

Paradis, E, Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics, 20(2), 289-290. doi: 10.1093/bioinformatics/btg412

Paradis, Emmanuel. (2010). pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics (Oxford, England), 26(3), 419-420. doi: 10.1093/bioinformatics/btp696

Pelletier, F., & Coltman, D. W. (2018). Will human influences on evolutionary dynamics in the wild pervade the Anthropocene? BMC Biology, 16(1), 1-10. doi: 10.1186/s12915-017-0476-1

Peña-Chocarro, L., Zapata, L., Iriarte, M. J., González Morales, M., & Straus, L. G. (2005). The oldest agriculture in northern Atlantic Spain: New evidence from El Mirón Cave (Ramales de la Victoria, Cantabria). Journal of Archaeological Science, 32(4), 579-587. doi: 10.1016/j.jas.2004.12.001

Pennings, P. S., Kryazhimskiy, S., & Wakeley, J. (2014). Loss and Recovery of Genetic Diversity in Adapting Populations of HIV. PLoS Genetics, 10(1), e1004000. doi: 10.1371/journal.pgen.1004000

Pérez-Díaz, S., López-Sáez, J. A., Núñez de la Fuente, S., & Ruiz-Alonso, M. (2018). Early farmers, megalithic builders and the shaping of the cultural landscapes during the Holocene in Northern Iberian mountains. A palaeoenvironmental perspective. Journal of Archaeological Science: Reports, 18, 463-474. doi: 10.1016/j.jasrep.2018.01.043

Petrovski, S., Wang, Q., Heinzen, E. L., Allen, A. S., & Goldstein, D. B. (2013). Genic Intolerance to Functional Variation and the Interpretation of Personal Genomes. PLoS Genetics, 9(8), e1003709. doi: 10.1371/journal.pgen.1003709

Pfeifer, B., Wittelsbuerger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. Molecular Biology and Evolution, 31, 1929-1936. doi: 10.1093/molbev/msu136

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. PLoS Genetics, 8(11), e1002967. doi: 10.1371/journal.pgen.1002967

Pilot, M., Greco, C., Vonholdt, B. M., Jędrzejewska, B., Randi, E., Jędrzejewski, W., … Wayne, R. K. (2014). Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. Heredity, 112(4), 428-442. doi: 10.1038/hdy.2013.122

Pilot, Magorzata, Branicki, W., Jędrzejewski, W., Goszczyski, J., Jędrzejewska, B., Dykyy, I., … Tsingarska, E. (2010). Phylogeographic history of grey wolves in Europe. BMC Evolutionary Biology, 10(1). doi: 10.1186/1471-2148-10-104

Polychronopoulos, D., King, J. W. D., Nash, A. J., Tan, G., & Lenhard, B. (2017). Conserved non-coding elements: Developmental gene regulation meets genome organization. Nucleic Acids Research, 45(22), 12611-12624. doi: 10.1093/nar/gkx1074

Pool, J. E., & Nielsen, R. (2007). Population size changes reshape genomic patterns of diversity. Evolution, 61(12), 3001-3006. doi: 10.1111/j.1558-5646.2007.00238.x

Pratto, F., Brick, K., Khil, P., Smagulova, F., Petukhova, G. V., & Camerini-Otero, R. D. (2014). Recombination initiation maps of individual human genomes. Science, 346(6211), 1256442-1256442. doi: 10.1126/science.1256442

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics, 26(6), 841-842. doi: 10.1093/bioinformatics/btq033

R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria.

Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. Nature reviews. Genetics, 16(6), 359-371. doi: 10.1038/nrg3936

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Systematic Biology, 67(5), 901-904. doi: 10.1093/sysbio/syy032

Ratkiewicz, M., Matosiuk, M., Kowalczyk, R., Konopiński, M. K., Okarma, H., Ozolins, J., … Schmidt, K. (2012). High levels of population differentiation in Eurasian lynx at the edge of the species' western range in Europe revealed by mitochondrial DNA analyses. Animal Conservation, 15(6), 603-612. doi: 10.1111/j.1469-1795.2012.00556.x

Ratkiewicz, M., Matosiuk, M., Saveljev, A. P., Sidorovich, V., Ozolins, J., Männil, P., … Schmidt, K. (2014). Long-range gene flow and the effects of climatic and ecological factors on genetic structuring in a large, solitary carnivore: The Eurasian lynx. PLoS ONE, 9(12), 1-29. doi: 10.1371/journal.pone.0115160

Reich, D., Thangaraj, K., Patterson, N., Price, A. L., & Singh, L. (2009). Reconstructing Indian population history. Nature, 461(7263), 489-494. doi: 10.1038/nature08365

Resendez, S. D., Bradley, J. R., Xu, D., & Gokcumen, O. (2019). Structural Variants in Ancient Genomes BT - Paleogenomics: Genome-Scale Analysis of Ancient DNA (C. Lindqvist & O. P. Rajora, Eds.). doi: 10.1007/13836_2018_34

Rey-Iglesia, A., García-Vázquez, A., Treadaway, E. C., van der Plicht, J., Baryshnikov, G. F., Szpak, P., … Lorenzen, E. D. (2019). Evolutionary history and palaeoecology of brown bear in North-East Siberia re-examined using ancient DNA and stable isotopes from skeletal remains. Scientific Reports, 9(1), 1-12. doi: 10.1038/s41598-019-40168-7

Rey-Iglesia, A., Grandal-d'Anglade, A., Campos, P. F., & Hansen, A. J. (2017). Mitochondrial DNA of pre-last glacial maximum red deer from NW Spain suggests a more complex phylogeographical history for the species. Ecology and Evolution, 7(24), 10690-10700. doi: 10.1002/ece3.3553

Robinson, D. (2018). broom: Convert Statistical Analysis Objects into Tidy Data Frames.

Robinson, J. A., Ortega-Del Vecchyo, D., Fan, Z., Kim, B. Y., Vonholdt, B. M., Marsden, C. D., … Wayne, R. K. (2016). Genomic Flatlining in the Endangered Island Fox. Current Biology, 26(9), 1183-1189. doi: 10.1016/j.cub.2016.02.062

Rodríguez-Varela, R., García, N., Nores, C., Álvarez-Lao, D., Barnett, R., Arsuaga, J. L., … Valdiosera, C. (2016). Ancient DNA reveals past existence of Eurasian lynx in Spain. Journal of Zoology, 298(2), 94-102. doi: 10.1111/jzo.12289

Rodríguez-Varela, R., Tagliacozzo, A., Ureña, I., García, N., Crégut-Bonnoure, E., Mannino, M. A., … Valdiosera, C. (2015). Ancient DNA evidence of Iberian lynx palaeoendemism. Quaternary Science Reviews, 112, 172-180. doi: 10.1016/j.quascirev.2015.01.009

Rodríguez, A, & Calzada, J. (2015). Lynx pardinus. The IUCN Red List of Threatened Species 2015, e.T12520A50655794. doi: http://dx.doi.org/10.2305/IUCN.UK.2015-2.RLTS.T12520A50655794.en

Rodríguez, Alejandro, & Delibes, M. (1992). Current range and status of the Iberian lynx Felis pardina Temminck, 1824 in Spain. Biological Conservation, 61(3), 189-196. doi: 10.1016/0006-3207(92)91115-9

Rodríguez, Alejandro, & Delibes, M. (2002). Internal structure and patterns of contraction in the geographic range of the Iberian lynx. Ecography, 25(3), 314-328. doi: 10.1034/j.1600-0587.2002.250308.x

Rodríguez, Alejandro, & Delibes, M. (2003). Population fragmentation and extinction in the Iberian lynx. Biological Conservation, 109, 321-331. doi: 10.1016/S0006-3207(02)00158-1

Rodríguez, R., Ramírez, O., Valdiosera, C. E., García, N., Alda, F., Madurell-Malapeira, J., … Dalén, L. (2011). 50,000 years of genetic uniformity in the critically endangered Iberian lynx. Molecular Ecology, 20(18), 3785-3795. doi: 10.1111/j.1365-294X.2011.05231.x

Rogers, R. L., & Slatkin, M. (2017). Excess of genomic defects in a woolly mammoth on Wrangel island. PLoS Genetics, 13(3), 1-16. doi: 10.1371/journal.pgen.1006601

Rohland, N., & Hofreiter, M. (2007). Comparison and optimization of ancient DNA extraction. BioTechniques, 42(3), 343-352. doi: 10.2144/000112383

Rubi, T. L., Knowles, L. L., & Dantzer, B. (2020). Museum epigenomics: Characterizing cytosine methylation in historic museum specimens. Molecular Ecology Resources, 20(5), 1161-1170. doi: 10.1111/1755-0998.13115

Rueness, E. K., Naidenko, S., Trosvik, P., & Stenseth, N. C. (2014). Large-scale genetic structuring of a widely distributed carnivore - The eurasian lynx (Lynx lynx). PLoS ONE, 9(4), 1-11. doi: 10.1371/journal.pone.0093675

Ryder, O. A. (2005). Conservation genomics: applying whole genome studies to species conservation efforts. Cytogenetic and genome research, 108(1-3), 6-15. doi: 10.1159/000080796

Schlolaut, G., Brauer, A., Nakagawa, T., Lamb, H. F., Tyler, J. J., Staff, R. A., … Tarasov, P. E. (2017). Evidence for a bi-partition of the Younger Dryas Stadial in East Asia associated with inversed climate characteristics compared to Europe. Scientific Reports, 7(1), 44983. doi: 10.1038/srep44983

Schmidt, K. (1998). Maternal behaviour and juvenil dispersal in the Eurasian lynx. Acta Theriologica, 43(4), 391-408.

Schmidt, K., Kowalczyk, R., Ozolins, J., Männil, P., & Fickel, J. (2009). Genetic structure of the Eurasian lynx population in north-eastern Poland and the Baltic states. Conservation Genetics, 10(2), 497-501. doi: 10.1007/s10592-008-9795-7

Schmitt, T., & Varga, Z. (2012). Extra-Mediterranean refugia: The rule and not the exception? Frontiers in Zoology, 9, 1-12. doi: 10.1186/1742-9994-9-22

Seguin-Orlando, A., Hoover, C. A., Vasiliev, S. K., Ovodov, N. D., Shapiro, B., Cooper, A., … Orlando, L. (2015). Amplification of TruSeq ancient DNA libraries with AccuPrime Pfx: consequences on nucleotide misincorporation and methylation patterns. STAR: Science & Technology of Archaeological Research, 1(1), 1-9. doi: 10.1179/2054892315Y.0000000005

Sheng, G., Basler, N., Ji, X.-P., Paijmans, J. L. A., Preick, M., Hartmann, S., … Barlow, A. (2019). Palaeogenome Reveals Genetic Contribution of Extinct Giant Panda to Extant Populations. SSRN Electronic Journal, 1-6. doi: 10.2139/ssrn.3316802

Siepel, A. (2019). Challenges in funding and developing genomic software: roots and remedies. Genome Biology, 20(1), 147. doi: 10.1186/s13059-019-1763-7

Simons, Y. B., & Sella, G. (2016). The impact of recent population history on the deleterious mutation load in humans and close evolutionary relatives. Current Opinion in Genetics & Development, 41, 150-158. doi: 10.1016/j.gde.2016.09.006

Simons, Y. B., Turchin, M. C., Pritchard, J. K., & Sella, G. (2014). The deleterious mutation load is insensitive to recent population history. Nature Genetics, 46(3), 220-224. doi: 10.1038/ng.2896

Sindičić, M., Gomerčić, T., Galov, A., Polanc, P., Huber, D., & Slavica, A. (2012). Repetitive sequences in Eurasian lynx (Lynx lynx L.) mitochondrial DNA control region. Mitochondrial DNA, 23(3), 201-207. doi: 10.3109/19401736.2012.668894

Skoglund, P., Ersmark, E., Palkopoulou, E., & Dalén, L. (2015). Ancient Wolf Genome Reveals an Early Divergence of Domestic Dog Ancestors and Admixture into High-Latitude Breeds. Current Biology, 25(11), 1515-1519. doi: 10.1016/j.cub.2015.04.019

Skoglund, P., Malmstrom, H., Omrak, A., Raghavan, M., Valdiosera, C., Gunther, T., … Jakobsson, M. (2014). Genomic Diversity and Admixture Differs for Stone-Age Scandinavian Foragers and Farmers. Science, 344(6185), 747-750. doi: 10.1126/science.1253448

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. Genetics, 195(3), 693-702. doi: 10.1534/genetics.113.154138

Slatkin, M. (1985). Gene Flow in Natural Populations. Annual Review of Ecology and Systematics, 16(1), 393-430. doi: 10.1146/annurev.ecolsys.16.1.393

Smeds, L., Warmuth, V., Bolivar, P., Uebbing, S., Burri, R., Suh, A., … Ellegren, H. (2015). Evolutionary analysis of the female-specific avian W chromosome. Nature communications, 6, 7330. doi: 10.1038/ncomms8330

Smith, T. C. A., Arndt, P. F., & Eyre-Walker, A. (2018). Large scale variation in the rate of germ-line de novo mutation, base composition, divergence and diversity in humans. PLOS Genetics, 14(3), e1007254. doi: 10.1371/journal.pgen.1007254

Smukowski, C. S., & Noor, M. A. F. (2011). Recombination rate variation in closely related species. Heredity, 107(6), 496-508. doi: 10.1038/hdy.2011.44

Sommer, R. S., & Benecke, N. (2006). Late Pleistocene and Holocene development of the felid fauna (Felidae) of Europe: a review. Journal of Zoology, 269(1), 7-19. doi: 10.1111/j.1469-7998.2005.00040.x

Sommer, R. S., & Nadachowski, A. (2006). Glacial refugia of mammals in Europe: Evidence from fossil records. Mammal Review, 36(4), 251-265. doi: 10.1111/j.1365-2907.2006.00093.x

Song, Y. S., & Steinrücken, M. (2012). A simple method for finding explicit analytic transition densities of diffusion processes with general diploid selection. Genetics, 190(3), 1117-1129. doi: 10.1534/genetics.111.136929

Srbek-Araujo, A. C., Haag, T., Chiarello, A. G., Salzano, F. M., & Eizirik, E. (2018). Worrisome isolation: noninvasive genetic analyses shed light on the critical status of a remnant jaguar population. Journal of Mammalogy, 99(2), 397-407. doi: 10.1093/jmammal/gyy007

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312-1313. doi: 10.1093/bioinformatics/btu033

Straus, L. G., Bicho, N., & Winegardner, A. C. (2000). The Upper Palaeolithic settlement of Iberia: first-generation maps. Antiquity, 74(285), 553-566. doi: 10.1017/S0003598X00059913

Sunquist, M., & Sunquist, F. (2002). Wild cats of the world. Chicago: The University of Chicago Press.

Sved, J. A., & Feldman, M. W. (1973). Correlation and probability methods for one and two loci. Theoretical population biology, 4(1), 129-132.

Svendsen, J. I., Alexanderson, H., Astakhov, V. I., Demidov, I., Dowdeswell, J. A., Funder, S., … Stein, R. (2004). Late Quaternary ice sheet history of northern Eurasia. Quaternary Science Reviews, 23(11-13), 1229-1271. doi: 10.1016/j.quascirev.2003.12.008

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology, 7(4), 453-464. doi: 10.1046/j.1365-294x.1998.00289.x

Terekhanova, N. V., Seplyarskiy, V. B., Soldatov, R. A., & Bazykin, G. A. (2017). Evolution of Local Mutation Rate and Its Determinants. Molecular Biology and Evolution, 34(5), 1100-1109. doi: 10.1093/molbev/msx060

Thatte, P., Joshi, A., Vaidyanathan, S., Landguth, E., & Ramakrishnan, U. (2018). Maintaining tiger connectivity and minimizing extinction into the next century: Insights from landscape genetics and spatially-explicit simulations. Biological Conservation, 218, 181-191. doi: 10.1016/j.biocon.2017.12.022

Timmermann, A., & Friedrich, T. (2016). Late Pleistocene climate drivers of early human migration. Nature, 538(7623), 92-95. doi: 10.1038/nature19365

Torres, R., Stetter, M. G., Hernandez, R. D., & Ross-Ibarra, J. (2020). The Temporal Dynamics of Background Selection in Nonequilibrium Populations. Genetics, 214(4), 1019-1030. doi: 10.1534/genetics.119.302892

Turvey, S. T., & Crees, J. J. (2019). Extinction in the Anthropocene. Current Biology, 29(19), R982-R986. doi: 10.1016/j.cub.2019.07.040

Urbanek, S. (2013). png: Read and write PNG images.

Valbuena-Carabaña, M., de Heredia, U. L., Fuentes-Utrilla, P., González-Doncel, I., & Gil, L. (2010). Historical and recent changes in the Spanish forests: A socio-economic process. Review of Palaeobotany and Palynology, 162(3), 492-506. doi: 10.1016/j.revpalbo.2009.11.003

Valdiosera, C. E., García-Garitagoitia, J. L., Garcia, N., Doadrio, I., Thomas, M. G., Hänni, C., … Götherström, A. (2008). Surprising migration and population size dynamics in ancient Iberian brown bears (Ursus arctos). Proceedings of the National Academy of Sciences of the United States of America, 105(13), 5123-5128. doi: 10.1073/pnas.0712223105

Valdiosera, C. E., García, N., Anderung, C., Dalén, L., Crégut-Bonnoure, E., Kahlke, R. D., … Barnes, I. (2007). Staying out in the cold: Glacial refugia and mitochondrial DNA phylogeography in ancient European brown bears. Molecular Ecology, 16(24), 5140-5148. doi: 10.1111/j.1365-294X.2007.03590.x

Vashon, J. (2016). Lynx canadensis. The IUCN Red List of Threatened Species 2016, 8235, e.T12518A101138963. doi: https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T12518A101138963.en

Vila, C., Amorim, I. R., Leonard, J. A., Posada, D., Castroviejo, J., Petrucci-Fonseca, F., … Wayne, R. K. (1999). Mitochondrial DNA phylogeography and population history of the grey wolf Canis lupus. Molecular Ecology, 8(12), 2089-2103. doi: 10.1046/j.1365-294x.1999.00825.x

Vilallonga, M. A. (1998). Labor migration during the first phase of Basque industrialization: The labor market and family motivations. The History of the Family, 3(2), 199-219. doi: 10.1016/S1081-602X(99)80242-5

Villaluenga, A., Castaños, P., Arrizabalaga, A., Antonio, J., & Alustiza, M. (2012). Cave Bear (Ursus spelaeus Rosenmüller Heinroth, 1794) and Humans During the Early Upper Pleistocene (Lower and Middle Palaeolithic) in Lezetxiki, Lezetxiki II and Astigarragako Kobea (Basque Country, Spain). Preliminary Approach. Journal of Taphonomy, 10, 3-4.

Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013). Detecting Natural Selection in Genomic Data. Annual Review of Genetics, 47(1), 97-120. doi: 10.1146/annurev-genet-111212-133526

Von Arx, M., Breitenmoser-Würsten, C., Zimmermann, F., & Breitenmoser, U. (2004). Status and conservation of the Eurasian lynx (Lynx lynx) in 2001. KORA Bericht, 19.

Wang, J., Street, N. R., Scofield, D. G., & Ingvarsson, P. K. (2016). Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related populus species. Genetics, 202(3), 1185-1200. doi: 10.1534/genetics.115.183152

Werdelin, L. (1981). The evolution of lynxes. Annalis Zoologici Fennici, 18, 37-71. doi: 10.2307/23734102

Westbury, M. V., Hartmann, S., Barlow, A., Preick, M., Ridush, B., Nagel, D., … Hofreiter, M. (2020). Hyena paleogenomes reveal a complex evolutionary history of cross-continental gene flow between spotted and cave hyena. Science Advances, 6(11), 1-11. doi: 10.1126/sciadv.aay0456

Wickham, H. (2007). Reshaping Data with the {reshape} Package. Journal of Statistical Software, 21(12), 1-20.

Wickham, H. (2009). ggplot2: elegant graphics for data analysis. Springer-Verlag New York.

Wickham, H. (2018). stringr: Simple, Consistent Wrappers for Common String Operations.

Wickham, H., & Henry, L. (2018). tidyr: Easily Tidy Data with «spread()» and «gather()» Functions.

Wickham, H., Francois, R., Henry, L., & Müller, K. (2017). dplyr: A Grammar of Data Manipulation.

Williams, A. L., Genovese, G., Dyer, T., Altemose, N., Truax, K., Jun, G., … Przeworski, M. (2015). Non-crossover gene conversions show strong GC bias and unexpected clustering in humans. eLife, 4. doi: 10.7554/eLife.04637

Williams, A. V, Nevill, P. G., & Krauss, S. L. (2014). Next generation restoration genetics: applications and opportunities. Trends in plant science, 19(8), 529-537. doi: 10.1016/j.tplants.2014.03.011

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proceedings of the VI International Congress of Genetics, 1, 356-366.

Xie, Y. (2014). knitr: A Comprehensive Tool for Reproducible Research in {R}. En V. Stodden, F. Leisch, & R. D. Peng (Eds.), Implementing Reproducible Computational Research. Chapman and Hall/CRC.

Xie, Y. (2015). Dynamic Documents with {R} and knitr (2nd ed.). Boca Raton, Florida: Chapman and Hall/CRC.

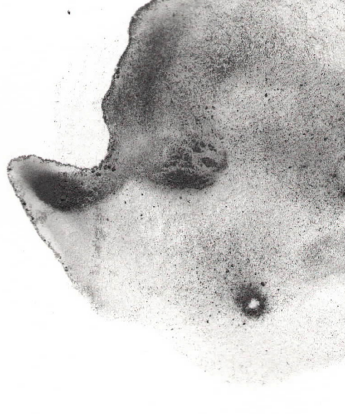
Xie, Y. (2017). knitr: A General-Purpose Package for Dynamic Report Generation in R.

Yang, M. A., Gao, X., Theunert, C., Tong, H., Aximu-Petri, A., Nickel, B., … Fu, Q. (2017). 40,000-Year-Old Individual from Asia Provides Insight into Early Population Structure in Eurasia. Current Biology, 27(20), 3202-3208.e9. doi: 10.1016/j.cub.2017.09.030

Yravedra, J. (2005). Aprovechamiento cárnico de lince (Lynx pardina) durante el Pleistoceno Superior en el interior de la Península Ibérica. Munibe, 75, 303–311.

Zimmermann, F., & Breitenmoser, U. (2007). Potential distribution and population size of the Eurasian lynx Lynx lynx in the Jura Mountains and possible corridors to adjacent ranges. Wildlife Biology, 13(4), 406-417.

AGRADECIMIENTOS



Para que mi ser pese sobre el suelo,
fue necesario un ancho espacio
y un largo tiempo:
hombres de todo mar y toda tierra,
fértiles vientres de mujer, y cuerpos
y más cuerpos, fundiéndose incesantes
en otro cuerpo nuevo.
Solsticios y equinoccios alumbraron
con su cambiante luz, su vario cielo,
el viaje milenario de mi carne
trepando por los siglos y los huesos.

"Para que yo me llame Ángel González"
Ángel González

Pocas veces he sentido tanto vértigo ante un folio en blanco, como lo siento ahora mismo. Es curioso como cuando llega ese momento que te has imaginado tanto tiempo te quedas sin palabras. Y es que es muy difícil que pueda hacer justicia a todas las personas que me han acompañado en este camino y han hecho que sea más agradable.

A José Antonio Godoy, director de esta tesis. Godo, gracias por haberme dado la oportunidad de trabajar todos estos años haciendo ciencia en un grupo que más que un grupo es una familia. Gracias también por tu generosidad y la confianza que has depositado en mi desde el principio. Siempre has estado ahí para contestar cualquier pregunta, por básica que fuera, una y mil veces, y esto no ha impedido que cuando tenía cualquier sugerencia la escuchases como si fuese la experta número uno. Ese valor que das a las palabras de quien trabaja contigo es muy poderoso. Gracias también por tu capacidad de escucha y la confianza que he sentido siempre para expresarte mis problemas o dudas. Tu labor como director ha ido más allá de una labor exclusivamente laboral, y esa relación de cariño me ha permitido en muchos momentos sacar fuerzas.

Gracias también a todas las personas que antes o después han formado parte de esta pequeña familia que es este grupo de investigación. Cuando me disponía a buscar tesis, tenía muy claro que quería rodearme de gente que sumara. Por suerte, llegué aquí, y siento que he sido tremendamente afortunada de que así haya sido. Desde que entre por la puerta, Laura, Maritxu y Elena me acogieron como si llevase toda la vida entre ellas. Fue muy fácil empezar así. En un momento especialmente duro para mi, siempre sentí su cariño. Ir a trabajar hacía mi día a día mucho más llevadero. GRACIAS. Laura, por tu ternura. Maritxu, por tu vitalidad. Elena, por tu amistad. Por los llantos y las risas. Por haberme acompañado desde el principio (¡y lo que nos queda!) y enseñarme gran parte de lo que sé. Porque trabajar en tándem es lo más divertido que hay, y yo sólo tenía que girar mi silla para encontrarte con tu sonrisa y tu rotulador en la boca diciendo "qué", siempre dispuesta a ayudar. Esta tesis es también un poco tuya. Tengo que reconocer que cuando Godo nos dijo que venía un nuevo doctorando al grupo me sentí algo nerviosa. Me encantaba la atmósfera que habíamos creado, y me daba miedo que la nueva persona rompiese esa magia. No podía ser un miedo más infundado. Dani, te uniste como una pieza perfecta para poner un poco de calma a ese remolino que formábamos Elena y yo. Has sido el ejemplo perfecto de paciencia y bondad. Siempre que estaba nerviosa encontraba en tus ojos un poquito de paz y tus palabras y consejos me han ayudado continuamente a lo largo de este camino. Gracias

por todo amigo. Gracias a Enrico; esos ratitos en la puerta me dieron la vida en un momento en el que me faltaba el aire. Sin apenas conocerme me escuchaste y me comprendiste, me diste mi espacio. Alex, por tu alegría continua (y tus chistes malos) y por darme la oportunidad de trabajar contigo. Bego (Martínez), por tu ejemplo de fuerza y constancia. Gracias también a todas aquellas personas que fueron aves de paso: Karolina, Bárbara y tantas otras personas de las que he aprendido tanto.

A los *carnívoros*, porque siempre nos han acogido con cariño y con ellos he sentido que formaba parte de un legado que trasciende el tiempo. Miguel (Delibes) gracias por ser un referente para mi en todos los aspectos. Miguel (Clavero) me siento muy afortunada de haberte conocido y de que aceptases mi propuesta para trabajar juntos. Desprendes alegría y sencillez en un mundo en el que a veces sobra ego. Bego (Adrados), gracias por tu cariño y tus ánimos constantes que tan importantes han sido para mi en tantos momentos, ¡como echo de menos esos desayunos! Rubén, contagias motivación a quien está a tu lado, es un gusto escucharte hablar de ciencia, y contar contigo como amigo. No dejes que nadie te quite nunca ese empuje, pero cuidado no te lleve a lesionarte (que nos conocemos). Gemma, gracias por contarme historias viejunis de la EBD y por estar a mi lado y apoyarme con esas maneras tan tuyas, entre la ternura y la firmeza. Miguel (Jácome), amigo, poca gente sabe cómo han sido estos años para mi mejor que tú. Doy gracias a la vida por haber coincidido contigo y por ese segundo contrato (gracias Pedro Jordano por tu ejemplo, tu ayuda y por traer a Miguel de vuelta) que te permitió pasar más tiempo aquí y que confirmó lo que ya sospechábamos, que nuestra amistad tiene que venir de otra vida y que ya se puede cruzar el Atlántico por medio, que lo que se ha fraguado en Sevilla no va a desaparecer.

A todas las personas con las que he coincidido primero en la etapa de máster, que fue tan importante para mi, y después gracias a la EBD y que han hecho que esta época haya sido más bonita y más divertida. No están todos los que son, pero son todos los que están, más o menos en orden cronológico: Javi Manzano (eres un ejemplo para mi), Noe, Isa, Rafa, Carlines (gracias por tu confianza en mi desde el principio), Marina, Esther, Chuso, Álvaro Dugo (¡deseando empezar esta nueva etapa con tu compañía!), Paloma, Rosa, Joan (gracias por esas calçotadas), Eneko (¡por muchas historias médicas y chistes malos más!), Irene (poquito a poco ganando

hueco hasta ser compañeras de reparto), Víctor y Marta (lo que ha unido el artisteo ¡que no lo separe el tiempo ni la distancia!), Vanina (gracias por todos esos ratitos en el despacho, tu confianza en mi me ha ayudado mucho), Ampi (¿dónde será el próximo reencuentro?), Sara, Isa, Christoph, Hyeunji, y a todos los nuevos fichajes (que ya no sois tan nuevos) y que han traído aire fresco y ¡mucha luz! especialmente a Elena Quintero que junto a Fer a este paso repueblan la península.

A toda esa gente que hace una eficiente labor que permite que la ciencia siga hacia delante. Desde los servicios de administración, pasando por informática, laboratorios, colecciones, limpieza, seguridad y mantenimiento. Sin vosotros/as no sería posible. Sobre todo al equipo de informática por la ayuda prestada siempre, a los técnicos del LAST y de colecciones por echarme una mano cuando lo he necesitado y a los técnicos del LEM, especialmente a Ana Piriz. Gracias porque en tu despacho siempre encontré complicidad y ayuda (y una persona con mi misma talla, lo cual ha sido bastante útil en un par de ocasiones). Esta tesis también le debe mucho a todas esas personas que se encargan de recoger y preservar muestras de todas las épocas a lo largo y ancho de Asia y Europa. Probablemente nunca leáis estos agradecimientos, pero muchas gracias por vuestro trabajo, en ocasiones tan desagradecido, pero tan necesario.

Gracias a la tesis he tenido la oportunidad de visitar lugares fascinantes y conocer a personas que me han aportado mucho científica y personalmente.

Thanks Michi for welcomed me in your lab in Potsdam. I had a beautiful experience there, and I have learnt a lot! Johanna, and Axel thank you for your time and your kind words. Working with you it's being a pleasure!!

Krzysztof thank you for your collaboration during all these years, and specially for showing us this magic place that Białowieża is. Being there has been one of the most special moments during my PhD.

Gracias a mi comité de tesis, Iván Gómez Mestre, Renaud Vitalis y Aida Andrés, por prestarme su tiempo y su conocimiento para mejorar esta tesis. Aida, gracias por haberme acogido en Londres durante un mes estupendo y abrirme las puertas a tu grupo, por ser siempre tan generosa conmigo y por ser un referente muy valioso para mi en muchos aspectos. Las charlas que hemos compartido, en Sevilla, en Suiza o en Londres me han ayudado mucho, y aunque no ha podido quedar reflejado en esta tesis todo el trabajo que hemos hecho juntas, si queda el aprendizaje que de él he adquirido.

Esta es la culminación de un camino que comencé a recorrer en el Departamento de Genética de la UGR de la mano de Pepi y Juan Pedro, pasando después por Barcelona con Elena Casacuberta y por Leuven en Bélgica, con Kevin Verstrepen. Sin ese largo recorrido no hubiese llegado aquí, sin la guía de todas las personas que aparecieron en el camino para aportar su punto de vista, su experiencia y de las que he aprendido tanto. Gracias a todas ellas, especialmente a las mujeres, por mostrarme el camino con su ejemplo.

A Manuel, por tu amistad de tantos años y por esa maravilla de ilustraciones que me va a acompañar siempre. Ojalá podamos pronto volver a hacer cositas juntos, ¡te admiro mucho amigo! A Manolo, por ayudarme a culminar con estilo.

Durante este tiempo en Sevilla he tenido la suerte de encontrarme con gente maravillosa. A Blanca, Lola, por compartir con nosotros la agenda cultural y ser fuente constante de anécdotas y risas. A Carmen, por acompañarme en los inicios. A Ana y Laura (¡os debemos una visita a la isla!). A Pedro y María, por hacernos formar parte de vuestra vida y compartir con nosotros el placer de una buena comida. Gracias por todos esos momentos que me han ayudado a desconectar a lo largo de todos estos años.

Mis clases, ensayos y proyectos de danza han sido una vía de escape y una fuente de motivación e inspiración constante. Gracias a Manuela y todas mis compañeras, muy especialmente a Anita, Camille, Macarena y Michelle. Amigas, habéis supuesto un sostén muy importante para mi. Bailando a vuestro lado me he sentido libre y poderosa y no veo el momento en el que podamos volver a improvisar juntas. Gracias por vuestro cariño y vuestra entrega.

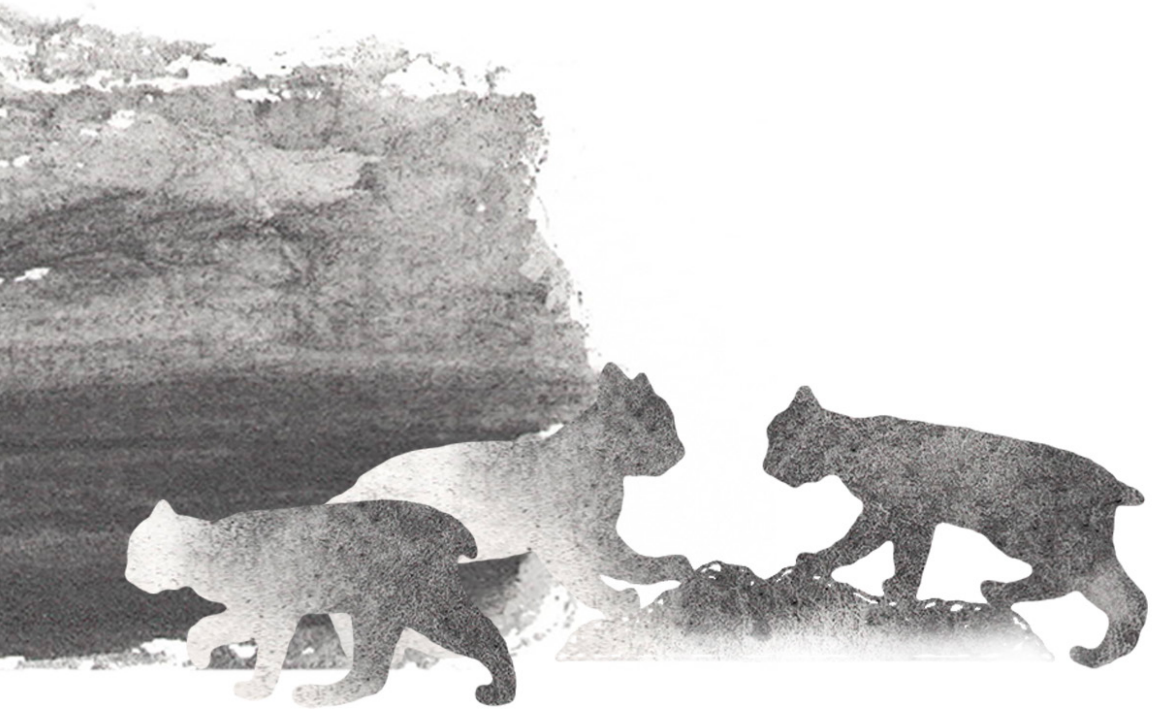
A las personas junto a las cuales descubrí mi amor por la biología. Marisa, Sara, Yedra, Alfredo, Antonio, Vito, Juan, Fernando, Marta, Jorge, Caro, María y tantas personas más. SOS necesito una escapada YA. Marisa y Sara, gracias por los viajes improvisados, por las risas y por tanto cariño siempre. Con vosotras al fin del mundo, en furgo, a dedo, o como haga falta. Yedra te admiro y te quiero por igual (¡mucho!). Alfredo gracias por las escapadas a Andújar a ver al "bicho" cada Navidad. Fernando gracias por recibirnos allí, y enseñarnos todos los rinconcitos de esa maravillosa sierra. ¡Qué ganas tengo de volver a ir! Juan y Antonio gracias por ser mis guías en vuestra casa, y enseñarme todos los secretos de esos lugares tan maravillosos.

A todas mis amigas en Granada. A Eli, Fati y Eva. A Pablo y Alba. Aunque nos veamos menos de lo que nos gustaría, sois el hogar al que da gusto volver. Me recargáis las pilas y me ayudáis a poner los pies en la tierra, recordar de donde vengo y hacia donde quiero ir.

A mi familia que la quiero tanto. A mis tías, Ana, Mari Carmen y Carmen y mis tíos Paco, Antonio y Paco. A mis primas, Ana, Elena y Elena y a mis primos, Jose Andrés, Antonio Jesús y Carlos. A mis abuelos, que aunque ya no están aquí han sido ejemplo para mi. Especialmente a mi abuela María, por su fuerza y a mi abuelo Alejandro por su buen humor. A Paqui y Manolo. Familia, los momentos juntos son lo más valioso que tengo en la vida. Sois para mi el refugio que me recuerda lo verdaderamente importante.

A los tres pilares de mi vida, mamá, Marta y Fer. Gracias por cuidarme tanto. Mamá, tu fuerza, tu empuje y tu libertad han sido mi mayor inspiración estos años. Tu comprensión y apoyo, el mejor refugio. Marta, hermana, mi otra mitad, mi mejor amiga. Tu alegría, valentía y firmeza en la vida son un ejemplo a seguir. Un silencio vuestro me calma y me consuela más que todas las palabras del mundo y la risa constante con vosotras me reconcilia con la vida. Fer, gracias por acompañarme en la vida todos estos años y por hacerme tanto bien. Soy tremendamente afortunada de tenerte a mi lado, y no hay palabras que puedan expresar el agradecimiento que siento por como me has cuidado todo este tiempo. Os quiero muchísimo.

A papá; aunque no pudiste ver como empezaba esta aventura por muy poquitos meses, me enseñaste a mirar la naturaleza con curiosidad. Gracias por haberme dado confianza en mi misma, y por educarme, junto con mamá, para luchar por un mundo mejor. Sé que hoy estarías orgulloso de mi.



TESIS DOCTORAL 2020

María Lucena Pérez

Estación Biológica Doñana CSIC

UNIVERSIDAD PABLO DE OLAVIDE SEVILLA AÑO 1997