

RESEARCH ARTICLE

Open Access



Retention of functional variation despite extreme genomic erosion: MHC allelic repertoires in the *Lynx* genus

Elena Marmesat¹, Krzysztof Schmidt², Alexander P. Saveljev³, Ivan V. Seryodkin^{4,5} and José A. Godoy^{1*} 

Abstract

Background: Demographic bottlenecks erode genetic diversity and may increase endangered species' extinction risk via decreased fitness and adaptive potential. The genetic status of species is generally assessed using neutral markers, whose dynamic can differ from that of functional variation due to selection. The MHC is a multigene family described as the most important genetic component of the mammalian immune system, with broad implications in ecology and evolution. The genus *Lynx* includes four species differing immensely in demographic history and population size, which provides a suitable model to study the genetic consequences of demographic declines: the Iberian lynx being an extremely bottlenecked species and the three remaining ones representing common and widely distributed species. We compared variation in the most variable exon of the MHC I and MHC II-DRB loci among the four species of the *Lynx* genus.

Results: The Iberian lynx was characterised by lower number of MHC alleles than its sister species (the Eurasian lynx). However, it maintained most of the functional genetic variation at MHC loci present in the remaining and genetically healthier lynx species at all nucleotide, amino acid, and supertype levels.

Conclusions: Species-wide functional genetic diversity can be maintained even in the face of severe population bottlenecks, which caused devastating whole genome genetic erosion. This could be the consequence of divergent alleles being retained across paralogous loci, an outcome that, in the face of frequent gene conversion, may have been favoured by balancing selection.

Keywords: Mhc, Lynx, Balancing selection, Natural selection, Genetic functional diversity, Bottleneck, Drift, Iberian lynx, Conservation genetics, Recombination

Background

It is crucial for conservation and evolutionary biology to understand how population bottlenecks shape genetic diversity, as its loss might increase extinction risk via decreased reproduction, survival and adaptive potential [1]. Genetic variation is commonly assessed using neutral markers like microsatellites, which have proven useful to infer demography, population structure, and kinship, among other applications. However, neutral markers by definition are not subjected to natural selection, so we cannot assume they adequately reflect the dynamics of functional genetic variation that is most relevant for

species persistence (e.g., [2, 3]). We thus need to focus on functionally relevant loci to fully understand how population declines may increase extinction risk through their effects on genetic variation [4–7]. A high concern in conservation, which is influenced by such functional genetic variation, is the species capacity to immunologically respond to pathogens, because it may largely determine their survival and long term viability [8, 9].

The Major Histocompatibility Complex (MHC) is a multigene family often described as the most important genetic component of the mammalian immune system [10]. The MHC encodes cell surface glycoproteins that bind antigens derived from pathogens or parasites and present them to T-lymphocytes, triggering the adaptive immune response. There are two major types of MHC gene families: MHC class I and MHC class II. MHC class I genes are expressed

* Correspondence: godoy@ebd.csic.es

¹Department of Integrative Ecology, Estación Biológica de Doñana (CSIC), C/ Américo Vespucio, 26, 41092 Sevilla, Spain

Full list of author information is available at the end of the article



on the cell membrane of all nucleated cells and defend against intracellular threats (as viral pathogens) and malignant cells. MHC class II genes are mainly expressed on specialized antigen-presenting cells of the immune system, which monitor the extracellular environment and use MHC receptors to present peptides of extracellular pathogens to the T-cells.

Previous studies on MHC have evidenced the long-term operation of natural selection (reviewed in [11]) and highlighted its important implications in evolutionary ecology and evolution (reviewed in [12]). It predisposes the MHC as a model for studying the non-neutral genetic evolution in non-model species.

The MHC genes are expected to evolve through natural selection in response to the selection pressures imposed by pathogens. The nature of the molecule, that interacts both with the antigens (i.e., the Antigen Binding Sites, onwards ABS) and the cellular and immune system components, makes it subject to different types of natural selection. On one hand, positive selection signals are expected in the ABS as they may evolve to bind novel antigens carried by pathogens, which in turn may evolve to escape recognition. On the other hand, purifying selection will act on non-ABS regions where most non-synonymous substitutions are likely to be deleterious. Additionally, MHC genes are one of the best characterized targets of balancing selection, which maintain high levels of polymorphisms by either overdominance, rare-allele advantage, fluctuating selection and divergent allele selection [13, 14]. All these forms of evolution are not mutually exclusive, on the contrary, they are likely to act at the same time and interact with each other. For example, divergent allele selection is expected to boost the effect of overdominance and rare-allele advantage [14]. Evidence for long-term balancing selection includes the occurrence of trans-species polymorphisms (TSP), i.e., alleles that are more similar between related species than alleles within each species and this similarity is not due to convergent evolution. TSP is generated by the passage of alleles from ancestral to descendant species due to increased coalescence times in the loci target of balancing selection. Another signature of balancing selection is the excess of non-synonymous polymorphisms, especially at ABS [15]. Since the net result of balancing selection is that it maintains polymorphisms, it is a matter of debate in the context of species conservation whether it can hamper the loss of genetic diversity driven by genetic drift and delay the fixation of alleles. The empirical evidence for this is conflicting, with some MHC studies finding high MHC diversity in otherwise genetically eroded populations (e.g., [2, 16–18]), and some others finding the opposite (e.g., [19–22]).

The MHC diversity and TSP can be studied at three different levels. First, the nucleotide level which is the

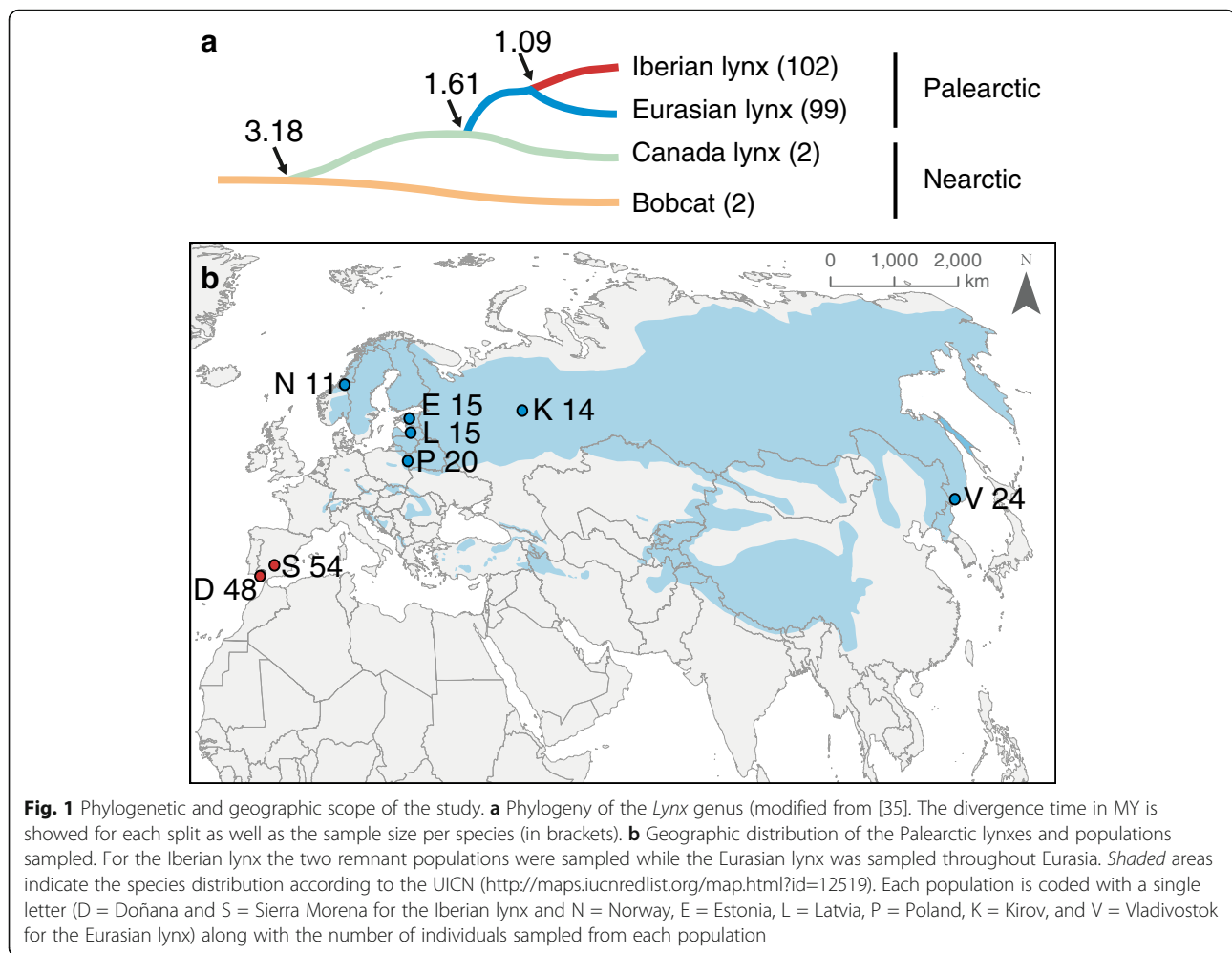
lower and more stringent one. Second, the protein level in which slightly different nucleotide sequences are functionally the same as they translate into the same protein sequence. Third, the supertype level, that which groups functionally similar alleles into the same superotypes. MHC alleles defined by nucleotide or even amino acid differences may be functionally similar if they bind similar repertoire of antigens. Hence, Functionally redundant alleles can be identified and grouped in the same super-type based on their biochemical similarities at amino acids known to interact with the antigen (i.e., ABS) [23].

In felids MHC genes have received special attention due to the role of domestic cat as a model for human diseases [24–27] and because of its possible impact on endangered wild felid conservation, including the paradigmatic case of the cheetah [18, 28–32].

The genus *Lynx* being composed of four species: the bobcat (*L. rufus*), the Canada lynx (*L. canadensis*), the Eurasian lynx (*L. lynx*) and the Iberian lynx (*L. pardinus*) ([33, 34]; Fig. 1a) provide a proper model on which to assess the genetic consequences of demographic decline and fragmentation on the functional genetic variation. The four species differ immensely in recent demographic history, population size, and genetic status as well as their phylogenetic relationships. In our model the Iberian lynx represents the highly bottlenecked species and the other three are common and widely distributed species. The Iberian lynx and the Eurasian lynx are sister species, both sharing their most recent common ancestor around 1.5–1.09 Mya, as estimated by phylogenomic methods ([33, 34], Fig. 1a), whereas coalescent methods applied to the whole-genome sequences suggest a more recent divergence (312.2 kya; 95% CI: 323.1–179.4 kya), followed by a long period of gene flow that ceased recently (2.473 kya; 95% CI: 126.8–0 kya) [35].

The Iberian lynx was recognized as the most endangered felid in the world, due to a dramatic decline of its entire population during the last two decades of twentieth century from above 1000 to 100 individuals [36, 37]. In sharp contrast to the Iberian, the Eurasian lynx shows one of the widest distributions among felids, covering a high variety of environmental, ecological, climatic and demographic conditions [38]. The current distribution of the Eurasian lynx extends across Eurasia with the core of its population in Russia. However, in Europe, the Eurasian lynx has also gone through a recent decline that has resulted in its extirpation from most of western Europe and the fragmentation of its westernmost range [39–41]. The bobcat and the Canada lynx are distributed across North America in southern and northern latitudes, respectively, with areas of range overlap in northern United States and southern Canada.

Concordant with the demographic history of the species the Iberian lynx neutral genetic variability is extremely reduced [42], its current genome- and species-



wide diversity being the lowest ever reported for any species [35]. In contrast, the bobcat, Canada, and Eurasian lynx, based on mtDNA and microsatellites have shown moderate levels of genetic diversity and little genetic structure across most of their ranges [39, 43, 44]. Hence, the *Lynx* genus arises as a good model to study patterns of functional genetic variation following extreme overall genetic erosion.

Here we characterize the allele repertoire of the most variable exon (exon 2) of MHC class I and class II-DRB genes in the *Lynx* genus, and test whether a genetically eroded species can maintain a similar range of functional diversity to demographically and genetically healthy related species. More specifically, we tackle the following questions: i) What are the levels of diversity at MHC class I and class II-DRB loci in the *Lynx* genus?, ii) are there signatures of positive or purifying selection in lynx MHC alleles?, iii) How are the different lynx alleles related among them and to those in other felids; what is the level of TSP across species?, iv) how does the allele repertoire translate into functional diversity in the form

of supertypes?, and v) how much of this functional diversity is maintained in Iberian lynx compared to the genetically healthier relatives?

Results

Data processing and allele validation

We obtained a total of 554,478 reads with a mean coverage of 2670 (range: 9–16,729) and 563 (range: 12–8964) for MHC I and MHC II-DRB per replicates respectively. Only replicates with coverages above 500 reads for MHC I and 100 for MHC II-DRB were taken into account. Within the *Lynx* genus we found a total of 37 MHC I alleles and 13 MHC II-DRB alleles (KY769287-KY769350 and KY769351-KY769367 respectively, for sequence alignments see Additional file 2: Tables S1 and S2). The number of alleles found within individuals ranged from 10 to 16 for MHC I and from 3 to 6 for MHC II-DRB suggesting a minimum of 5 and 3 amplified loci, respectively. Six of the *L. lynx* MHC II-DRB alleles were previously found on Chinese Eurasian lynxes (*Lyly-DRB*1*, **2*, **3*, **7*, **8* and **10*) [29]. The alleles not found in our samples but found in Wang et

al. [29] were not included in the analysis, as we considered that the use of a different genotyping methodology (both in the wet lab and filtering steps) could bias the diversity of Eurasian lynx compared with the other lynxes. Nevertheless, we included them in the analysis comparing the lynx alleles sequences with published alleles of Felidae, as most of the published sequences also differ in how they were generated.

None of the MHCII-DRB alleles showed signals of pseudogenization. Therefore, all of them were considered as functional alleles in downstream analyses. On the other hand, seven MHCI alleles presented pseudogenization signals and were excluded from diversity, recombination, selection and supertype analyses. Three of these alleles included deletions, another three insertions (two of them causing a stop codon to appear within the sequenced exon) and the last one harboured a premature stop codon but no indels. They were found across all lynx species (3 in *L. pardinus*, 3 in *L. lynx*, 2 in *L. canadensis* and 5 in *L. rufus* for) and the pattern of sharedness was similar to that found in functional ones (further details in Additional file 1: Table S3).

Recombination and phylogenetic relationships of the alleles

Both MHC genes showed signals of recombination. For the MHCI sequences SBP inferred recombination with 100% support and suggested position 201 as the recombination breakpoint, whereas GARD lacked power to infer recombination. SBP also inferred recombination in MHCII-DRB sequences and indicated nucleotide position 114 as a recombination breakpoint; GARD confirmed such breakpoint and pointed to an additional one at nucleotide 209. Recombination events contribute to the complexity of the phylogenetic networks (Additional file 2: Figures S1 and S2),

with multiple connection links between alleles. The phylogenetic networks also show little structuring of MHC alleles across species, with most clades containing alleles of all lynx species.

We found many alleles to be shared among species within the genus (Fig. 3; nb, even though Fig. 3 focus on supertypes description nucleotide and amino acid differences are also depicted). The percentage of alleles within each species found in any other species was: 100% in *L. pardinus*, 82% in *L. lynx*, 69% in *L. canadensis* and 10% in *L. rufus* for MHCI, and 71% in *L. pardinus*, 75% in *L. lynx*, 100% in *L. canadensis* and 40% in *L. rufus* for MHCII-DRB (Figs. 2 and 3). Shared alleles were less frequent in *L. rufus*, as expected by its basal position in the phylogeny and its older separation from the rest (3.48 MYA versus the *L. canadensis* split which occurred 1.61 MYA) [34]. In contrast, the trio *L. canadensis*, *L. lynx* and *L. pardinus* showed similar high levels of shared polymorphisms despite *L. canadensis* splitting from the Palearctic species 0.52 MY before *L. lynx* and *L. pardinus* separated. The only two MHC alleles that are private of *L. pardinus* are further discussed below (see Functional diversity retention section).

Signatures of selection

To search for signatures of long-term positive and purifying selection we compared the rates of nonsynonymous (d_N) and synonymous (d_S) substitutions for ABS and non-ABS (Table 1), which were congruent across the genus. None of the MHC gene families studied showed d_N/d_S ratios above one (one of the signals of positive selection) neither at ABS nor at non-ABS. On the contrary, they showed purifying selection signals (i.e., d_N and d_S ratios below one). At the MHCI exon 2 a slight signature of purifying selection was detected at the ABS of *L. canadensis* and *L. pardinus* and

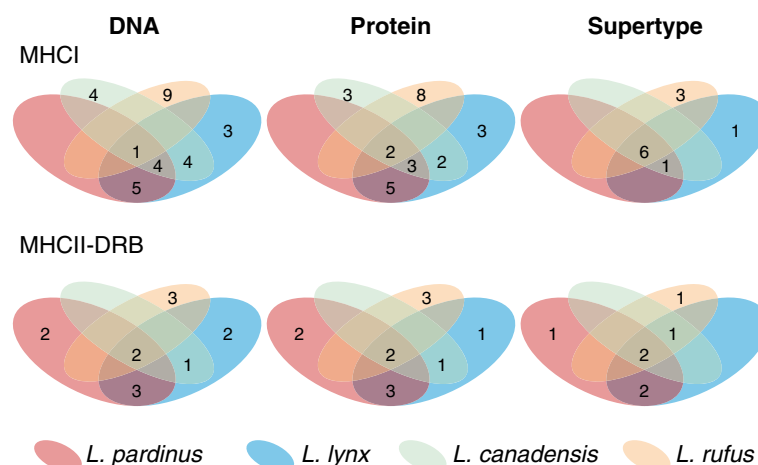


Fig. 2 Trans-species polymorphism at MHCI and MHCII-DRB in the Lynx genus. The number of shared nucleotide and protein sequences and functional types (supertypes) between lynx species is shown

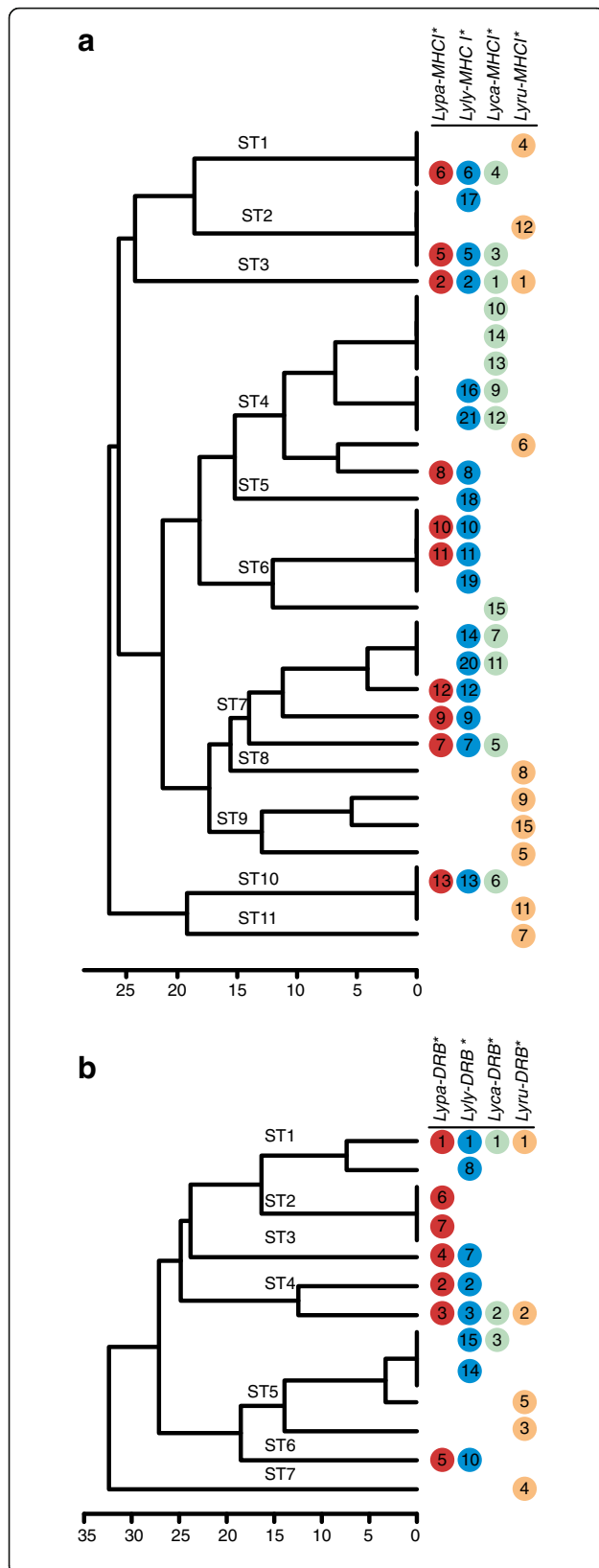


Fig. 3 Lynx MHC class I and class II DRB supertypes. We used average hierarchical clustering and a cut-off threshold of Euclidean distance ≥ 15 to classify alleles into distinctive functional groups (supertypes, ST) based on functional similarity of their ABS. Alleles with identical amino acids at ABS are represented as polytomies at 0 Euclidean distance. Alleles in the same line are identical at the nucleotide level. Alleles are specified with their allele number in each species

at the non-ABS of *L. canadensis* and *L. lynx*. On the other hand, MHCII-DRB exon 2 showed strong signals of purifying selection only at the non-ABS for all species but *L. canadensis*.

Even though we did not find positive selection at ABS in the form of an elevated d_N/d_S ratio), we found that for MHC I d_S at the ABS were from five to ten times higher than d_S at the non-ABS. This indicates an older origin of ABS regions when compared to non-ABS, what has been interpreted as a signal of gene-conversion coupled with positive selection [45].

Regarding the site-by-site set of analyses, we found SLAC test to lack power to detect selected sites on our small datasets and REL to suffer from a high rate of false positives- probably because our alignment is too short [46]. Interestingly, we found a good correlation between FEL and FURBAR results. As such analyses are the most suitable for intermediate size datasets, we considered under selection those sites identified by both tests or those identified in one and with borderline values on the other. For MHC I only one amino acid showed signals of positive selection, and ten of negative selection (Additional file 1: Tables S4 and S5). Similarly, for MHCII-DRB the signals of negative (six amino acids) were more pervasive than of positive selection (one amino acid) (Additional file 1: Tables S6 and S7). As expected, we found all negatively selected codons to be non-ABS and the positively selected site in MHCII-DRB to be an ABS. However, the positively selected site in MHC I was not an ABS.

Functional diversity retention: Genetic diversity and superotypes

The alignment of lynx MHC I (229 bp) and MHCII-DRB (247 bp) nucleotide sequences revealed 86 and 45 variable sites, respectively (Additional file 1: Tables S1 and S2). At the protein level, the sequenced MHC I α_1 domain (i.e., the translation of the MHC I exon 2) consisted of 76 residues (positions 8–83; Additional file 1: Table S5) with 45 segregating sites, and the sequenced MHCII-DRB β_1 domain (i.e., the translation of the MHCII-DRB exon 2) consisted of 82 amino acids (positions 9–90; Additional file 1: Table S7) with 28 variable sites. In both cases, 14 of the variable sites were ABS. Genetic distance patterns along the molecule were similar for the four lynx species (Table 2). The overall mean genetic distance was ~ 0.1 for both gene classes at nucleotide sequences, whereas it increased to ~ 0.2

Table 1 Nonsynonymous (d_N) and synonymous (d_S) substitutions (\pm standard error) and their ratio in ABS and non-ABS for MHC I and MHCII-DRB sequences

Species	ABS				non-ABS				ABS vs non-ABS	
	d_N	d_S	d_N/d_S	P	d_N	d_S	d_N/d_S	P	$d_{S,ABS}/d_{S,non-ABS}$	P
MHC I (17/59)										
Total	0.26 \pm 0.022	0.6 \pm 1.39	0.43	n.s	0.06 \pm 0.001	0.08 \pm 0	0.73	***	7.706	***
<i>L. pardinus</i>	0.3 \pm 0.027	0.68 \pm 2.35	0.44	*	0.07 \pm 0.001	0.1 \pm 0.01	0.76	n.s	6.972	***
<i>L. lynx</i>	0.26 \pm 0.022	0.43 \pm 1.36	0.59	n.s	0.06 \pm 0.001	0.08 \pm 0	0.73	*	5.443	***
<i>L. canadensis</i>	0.28 \pm 0.024	0.68 \pm 1.23	0.4	**	0.06 \pm 0.001	0.09 \pm 0	0.67	**	7.466	***
<i>L. rufus</i>	0.33 \pm 0.032	1.01 \pm 1.55	0.33	n.s	0.08 \pm 0.002	0.1 \pm 0.01	0.81	n.s	9.89	***
MHCII-DRB (24/58)										
Total	0.19 \pm 0.011	0.17 \pm 0.03	1.09	n.s	0.06 \pm 0.001	0.13 \pm 0.01	0.46	***	1.354	**
<i>L. pardinus</i>	0.14 \pm 0.008	0.16 \pm 0.03	0.88	n.s	0.06 \pm 0.001	0.14 \pm 0.01	0.4	***	1.166	n.s
<i>L. lynx</i>	0.16 \pm 0.009	0.15 \pm 0.02	1.1	n.s	0.06 \pm 0.001	0.14 \pm 0.01	0.43	***	1.057	n.s
<i>L. canadensis</i>	0.19 \pm 0.011	0.2 \pm 0.03	0.99	n.s	0.07 \pm 0.002	0.16 \pm 0.01	0.48	n.s	1.251	n.s
<i>L. rufus</i>	0.26 \pm 0.016	0.23 \pm 0.03	1.11	n.s	0.07 \pm 0.001	0.12 \pm 0.01	0.59	*	1.907	**

The number of amino acids in each typology appears in brackets (ABS/non-ABS). d_N is the number of nonsynonymous substitutions and d_S the number of synonymous substitutions per site category (ABS or Non-ABS). P is the probability (using a Mann-Whitney U-test) that d_N and d_S are different at each site type or that d_S differ between ABS and Non-ABS

when considering amino acid distances. Moreover, the mean nucleotide and amino acid distances were much higher in the ABS than in the non-ABS, for both gene classes and all four species, indicating that genetic variation at the studied exons concentrate at ABS, as expected.

Admittedly, we underestimated diversity indices for Neartic lynxes due to our limited sampling, so we focus on the two Palearctic species for interspecies comparisons. The allelic richness of *L. pardinus* for MHC I exon 2 was by 41% lower than in its sister species *L. lynx* (10 and 17, respectively; Table 2). A similar pattern was found at the protein level, with 10 alleles in *L. pardinus* and 15 in *L. lynx*. However, the number of variable sites at both levels were similar- with 72 and 73 at MHC I exon 2, and 38 and 39 at MHC I α_1 domain, for Iberian and Eurasian lynx, respectively. Hence, the mean number of pairwise differences between MHC I alleles was higher for *L. pardinus* than for *L. lynx*, both at the protein ($19.97 \pm 1.36 > 16.78 \pm 0.79$) and the nucleotide levels ($24.11 \pm 1.74 > 20.07 \pm 1.00$). This pattern points to divergent allele selection as the retained subset of alleles retains most of the variable sites, despite being nearly half of the original one. The two Palearctic species showed a similar number of alleles and variable sites for MHCII-DRB exon 2 nucleotide sequences (AR = 7 and 8 and S = 47 and 49) and identical numbers in protein sequences (AR = 7 and S = 23). Therefore, the mean number of differing positions was similar for both levels (Table 2). The MHCII-DRB diversity observed in *L. pardinus* was comparable to that of *L. lynx* in all the indices calculated; indicating a remarkable retention of MHCII-DRB alleles in the former despite its extreme overall genomic erosion [35].

The MHC alleles were grouped into distinctive functional groups (supertypes) according to physicochemical similarities at ABS and using an arbitrary threshold Euclidean distance of ≥ 15 (Fig. 3). We found 11 MHC I supertypes: five shared among the four species in the genus, two shared by all species but *L. rufus*, three exclusive to *L. rufus* and another one exclusive to *L. lynx*. In the MHCII-DRB case we found 7 supertypes: two shared among the four lynx species, another two shared only among the Palearctic ones, one shared by all but *L. pardinus*, one exclusive to *L. rufus* and another one to *L. pardinus*. The total number of supertypes (MHC I/MHCII-DRB) by species was *L. pardinus* (7/5), *L. lynx* (8/5), *L. canadensis* (7/3), *L. rufus* (8/4).

The only example of alleles private to Iberian lynx were *Lypa-DRB*6* and *Lypa-DRB*7*, which differ only by one base from each other (Figs. 2 and 3). This pair of alleles only add one segregating site to the alignment (the one differentiating *Lypa-DRB*6* from *Lypa-DRB*7*) and can be explained as the result of a single recombination event between some pairs of the other extant alleles. The required recombination point is located between positions 92 and 145, within the range of sites identified in the recombination site analysis. These Iberian lynx private alleles form a separate clade in the phylogenetic networks and a unique supertype cluster.

Discussion

Here we provide evidence that species-wide functional genetic diversity can be maintained even in the face of severe population bottlenecks which caused devastating whole genome genetic erosion. Our results suggest that

Table 2 Genetic diversity in MHC I and MHCII-DRB in the *Lynx* genus. Diversity at the nucleotide, amino acid and supertype levels are estimated as the number of observed alleles (allelic richness, AR), number (S) and proportion of segregating sites [S(%)], mean number (\pm standard error) of differences among alleles (K). The latter is calculated for all sites, ABS and non-ABS

Indice	Total	<i>L. pardinus</i>	<i>L. lynx</i>	<i>L. canadensis</i>	<i>L. rufus</i>
MHC Class I					
<i>Diversity level</i>					
MHC Class I					
<i>DNA (299)</i>					
AR	30	10	17	13	10
Private		0	3	4	9
S	86	72	73	76	76
S (%)	37.55	31.44	31.88	33.19	33.19
K	19.77 \pm 0.52	24.11 \pm 1.74	20.07 \pm 1	21.51 \pm 1.42	25.8 \pm 1.64
Gen. dist. All	0.094 \pm 0.003	0.116 \pm 0.009	0.096 \pm 0.005	0.103 \pm 0.007	0.125 \pm 0.009
Gen. dist. ABS	0.23 \pm 0.01	0.27 \pm 0.02	0.23 \pm 0.01	0.24 \pm 0.02	0.29 \pm 0.02
Gen. dist. Non-ABS	0.06 \pm 0	0.08 \pm 0.01	0.06 \pm 0	0.07 \pm 0.01	0.08 \pm 0.01
<i>Protein (76)</i>					
AR	26	10	15	10	10
Private		0	3	3	8
S	45	38	39	39	40
S (%)	59.21	50	51.32	51.32	52.63
K	16.68 \pm 0.4	19.97 \pm 1.36	16.78 \pm 0.79	17.27 \pm 1.16	21.93 \pm 1.28
Gen. dist. All	0.189 \pm 0.005	0.231 \pm 0.018	0.191 \pm 0.01	0.199 \pm 0.015	0.256 \pm 0.017
Gen. dist. ABS	0.53 \pm 0.01	0.58 \pm 0.04	0.5 \pm 0.02	0.54 \pm 0.04	0.68 \pm 0.04
Gen. dist. Non-ABS	0.11 \pm 0	0.15 \pm 0.01	0.12 \pm 0.01	0.13 \pm 0.01	0.17 \pm 0.01
<i>Supertypes</i>					
Number	11	7	8	7	8
Private		0	1	0	3
MHC Class II-DRB					
<i>DNA (247)</i>					
AR	13	7	8	3	5
Private		2	2	0	3
S	61	47	49	41	56
S (%)	24.7	19.03	19.84	16.6	22.67
K	23.71 \pm 1.01	22.33 \pm 1.81	23.39 \pm 1.76	27.67 \pm 4.37	29.1 \pm 2.13
Gen. dist. All	0.104 \pm 0.005	0.097 \pm 0.008	0.103 \pm 0.008	0.122 \pm 0.021	0.129 \pm 0.01
Gen. dist. ABS	0.19 \pm 0.01	0.15 \pm 0.01	0.16 \pm 0.01	0.2 \pm 0.04	0.26 \pm 0.03
Gen. dist. Non-ABS	0.07 \pm 0	0.07 \pm 0.01	0.08 \pm 0.01	0.09 \pm 0.01	0.08 \pm 0.01
<i>Protein (82)</i>					
AR	12	7	7	3	5
Private		2	1	0	3
S	28	23	23	23	27
S (%)	34.15	28.05	28.05	28.05	32.93
K	16.17 \pm 0.7	14.63 \pm 1.23	15.72 \pm 1.25	19.51 \pm 3.73	20 \pm 1.51
Gen. dist. All	0.18 \pm 0.008	0.161 \pm 0.014	0.175 \pm 0.015	0.221 \pm 0.046	0.226 \pm 0.018
Gen. dist. ABS	0.38 \pm 0.02	0.31 \pm 0.03	0.36 \pm 0.04	0.46 \pm 0.15	0.5 \pm 0.06
Gen. dist. Non-ABS	0.11 \pm 0.01	0.11 \pm 0.01	0.11 \pm 0.01	0.14 \pm 0.02	0.14 \pm 0.01
<i>Supertypes</i>					
Number	7	5	5	3	4
Private		1	0	0	1

this could be the consequence of divergent alleles being retained across paralogous loci, an outcome that, in the face of frequent gene conversion, may have been favoured by balancing selection.

We characterized MHC variation in a quite extensive sampling of *L. pardinus* and *L. lynx* comprehending most of the species range. All previous characterization of MHC variation in lynx and other felids have used PCR with degenerate primers, followed by cloning and Sanger sequencing or SSCP, methods that are probably affected by amplification bias and other artefacts [47, 48]. Our application of improved amplification strategies, high-throughput sequencing approaches and efficient validation algorithms should minimize these problems. The combination of improved genotyping protocols and extensive Palearctic species sampling should have resulted in a rather complete and artefact-free dataset. Even though our sampling of the Nearctic species was poor, the inclusion of these two species enabled a genus-wide evaluation of genetic variation at MHC loci. To the best of our knowledge, this is the first time that the MHC variation of all the species of a genus in the Felidae family has been characterized.

We also believe that we have made a reasonably complete sampling of paralogs in each MHC family. The number of MHC loci described of the Iberian lynx genome annotation [35] is compatible with the number of alleles we found per individual in all lynx species (i.e., MHCI: 10 annotated paralogs and 10–13 alleles per individual and MHCII-DRB: 3 paralogs and 3–6 alleles per individual). Admittedly, the genome assembly could have collapsed some paralogs into one or split some others. However, we do not think this is very likely as many of the paralogs were assembled from different fosmid pools [35]. Moreover, these numbers of paralogs is similar to the numbers found in domestic cat, where the genomic region containing the MHC genes was characterized in detail (MHCI: 12 functional paralogs and MHCII-DRB: 3 functional paralogs) [27], and in other felids; e.g. three paralogs were also reported for the MHCII-DRB in cheetah and leopard [30, 31]. Surprisingly, Wang et al. [29] reported from 1 to 6 alleles per *L. lynx* individual in MHCII-DRB suggesting allelic dropout or copy number variation. However, by applying improved amplification and validation methods we found no evidence of copy number variation at MHCII-DRB in any lynx species, as we consistently found from 3 to 6 alleles per individual. Lower than three allele per individual is thus most likely the result of alleles being missed by their genotyping approach based on degenerate primer amplification, cloning and Sanger sequencing [47, 49].

Patterns of TSP and their possible origin

The four species in the *Lynx* genus showed a high level of shared MHC polymorphisms across species, a feature in

common with other recently diverged taxa [28, 29, 50, 51]. As Wei et al. [28] found for several felids in MHCII-DRB, we also detected instances of TSP between lynxes and other distant felids for both MHCI and MHCII-DRB (Additional file 1: Tables S8-S13). TSP across such large divergence times are unlikely due only to neutral processes and are thus more likely the consequence of long-term balancing selection or convergent evolution (i.e., balanced TSP). However, in the case of recently diverged species it is more parsimoniously explained by incomplete lineage sorting during speciation and/or posterior introgression events (i.e., neutral TSP) [52]. Both processes have been described for the *Lynx* genus. A discordant phylogenetic pattern for mitochondrial DNA was found whereby *L. canadensis* appeared as *L. pardinus* sister species, a signal of incomplete lineage sorting, and pervasive signals of introgression between the *L. canadensis* and *L. lynx* and between *L. pardinus* and *L. lynx* have been reported [34, 35]. A more detailed study, taking into account broader genomic regions and ideally MHC alleles assigned to loci, would be required to tackle whether the origin of each trans-species polymorphism in lynx is due to introgression, incomplete lineage sorting, convergent evolution or mere lack of genetic divergence caused by recent species divergence times.

Evidence of natural selection

Both MHCI and MHCII-DRB alleles showed signals of purifying selection at non-ABS, stronger in MHCII-DRB than in MHCI. However, we did not find signatures of long-term positive selection in the form of an increased d_N/d_S ratio at ABS for any MHC family. Contrastingly, such positive selection signals were previously described in analyses of MHCI in cheetah [30], bengal tiger [32] and leopard [31], and for MHCII-DRB in eight different Felidae lineages [28], cheetah [30], but not in bengal tiger [32], leopard [31] or Eurasian lynx [29]. Interestingly, for MHCI we found d_S at the ABS to be much larger than d_S at non-ABS, this indicates that the ABS regions are older than non-ABS and points to positive selection acting in presence of recurrent gene conversion events [45].

In the site-by-site selection test, the amino acids found to be under positive selection in each class had previously been reported in felids (i.e., the 77th amino acid of the MHCI was described as one of the highly polymorphic positions located at the α_1 by Yuhki et al. [24, 25] and the 86th amino acid of the MHCII-DRB was found to be under positive selection by Wei et al. [28] when comparing different Felidae species). It must be noted that the pooling of alleles across loci will generally reduce the power of selection detection tests, so ours and others results might be too conservative, but not prone to false positives [51]. The fact that we don't find the Iberian lynx

to be enriched for pseudogenes also suggests that the relaxation of purifying selection observed across the genome has not affected MHC variation in this way [35].

Genetic drift and balancing selection

One of the main objectives of this study was assessing the possible role of balancing selection in maintaining genetic diversity in declining species by comparing allele repertoires in the two sister lynx species with contrasting demographic histories and overall genetic diversity. We can use the combination of alleles observed in Iberian and Eurasian lynx as a proxy for the ancestral variation in the Iberian lynx. This seems reasonable given the recent divergence and posterior gene flow, and is somehow supported by the nearly complete overlap of allele repertoires between these two species (also between these and Canada lynx) (Fig. 3). Loss of diversity at MHC in bottlenecked populations is a common observation, and suggests that genetic drift is able to override any ongoing balancing selection in most scenarios [19–22]. However, the retention of allelic and nucleotide diversity at MHC exons is quite striking given the extreme overall genetic erosion reported in the Iberian lynx genome, with over 80% of the coding sequences showing no variation at all and overall SNP density being at most one third of that of Eurasian lynx [35]. Furthermore, despite the loss of allelic diversity, Iberian lynx conserved a similar number of segregating sites and maintained high average levels of divergence between alleles, with most supertypes being represented by at least one allele. In fact, the retention of variation—especially of a non-random highly-divergent subset of alleles—following a bottleneck has been observed in several other studies (e.g., [53–56]). In these cases, this has been interpreted as the consequence of balancing selection favouring the retention of more divergent vs. less divergent alleles (divergent allele selection [14]).

An often overlooked fact is that MHC are multigene families with rather complex evolution, and the outcome of increased drift imposed by a bottleneck could depend on the evolutionary relationships of alleles across paralogs. In the absence of gene conversion, which shuffles and homogenizes alleles both within and across loci, the duplicated gene copies would evolve reciprocally monophyletic alleles sets, so that the action of drift would necessarily result in the retention of divergent alleles even if all copies become fixed for one allele, as at least one representative of each divergent lineage would be kept in each locus. It is thus not parsimonious to invoke balancing selection without first discarding this pre-existing pattern (as suggested by Van Oosterhout [57] for Ellison et al. data [58]). Although, as most other studies in non-model species, we could not attribute alleles to particular loci, several lines of evidence suggest that

gene-conversion and recombination are pervasive in the MHC region, preventing the formation of reciprocal monophyly across loci. First, in domestic cat allelic lineages do not correspond to loci, as all alleles belonged to the same phylogenetic lineage in some individuals [26]. Second, in humans MHC I introns have been homogenized relative to exons, as expected when gene conversion is present but balancing selection operates to maintain diversity in the latter but not in the former [59]. Third, the “patchwork pattern” found in the MHC exons is explained by gene conversion mediated minor segmental exchange [24, 25, 60, 61]. Fourth, the highly elevated $d_{S,ABS}/d_{S,non-ABS}$ ratio found here at MHC I in lynx and elsewhere in other species [30] indicates that gene-conversion occurred recurrently in evolutionary times, at least following locus duplication [45].

If recombination and gene-conversion events indeed have shaped the lynxes MHC region up to a point in which allelic lineages are no longer monophyletic by locus, the retention of divergent alleles may be taken as evidence to support the action of balancing selection. Balancing selection could be operating within and also between loci, allowing the extension of the target of balancing selection to multilocus MHC haplotypes [14, 62]. In this regard, MHC haplotypes (i.e., allele combinations at several tightly linked loci) could thus be selected to harbour a set of highly divergent (and thus functionally complementary) alleles giving the individuals with such haplotypes a “permanent heterozygote” advantage [14, 62–64]. Such process will be especially relevant for small or declining populations where due to low within locus allelic diversity and heterozygosity, interlocus diversity will become the main source of intraindividual functional variation.

Disentangling which mechanisms are the main drivers of the retention of functional variation in the Iberian lynx will require data at the population and individual levels, as well as locus-specific genotypes. Such study would expand our empirical knowledge on how the evolutionary forces shape the genomic landscape of endangered species.

Generation of novel functional variation

The Iberian lynx allelic repertoire was mostly a subset of the Eurasian lynx alleles. However, it included a pair of Iberian-private alleles (Fig. 3). One of these private alleles (*Lypa-DRB*6*) appears to have been originated by recombination of alleles within the species repertoire while the other (*Lypa-DRB*7*) derived from it and contained a single non-synonymous point mutation. Although we cannot discard the possibility of these alleles being ancestral to both species and being lost in Eurasian, an alternative explanation is that it was originated within the Iberian lynx lineage. In fact, these two

alleles define a novel and exclusive lineage and super-type, not represented in any of the other felids analysed. Indeed, recombination and gene conversion have long been described as a motor of rapid evolution compared to point mutation in multigene families [65, 66], with special mention to the MHC families [45, 67]. Short exonic fragments passed to different alleles by gene-conversion or interchanged by reciprocal recombination events would increase the number of alleles faster compared to point mutation acting alone, especially in the presence of interlocus recombination [67]. Only recently, such processes have been invoked as a process by which genetically depauperated populations could recover some of the variability lost through genetic erosion [68]. Additionally, the non-synonymous point mutation that created *Lypa-DRB*7* from *Lypa-DRB*6* originated a similar but functionally distinct allele, suggesting that mutation did also contribute, although to a lesser extent, to create new functional variation (Fig. 3, Additional file 1: Tables S2 and S13). This amino acid change is at a non-ABS site though, so we cannot gauge whether it has increased its frequency because of positive selection or instead by the relaxation of purifying selection. Non-ABS sites have clear signals of long-term purifying selection (Table 1), but the Iberian lynx genome contains pervasive signals of recent relaxation, as reflected in excess of non-synonymous variants and substitutions [35]. The fact that these new alleles were able to attain allelic frequencies high enough to persist in the species despite the series of demographic bottlenecks it has suffered since its speciation [35] points to some form of positive or balancing selection [13, 14].

Iberian lynx immunity and population viability

Irrespective of the evolutionary mechanisms involved, we showed retention of substantial functional variation at MHC despite severe overall genomic erosion in the Iberian lynx. These are good news for the conservation of Iberian lynx. Notwithstanding, several studies suggest that the immune system of the Iberian lynx is compromised and have alerted on its consequences for population viability. A recent increase in disease-associated mortality rates for the Doñana population and the species as a whole has been reported [69, 70]. In particular, a feline leukaemia virus (FeLV) outbreak in 2007 killed more than half of the twelve infected animals in less than six months, leaving the main reproductive nucleus in Doñana without males [71–73]. A latter study confirmed that the strain responsible of the outbreak was not specially virulent [74], as it did not exhibit genetic variants known to confer increased virulence and it did not severely affect cats in a *in vivo* transmission study, highlighting the special susceptibility of Iberian lynxes to infectious diseases. Based on that apparent deficiency of

immunity in Iberian lynxes and the frequent detection of infectious agents among sympatric carnivores [75], a disease outbreak has been considered a serious risk for Iberian lynx persistence [71, 75, 76].

We should not forget that the diversity measures presented in this study are referred to the species level and may not properly reflect the diversity at population or individual levels, which are the levels at which natural selection act. If the repertoire of alleles of a species is sorted out in different sets corresponding to distinct populations, the levels of each population's diversity will be low- even when the species diversity is high. Likewise, if the allelic frequencies within a population are very skewed, the levels of individual variation will be low- even when the population allelic richness is high. Hence, we need to extend our characterization of MHC variation to Iberian lynx populations and individuals before we can fully evaluate its impact on immune function and ultimately on disease susceptibility and population viability. Furthermore, the immune response is a complex physiological process involving many genes, among which MHC are an important but limited part [77]. Thus, a more robust evaluation of immunocompetence based on genetic information would require the assessment of the whole immunome.

Conclusions

The fate of functional variation following a bottleneck could be to follow the neutral variation dynamics towards depletion, or to be, at least partially, maintained by balancing selection. Based on MHC data from the *Lynx* genus, we argue that functional genetic variation can be retained even in the face of extreme genetic erosion through the functional redundancy provided by multigene families and the possible contribution of different forms of balancing selection. If generalized to other species and functional components, especially those based on multigene families, it would mean that endangered species might retain more adaptive capacity than we would predict from diversity at neutral markers.

Methods

Sampling

Our sampling covered the whole distribution of *L. pardinus* and a remarkable part of the distribution of *L. lynx*. We used 102 samples of the two remnant populations of Iberian lynxes (Doñana ($N = 48$) and Sierra Morena ($N = 54$)), and 99 samples of six populations of Eurasian lynxes (Estonia ($N = 15$), Latvia ($N = 15$), Norway ($N = 11$), Poland ($N = 20$), Russia-Kirov ($N = 14$), and Russia-Vladivostok ($N = 24$)) (Fig. 1). Admittedly, some of these Eurasian lynx populations went through bottlenecks and show reduced genetic diversity [39] but this will not bias our species level estimates. Given that

our aim here is to get the most complete allelic repertoire for each species, the inclusion of bottlenecked populations can only add some new alleles or not, but will not artificially reduce the allelic repertoire of the species. *L. canadensis* and *L. rufus* are represented for comparative purposes by two individuals per species.

Primer design and amplification strategy

Two sets of primers were designed to amplify the second exon of both MHC class I and MHC class II-DRB genes. Primer design used for MHCI amplification is described elsewhere [49] and MHCII-DRB design followed the same rationale. In short, primer design was based on a set of variants that included all MHC class I or class II-DRB exon 2 alleles for species closely-related to lynxes available in GenBank as of May of 2011 (i.e., all Felidae alleles.) and the annotated region of the cat [27]. Also the transcripts from two different Iberian lynxes annotated as MHCI or MHCII-DRB by the Iberian lynx genome project [35] were included. Primers were designed in approximately the same regions as the ones used in previous studies on MHC variation in the Felidae to enable comparisons [28, 30–32]. More specifically, MHCI primers spanned bases 2 to 21 (forward) and 253 to 271 (reverse) of the human MHC class I exon 2 and MHC class II-DRB spanned bases 4 to 23 (forward) and 270 to 288 (reverse) of the human homologues [78–80].

For MHCI we used a PCR amplification strategy we called Pooled-primers designed to maximize the number of detected alleles by minimizing amplification bias among alleles [49]. On the contrary, given that the MHCII-DRB required less degenerated primers, we used a standard PCR using a single pair of degenerate primers: Fel_MhcII_ex2_F [5'-TGTBYCCACAGCACATTTTCY-3'] and Fel_MhcII_ex2_R [5'-CTCAMCTCGCCGSGTGCAC-3'].

Amplification and sequencing of MHC loci

Genomic DNA was extracted from blood or tissue using standard phenol-chloroform methods [81]. To produce sequencing libraries, we used the Universal Tailed Amplicon Sequencing design by Roche consisting of a two-round PCR approach in which the first round amplifies the target loci using primers with a universal 5' extension and the second adds 454 sequencing adapters and a sequence tag to the amplicons generated in the first PCR to unambiguously identify each sample. Artefact formation during PCR was minimized by using the Phusion High-Fidelity PCR Kit by Roche, which reduces nucleotide misincorporation [82], and by implementing long extension times and no final extension step, which prevents chimera formation. PCRs were run in a final volume of 10 μ l following manufacturer indications. Cycling conditions were the same for all PCRs: an initial denaturation at 98° for 30 s and 25 cycles of 98 °C for

10 s, 57 °C for 30 s and 72 °C for 2 min. PCR products were equimolarly pooled and sequenced on a 454 GS Junior system.

Data processing and allele validation

High throughput reads were sorted by individual, quality filtered, collapsed into unique sequences and assigned to alleles or artefacts following Sommer et al. 2013 (refer to [83] for further details). This protocol assumes different amplification efficiencies for different alleles and does not assume that reads representing true alleles are more abundant in the read pool than any artefact. These two aspects are important when primers are not guaranteed to amplify alleles with the same efficiency. The sole modification introduced to the method was that we did not systematically replicate each individual in the case of *L. lynx* and *L. pardinus*. Due to the low genetic variation of the species [35, 42], and the use of several individuals belonging to the same population, we expected every allele to be found in more than one individual. Nevertheless, a minimum of 10% of the samples were replicated as quality control. Alleles were named following Klein et al. [84]. To be considered as functional and taken into account in downstream analysis an allele had to lack signs of pseudogenization (i.e., they should not present indels, reading frame shifts or premature stop codons). All Iberian lynx alleles without pseudogenization signals, were confirmed to be transcribed by RNA-Seq or RT-PCR amplicon (data not shown); no similar data was available for non-Iberian lynx alleles.

Data analysis

Recombination and phylogenetic relationships of the alleles

Single Break Point (SBP) and Genetic Algorithm Recombination Detection (GARD) tests [85] were used to identify recombination breakpoints. Such breakpoints and the phylogenetic trees produced as output were used to inform subsequent site-by-site selection tests. To evaluate relationships among sequences, we constructed a phylogenetic network using SplitsTree4 [86, 87]. We used the Neighbor-Net method and Jukes-Cantor distances. Contrary to traditional phylogenetic trees, phylogenetic networks enable visualization of conflicting signals from processes such as gene conversion and recombination [87].

Alleles found in the *Lynx* genus were compared among them and to all the Felidae alleles already available in GenBank to detect instances of functional variation shared across the genus and across the Felidae family. For this purpose we first looked for shared alleles at both nucleotide and amino acid levels. Even though phylogenetic trees are inappropriate to represent phylogenetic relationships of alleles that suffered recombination or gene-conversion, we used them to explore how lynx

alleles are related to others in the Felidae family. For MHCII-DRB, for which previous allele lineages have been described [26, 28], we recorded which of those lineages are represented in the *Lynx* genus. Trees were constructed in the Geneious suite [88] using RAxML [89] under the GTR GAMMA I nucleotide model and 1000 bootstraps. Consensus tree was calculated using the suite's consensus tree builder tool.

The *Lynx* genus showed shared polymorphisms also with other felids, as we expected both alleles phylogenetic trees and the supertypes clustering show the alleles of different species clustering together rather than clustering them by species (Additional file 2: Figures S3-S6). We even found lynx alleles to be completely identical at the nucleotide and/or amino acid level to alleles present in other felid species. For the MHCI, one allele was shared between *L. rufus* and *F. catus* at the nucleotide level (*FLAI-O = Lyru-MHCI*11*, see Additional file 1: Table S8), while two sets of alleles were shared at the protein level: one including all the lynxes, cat and tiger and the second one with all the lynxes and cat only (Additional file 1: Table S9). At the supertype level, 12 supertypes contained lynx alleles and 11 of them also contained alleles from other felid species (Additional file 1: Table S10). No MHCII-DRB allele was shared between lynxes and the rest of the felids at a nucleotide level (Additional file 1: Table S11) and only one pair was shared between *Lynx lynx* and *Neofelis nebulosa* at the protein level (Nene-DRB*202 = Lyly-DRB*09, Additional file 1: Table S12). Such alleles shared across distant species and showing no or very low divergence, pose a typical example of TSP and are likely to have been subject of balancing selection. When compared to the five MHCII-DRB lineages described by Yuhki and O'Brien [26] we found most lynx alleles to cluster in lineages 2 and 5, as found in Wei et al. [28] for *L. lynx* alleles (Additional file 2: Figure S4). Likewise, we found no lynx alleles in lineages 4 and 1. However, we did not find many lynx alleles on an uncertain lineage, as reported by Wei et al. [28], but those alleles grouped instead in lineage 5, and *Pati-DRB*1* (which was the only non-lynx allele on such uncertain clade) on the base of the tree. Additionally, we found a bobcat allele clustering in lineage 3, where no lynx alleles had been described before.

Signatures of selection

Antigen binding sites (ABS) codon positions were inferred from human MHC class I [79, 80] and from MHCII-DRB [78] homologs. Additionally, we checked that the positions attributed to the ABS were equivalent to those reported in other felids [30, 31]. Positive and purifying selection at putative ABS and non-ABS codons were inferred by an unbiased estimation of the rates of nonsynonymous (d_N) and synonymous (d_S) substitution [90], calculated by the *kaks* function in the *Seqinr* R package [91]. The significance of the deviation above or

below zero of the d_N/d_S ratio was calculated using a Mann-Whitney U-test.

To infer site-by-site positive and negative selection in our alleles, we ran the four tests available at *datamonkey* web server [92] (i.e., SLAC,REL, FEL [46] and FURBAR [93]). For all analyses, we used the model of evolution that best fitted the data and the recombination information when the alignment spanned recombination breakpoints. Finally, we used the Integrative Selection Results tool to compare the results.

Functional diversity retention: Genetic diversity and supertypes

All analysis were performed for the whole dataset of alleles found across species as well as for each species on its own using different R packages. We used *ape* [94] to calculate the number of segregating sites as well as the number of nucleotide differences and *pegas* [95] to calculate nucleotide diversity. Nucleotide genetic distances were estimated using *ape* while *Phangorn* [96] was used to estimate amino acid distances for the whole sequence, for ABS and for non-ABS. The evolutionary model to be used in each case was inferred using *Model Selection* package implemented in the web server hosted at <http://www.datamonkey.org> [92].

Second, to evaluate the retention of functional variation beyond TSP, we grouped alleles into supertypes based on functional similarity [97]. We used a pocket-based approach in which only the ABS are taken into account to group the alleles. Each ABS amino acid was transcribed into five physicochemical relevant descriptor variables: z1 (hydrophobicity), z2 (steric bulk), z3 (polarity), z4 and z5 (electronic effects) [98]. Later, hierarchical clustering with euclidean distance was applied to the resulting allele matrix using *dist* and *hclust* functions from the *stats* R package to group alleles based on their functional (antigen-binding) properties.

Additional files

Additional file 1: Table S1. Nucleotide sequence alignment of the second exon of MHC class I alleles observed in *Lynx*. Numbers indicate the nucleotide positions inferred from human MHC class I loci (Bjorkman et al. 1987; Bjorkman & Parham 1990). Dots indicate identity to allele *Lyra_MHCI*10/Lyly_MHCI*10*, which is used as reference. **Table S2.** Nucleotide sequence alignment of the second exon of MHC class II-DRB alleles observed in *Lynx*. Alleles are aligned to *Lyra_DRB*1/Lyly_DRB*1/Lyca_DRB*1/Lyru_DRB*1*. Numbers indicate the nucleotide positions of exon 2 inferred from the human sequence (Brown et al. 1993). Dots indicate identity to the top sequence. **Table S3.** *Lynx* MHCI pseudogenization signatures. For each allele we show whether it contains insertion and/or deletions (along with the inserted base and its position), whether it shows an Open Reading Frame (ORF), and if it was considered as a pseudogene in this study. **Table S4.** Integrative MHCI positive and negative selection tests. d_N/d_S ratios for codon showing signals of selection in any of the test performed (SLAC, FEL, REL, and FURBAR) are reported along with their significance values (either *p*-value or Bayes factor), whether they were significant for each test (+) or not (-), and if they were considered in this study (Yes/No), with those considered also shaded in light blue. **Table S5** Amino

acid sequence alignment of the second exon of MHC class I alleles observed in *Lynx*. Alleles are aligned to *Lypa_DRB*10/Lyly_DRB*10*. Numbers indicate the amino acid positions and asterisks putative ABS inferred from human MHC class I loci (Bjorkman et al. 1987; Bjorkman & Parham 1990). Dots indicate identity to the top sequence. The amino acids inferred to be under positive or negative selection are marked with a + or a - sign, respectively. **Table S6.** Integrative MHCII-DRB positive and negative selection tests. dN-dS ratios for codon showing signals of selection in any of the test performed (SLAC, FEL, REL, and FURBAR) are reported along with their significance values (either *p*-value or Bayes factor), whether they were significant for each test (+) or not (-), and if they were considered in this study (Yes/No), with those considered also shaded in light blue. **Table S7.** Amino acid sequence alignment of the second exon of MHC class II-DRB alleles observed in *lynxes*. Alleles are aligned to *Lypa_DRB*1/Lyly_DRB*1/Lyca_DRB*1/Lyru_DRB*1*. Numbers indicate the amino acid positions and asterisks putative ABS inferred from human MHC class I loci (Brown et al. 1993). Dots indicate identity to the top sequence. The amino acids inferred to be under positive or negative selection are marked with a + or a - sign, respectively. **Table S8.** Unique MHC I nucleotide alleles found across Felidae. Each Felidae allele were collapsed into unique nucleotide sequences and the instances of identical sequences being shared by different species were identified. For each shared sequence we show an identification number (Seq), the number of species in which it appears (#Species), the acronym of such species (Species), whether it is present in any *Lynx* species (*Lynx*), whether it is shared by any *Lynx* species and one non-*Lynx* species (*Lynx_Other*), the number of alleles collapsed (#Alleles), and the alleles names (Alleles names). **Table S9.** Unique MHC I amino acid alleles found across Felidae. Each Felidae allele were collapsed into unique amino acids sequences and the instances of identical sequences being shared by different species were identified. For each shared sequence we show an identification number (Seq), the number of species in which it appears (#Species), the acronym of such species (Species), whether it is present in any *Lynx* species (*Lynx*), whether it is shared by any *Lynx* species and one non-*Lynx* species (*Lynx_Other*), the number of alleles collapsed (#Alleles), and the alleles names (Alleles names). **Table S10.** Felidae MHC I supertype definition based on hierarchical clustering. Each Felidae allele was clustered following a pocket based approach that considered only ABS. For each allele we show its name (Allele) and the supertype it was assigned to within the Felidae analysis (Supertype). *Lynx* alleles are shown in bold. For each supertype we report whether it contains at least one *Lynx* allele, and if so we indicate the supertype cluster they correspond to in Fig. 3a (Supertype present in *Lynx* genus) and whether this was exclusive to *Lynx* or shared with other felid species (Supertype shared with other felids). **Table S11.** Unique MHCII-DRB nucleotide alleles found across Felidae. Each Felidae allele were collapsed into unique nucleotide sequences and the instances of identical sequences being shared by different species were identified. For each shared sequence we show an identification number (Seq), the number of species in which it appears (#Species), the acronym of such species (Species), whether it is present in any *Lynx* species (*Lynx*), whether it is shared by any *Lynx* species and one non-*Lynx* species (*Lynx_Other*), the number of alleles collapsed (#Alleles), and the alleles names (Alleles names). **Table S12.** Unique MHCII-DRB amino acid alleles found across Felidae. Each Felidae allele were collapsed into unique amino acids sequences and the instances of identical sequences being shared by different species were identified. For each shared sequence we show an identification number (Seq), the number of species in which it appears (#Species), the acronym of such species (Species), whether it is present in any *Lynx* species (*Lynx*), whether it is shared by any *Lynx* species and one non-*Lynx* species (*Lynx_Other*), the number of alleles collapsed (#Alleles), and the alleles names (Alleles names). **Table S13.** Felidae MHCII-DRB supertype definition based on hierarchical clustering. Each Felidae allele was clustered following a pocket based approach that considered only ABS. For each allele we show its name (Allele), the supertype it was assigned to (Supertype). *Lynx* alleles are shown in bold. For each supertype we report whether it contains at least one *Lynx* allele, and if so we indicate the supertype cluster they correspond to in Fig. 3a (Supertype present in *Lynx* genus) and whether this was exclusive to *Lynx* or shared with other felid species (Supertype shared with other felids). (XLS 350 kb)

Additional file 2: Figure S1. Phylogenetic network of *Lynx* MHC I alleles. We constructed a network with the Neighbor-Net method and Jukes-Cantor distances in SplitsTree4. Alleles are colored by species of

origin. **Figure S2.** Phylogenetic network of *Lynx* MHCII-DRB alleles. We constructed a network with the Neighbor-Net method and Jukes-Cantor distances in SplitsTree4. Alleles are colored by species of origin. **Figure S3.** Phylogenetic tree of Felidae MHC I alleles. The tree was constructed with RaxML. Tips are colored by genus of origin. **Figure S4.** Phylogenetic tree of Felidae MHCII-DRB alleles. The tree was constructed with RaxML. Tips are colored by genus of origin. **Figure S5.** Felidae MHC I supertype definition. We used average hierarchical clustering and a cut-off threshold of Euclidean distance ≥ 15 to classify alleles into distinctive functional groups (supertypes, red boxes) based on functional similarity at their ABS. Alleles with identical amino acids at ABS are represented as tip polytomies at 0 Euclidean distance. **Figure S6.** Felidae MHCII-DRB supertype definition. We used average hierarchical clustering and a cut-off threshold of Euclidean distance ≥ 15 to classify alleles into distinctive functional groups (supertypes, red boxes) based on functional similarity at their ABS. Alleles with identical amino acids at ABS are represented as polytomies at 0 euclidean distance. (PDF 408 kb)

Acknowledgements

Thanks to the Junta de Andalucía, the Iberian lynx Captive Breeding Program and the LIFE projects LIFE-02NAT/8609, LIFE 06NAT/E/209, LIFE 10NAT/ES/570 for kindly providing Iberian lynx samples used in this study. We are grateful to Janis Ozolins, Peep Männil, and Ingrid Reinkind for kind sharing the Eurasian lynx samples. Logistical support was provided by Laboratorio de Ecología Molecular (LEM-EBD) and by Laboratorio de Sistemas de Información Geográfica y Teledetección (LAST-EBD) at the Estación Biológica de Doñana, CSIC. Thanks to Laura Soriano, María Mendez, José María Gassent and Ana Píriz for support with laboratory work and Isabel Afán and David Aragonés for their help with the map. We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI). We are grateful to Janis Ozolins, Peep Männil and Ingrid Reinkind for kind providing the Eurasian lynx samples from Latvia, Estonia and Norway, respectively.

Funding

Funding for this project was provided by the Spanish Dirección General de Investigación Científica y Técnica (CGL2010-21540/BOS and CGL2013-47755-P), project "Adaptive variation in declining species: Survey of MHC variation in Eurasian lynx populations at the western edge of its range" funded by the internal EBD proposal call "Microproyectos" financed by the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Program for Centres of Excellence in R + D + I (SEV-2012-0262), and project 2014/15/B/NZ8/00212 funded by the National Science Center, Poland. Elena Marmesat received a JAE predoctoral grant from CSIC (Spanish National Research Council).

Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files. Additionally, the datasets supporting the conclusions of this article are available in:

-Raw reads: the ENA repository under project accession PRJEB20088.

-*Lynx* MHC alleles: Genbank repository, Genbank accessions MHC I (KY769287 to KY769350) and MHCII-DRB (KY769351 to KY769367).

Authors' contributions

EM designed the study, did the lab work, collected the data, performed NGS data processing and analyses and led the writing of the manuscript. KS designed the *L. lynx* sampling and collected samples, participated in the study design and critically reviewed the manuscript. MR provided valuable input in data analyses and interpretation and critically reviewed the manuscript. APS and IVS collected part of the Eurasian lynx samples. JAG conceived, designed and supervised the study and contributed to the interpretation of the data and the writing of the manuscript. All authors drafted and approved the final manuscript.

Ethics approval and consent to participate

Iberian lynx samples were donated to the EBD group and were collected by the veterinary staff at Centro de Análisis y Diagnóstico de la Fauna Silvestre-CAD in the course of necropsies of dead animals, by the resident veterinary at Iberian lynx Breeding centers during regular health revisions of captive animals, or by

veterinary staff of LIFE10NAT/ES/570 from animals captured for management and monitoring. The procedures were licensed by the corresponding local Competent Authority to comply with Spanish legislation on the protection of animals used for scientific purposes. 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 from Russia and Norway was licensed by CITES permissions no 12RU000512 and 15NO-046-EX, respectively.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Integrative Ecology, Estación Biológica de Doñana (CSIC), C/ Américo Vespucio, 26, 41092 Sevilla, Spain. ²Mammal Research Institute, Polish Academy of Sciences, 17-230 Białowieża, Poland. ³Department of Animal Ecology, Russian Research Institute of Game Management and Fur Farming, 79 Preobrazhenskaya Str, Kirov 610000, Russia. ⁴Laboratory of Ecology and Conservation of Animals, Pacific Institute of Geography of Far East Branch of Russian Academy of Sciences, 7 Radio Street, Vladivostok 690041, Russia. ⁵Far Eastern Federal University, 8 Sukhanova Street, Vladivostok 690091, Russia.

Received: 10 April 2017 Accepted: 23 June 2017

Published online: 04 July 2017

References

- Allendorf FW, Luikart G, Aitken SN. Conservation and the Genetics of Populations. 2nd edn. Oxford: Wiley-Blackwell; 2013: XVIII, 610 p.
- Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *PNAS*. 2004;101:3490–4.
- Väli Ü, Einarsson A, Waits L, Ellegren H. To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Mol Ecol*. 2008;17:3808–17.
- Joop Ouborg N, Angeloni F, Vergeer P. An essay on the necessity and feasibility of conservation genomics. *Conserv Genet*. 2010;11:643–53.
- Primmer CR. From conservation genetics to conservation genomics. *Ann N Y Acad Sci*. 2009;1162:357–68.
- Allendorf FW, Hohenlohe P. A, Luikart G. Genomics and the future of conservation genetics. *Nat. Rev. Genet. Nat Publ Group*. 2010;11:697–709.
- Steiner CC, Putnam AS, Hoek PEA, Ryder OA. Conservation genomics of threatened animal species. *Annu Rev Anim Biosci*. 2013;1:261–81.
- Smith KF, Sax DF, Lafferty KD. Evidence for the role of infectious disease in species extinction and endangerment. *Conserv Biol*. 2006;20:1349–57.
- McCallum H. Disease and the dynamics of extinction. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. 2012;367:2828–39.
- Klein J. Natural history of the major histocompatibility complex. New York: Wiley; 1986.
- Piertney SB, Oliver MK. The evolutionary ecology of the major histocompatibility complex. *Heredity*. 2006;96:7–21.
- Sommer S. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Front Zool*. 2005;2:16.
- Spurgin LG, Richardson DS. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proc Biol Sci*. 2010;277:979–88.
- Wakeland E, Boehme S, She J, Lu C, McIndoe R, Cheng I, et al. Ancestral polymorphisms of MHC class-II genes: divergent allele advantage. *Immunol Res*. 1990;9:115–22.
- Klein J, Sato A, Nikolaidis N. MHC, TSP, and the origin of species: from immunogenetics to evolutionary genetics. *Annu Rev Genet*. 2007;41:281–304.
- Wenink PW, Groen AF, Roelke-Parker ME, Prins HH. African buffalo maintain high genetic diversity in the major histocompatibility complex in spite of historically known population bottlenecks. *Mol Ecol*. 1998;7:1315–22.
- Jarvi SI, Tarr CL, McIntosh CE, Atkinson CT, Fleischer RC. Natural selection of the major histocompatibility complex (Mhc) in Hawaiian honeycreepers (Drepanidinae). *Mol Ecol*. 2004;13:2157–68.
- Sachdev M, Sankaranarayanan R, Reddanna P, Thangaraj K, Singh L. Major histocompatibility complex class I polymorphism in Asiatic lions. *Tissue Antigens*. 2005;66:9–18.
- O'Brien SJ, Wildt DE, Goldman D, Merrill CR, Bush M. The cheetah is depauperate in genetic variation. *Science*. 1983;221:459–62.
- Weber DS, Van Coeverden De Groot PJ, Peacock E, Schrenzel MD, Perez DA, Thomas S, et al. Low MHC variation in the polar bear: implications in the face of Arctic warming? *Anim Conserv*. 2013;16:671–83.
- Mainguy J, Worley K, Côté SD, Coltman DW. Low MHC DRB class II diversity in the mountain goat: past bottlenecks and possible role of pathogens and parasites. *Conserv Genet*. 2007;8:885–91.
- Miller HC, Lambert DM. Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). *Mol Ecol*. 2004;13:3709–21.
- Doytchinova IA, Guan P, Flower DR. Identifying human MHC supertypes using bioinformatic methods. *J Immunol*. 2004;172:4314–23.
- Yuhki N, O'Brien SJ. DNA recombination and natural selection pressure sustain genetic sequence diversity of the feline MHC class I genes. *J Exp Med*. 1990;172:621–30.
- Yuhki N, O'Brien SJ. Exchanges of short polymorphic DNA segments predating speciation in feline major histocompatibility complex class I genes. *J Mol Evol*. 1994;39:22–33.
- Yuhki N, O'Brien SJ. Nature and origin of polymorphism in feline MHC class II DRA and DRB genes. *J Immunol*. 1997;158:2822–33.
- Yuhki N, Mullikin JC, Beck T, Stephens R, O'Brien SJ. Sequences, annotation and single nucleotide polymorphism of the major histocompatibility complex in the domestic cat. *PLoS One*. 2008;3:e2674.
- Wei K, Zhang Z, Wang X, Zhang W, Xu X, Shen F, et al. Lineage pattern, trans-species polymorphism, and selection pressure among the major lineages of feline Mhc-DRB peptide-binding region. *Immunogenetics*. 2010;62:307–17.
- Wang X, Wei K, Zhang Z, Xu X, Zhang W, Shen F, et al. Major histocompatibility complex class II DRB exon-2 diversity of the Eurasian lynx (*Lynx lynx*) in China. *J Nat Hist Taylor & Francis*. 2009;43:245–57.
- Castro-Prieto A, Wachter B, Sommer S. Cheetah paradigm revisited: MHC diversity in the World's largest free-ranging population. *Mol Biol Evol*. 2011;28:1455–68.
- Castro-Prieto A, Wachter B, Melzheimer J, Thalwitzer S, Sommer S. Diversity and evolutionary patterns of immune genes in free-ranging Namibian leopards (*Panthera pardus pardus*). *J Hered*. 2011;102:653–65.
- Pokorny I, Sharma R, Goyal SP, Mishra S, Tiedemann R. MHC class I and MHC class II DRB gene variability in wild and captive Bengal tigers (*Panthera tigris tigris*). *Immunogenetics*. 2010;62:667–79.
- Johnson WE, Eizirik E, Pecon-Slattery J, Murphy WJ, Antunes A, Teeling E, et al. The late Miocene radiation of modern Felidae: a genetic assessment. *Science American Association for the Advancement of Science*. 2006;311:73–7.
- Li G, Davis BW, Eizirik E, Murphy WJ. Phylogenomic evidence for ancient hybridization in the genomes of living cats (Felidae). *Genome Res*. 2016;26:1–11.
- Abascal F, Corvelo A, Cruz F, Villanueva-Cañas JL, Vlasova A, Marcet-Houben M, et al. Extreme genomic erosion after recurrent demographic bottlenecks in the highly endangered Iberian lynx. *Genome Biol Genome Biology*. 2016;17:251.
- Von Arx M, Breitenmoser-Würsten C. *Lynx pardinus*. The IUCN Red List of Threatened Species 2008. Version 2008 edn: IUCN; 2008.
- Rodríguez A, Calzada J. *Lynx pardinus*. The IUCN Red List of Threatened Species 2015. Version 2015.2 edn: IUCN; 2015.
- Breitenmoser U, Breitenmoser-Würsten C, Lanz T, von Arx M, Antonevich A, Bao W, Avgan B. *Lynx lynx*. The IUCN Red List of Threatened Species 2015. Version 2015.2 edn: IUCN; 2015.
- Ratkiewicz M, Matosiuk M, Kowalczyk R, Konopiński MK, Okarma H, Ozolins J, et al. High levels of population differentiation in Eurasian lynx at the edge of the species' western range in Europe revealed by mitochondrial DNA analyses. *Anim Conserv*. 2012;15:603–12. Katzner T, Davies J, editors
- Kowalczyk R, Gorny M, Schmidt K. Edge effect and influence of economic growth on Eurasian lynx mortality in the Białowieża primeval Forest, Poland. *Mammal Res*. 2015;60:3–8.
- Schmidt K, Ratkiewicz M, Konopiński MK. The importance of genetic variability and population differentiation in the Eurasian *Lynx lynx lynx* for conservation, in the context of habitat and climate change. *Mamm Rev*. 2011;41:112–24.

42. Casas-Marce M, Soriano L, López-Bao JV, Godoy JA. Genetics at the verge of extinction: insights from the Iberian lynx. *Mol Ecol*. 2013;22(22):5503–15.
43. Croteau EK, Heist EJ, Nielsen CK, Hutchinson JR, Hellgren EC. Microsatellites and mitochondrial DNA reveal regional population structure in bobcats (*Lynx rufus*) of North America. *Conserv Genet*. 2012;13:1637–51.
44. Schwartz MK, Mills LS, Ortega Y, Ruggiero LF, Allendorf FW. Landscape location affects genetic variation of Canada lynx (*Lynx canadensis*). *Mol Ecol*. 2003;12:1807–16.
45. Ohta T. Gene conversion vs point mutation in generating variability at the antigen recognition site of major histocompatibility complex loci. *J Mol Evol*. 1995;41:115–9.
46. Kosakovsky Pond SL, Frost SDW, Pond SLK, Frost SDW. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol*. 2005;22:1208–22.
47. Babik W. Methods for MHC genotyping in non-model vertebrates. *Mol Ecol Resour*. 2010;10:237–51.
48. Babik W, Taberlet P, Ejsmond MJ, Radwan J. New generation sequencers as a tool for genotyping of highly polymorphic multilocus MHC system. *Mol Ecol Resour*. 2009;9:713–9.
49. Marmesat E, Soriano L, Mazzoni CJ, Sommer S, Godoy JA. PCR strategies for complete allele calling in Multigene families using high-throughput sequencing approaches. *PLoS One*. 2016;11:e0157402.
50. Sato A, Tichy H, Grant PR, Grant BR, Sato T, O'Huigin C. Spectrum of MHC class II variability in Darwin's finches and their close relatives. *Mol Biol Evol*. 2011;28:1943–56.
51. Gillingham MAF, Courtiol A, Teixeira M, Galan M, Bechet A, Cezilly F. Evidence of gene orthology and trans-species polymorphism, but not of parallel evolution, despite high levels of concerted evolution in the major histocompatibility complex of flamingo species. *J Evol Biol*. 2016;29:438–54.
52. Klein J, Sato A, Nagl S, O'Huigin C. MOLECULAR TRANS-SPECIES POLYMORPHISM. *Annu Rev Ecol Syst*. 1998;29:1–21.
53. Hedrick PW. Pathogen resistance and genetic variation at Mhc loci. *Evolution (N Y)*. 2002;56:1902–8.
54. Richardson DS, Westerdahl H. MHC diversity in two *Acrocephalus* species: the outbred great reed warbler and the inbred Seychelles warbler. *Mol Ecol*. 2003;12(12):3523–9.
55. Van Oosterhout C, Joyce DA, Cummings SM, Blais J, Barson NJ, Ramnarine IW, et al. Balancing selection, random genetic drift, and genetic variation at the major histocompatibility complex in two wild populations of guppies (*Poecilia reticulata*). *Evolution*. 2006;60:2562–74.
56. Viček J, Hoecq PEA, Keller LF, Wayhart JP, Dolinová I, Štefka J. Balancing selection and genetic drift create unusual patterns of MHCII β variation in Galápagos mockingbirds. *Mol Ecol*. 2016;25(19):4757–72.
57. van Oosterhout C. Maintenance of major histocompatibility supertype variation in selfing vertebrate is no evidence for overdominant selection. *Proc Biol Sci*. 2013;280:20122501.
58. Ellison A, Allaingillaume J, Girdwood S, Pachebat J, Peat KM, Wright P, et al. Maintaining functional major histocompatibility complex diversity under inbreeding: the case of a selfing vertebrate. *Proc Biol Sci*. 2012;279:5004–13.
59. Hughes AL, Yeager M. Natural selection and the evolutionary history of major histocompatibility complex loci. *Front Biosci*. 1998;3:d509–16.
60. Kappes D, Strominger JL. Human class II major Histocompatibility complex genes and proteins. *Annu Rev Biochem*. 1988;57:991–1028.
61. Lawlor DA, Zemmour J, Ennis PD, Parham P. Evolution of class-I genes and proteins: from natural selection to Thymic selection. *Annu Rev Immunol*. 1990;8:23–63.
62. Innan H, Kondrashov F. The evolution of gene duplications: classifying and distinguishing between models. *Nat Rev Genet*. 2010;11:4.
63. Otto SP, Yong P. The evolution of gene duplicates. *Adv Genet*. 2002;46:451–83.
64. Spofford JB. Heterosis and the evolution of duplications. *Am Nat*. 1969;103:407–32.
65. Basten CJ, Ohta T. Simulation study of a multigene family, with special reference to the evolution of compensatory advantageous mutations. *Genetics*. 1992;132:247–52.
66. Ohta T, Basten CJ. Gene conversion generates hypervariability at the variable regions of kallikreins and their inhibitors. *Mol Phylogenet Evol*. 1992;1:87–90.
67. Ohta T. Role of diversifying selection and gene conversion in evolution of major histocompatibility complex loci. *Proc Natl Acad Sci U S A*. 1991;88:6716–20.
68. Spurgin LG, van Oosterhout C, Illera JC, Bridgett S, Gharbi K, Emerson BC, et al. Gene conversion rapidly generates major histocompatibility complex diversity in recently founded bird populations. *Mol Ecol*. 2011;20:5213–25.
69. Palomares F, Godoy JA, López-Bao JV, Rodríguez A, Roques S, Casas-Marce M, et al. Possible extinction vortex for a population of Iberian lynx on the verge of extirpation. *Conserv Biol*. 2012;26:689–97.
70. López G, López-Parra M, Garrote G, Fernández L, del Rey-Wamba T, Arenas-Rojas R, et al. Evaluating mortality rates and causalities in a critically endangered felid across its whole distribution range. *Eur J Wildl Res*. 2014;60:359–66.
71. Meli ML, Cattori V, Martínez F, López G, Vargas A, Simón MA, et al. Feline leukemia virus and other pathogens as important threats to the survival of the critically endangered Iberian lynx (*Lynx pardinus*). *PLoS One*. 2009;4:e4744.
72. López G, López-Parra M, Fernández L, Martínez-Granados C, Martínez F, Meli ML, et al. Management measures to control a feline leukemia virus outbreak in the endangered Iberian lynx. *Anim Conserv*. 2009;12:173–82.
73. Meli ML, Cattori V, Martínez F, López G, Vargas A, Palomares F, et al. Feline leukemia virus infection: a threat for the survival of the critically endangered Iberian lynx (*Lynx pardinus*). *Vet Immunol Immunopathol*. 2010;134:61–7.
74. Geret CP, Cattori V, Meli ML, Riond B, Martínez F, López G, et al. Feline leukemia virus outbreak in the critically endangered Iberian lynx (*Lynx pardinus*): high-throughput sequencing of envelope variable region a and experimental transmission. *Arch Virol*. 2011;156:839–54.
75. Millán J, Candela MG, Palomares F, Cubero MJ, Rodríguez A, Barral M, et al. Disease threats to the endangered Iberian lynx (*Lynx pardinus*). *Vet J*. 2009;182:114–24.
76. Roelke ME, Johnson WE, Millán J, Palomares F, Revilla E, Rodríguez A, et al. Exposure to disease agents in the endangered Iberian lynx (*Lynx pardinus*). *Eur J Wildl Res*. 2007;54:171–8.
77. Acevedo-Whitehouse K, Cunningham AA. Is MHC enough for understanding wildlife immunogenetics? *Trends Ecol Evol*. 2006;21:433–8.
78. Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*. 1993;364:33–9.
79. Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, Wiley DC. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature*. 1987;329:506–12.
80. Bjorkman PJ, Parham P. Structure, function, and diversity of class I major histocompatibility complex molecules. *Annu Rev Biochem Annual Reviews*. 1990;59:253–88. 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA
81. Sambrook J, Russell DW. *Molecular cloning: a laboratory manual*, 3rd edn. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory; 2001.
82. Dabney J, Meyer M. Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *BioTechniques*. 2012;52:87–94.
83. Sommer S, Courtiol A, Mazzoni CJ. MHC genotyping of non-model organisms using next-generation sequencing: a new methodology to deal with artefacts and allelic dropout. *BMC Genomics BMC Genomics*. 2013;14:542.
84. Klein J, Bontrop R, Dawkins R, Erlich H, Gyllensten U, Heise E, et al. Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics*. 1990;31:217–9.
85. Kosakovsky Pond SL, Posada D, Gravenor MB, Woelk CH, Frost SDW. Automated phylogenetic detection of recombination using a genetic algorithm. *Mol Biol Evol*. 2006;23:1891–901.
86. Huson DH, Klopper TH. Computing recombination networks from binary sequences. *Bioinformatics*. 2005;21:159–65.
87. Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol*. 2006;23:254–67.
88. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28:1647–9.
89. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22:2688–90.
90. Li W-H. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J Mol Evol*. 1993;36:96–9.
91. Charif D, Lobry JR. In: Bastolla U, Porto M, Roman HE, Vendruscolo M, editors. Seqin{R} 1.0–2: a contributed package to the {R} project for statistical computing devoted to biological sequences retrieval and analysis. New York: Springer Verlag; 2007. p. 207–32. Struct. approaches to Seq. Evol. Mol. networks, Popul.
92. Delpont W, Poon AFY, Frost SDW, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*. 2010;26:2455–7.

93. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast, unconstrained bayesian AppRoximation for inferring selection. *Mol Biol Evol.* 2013;30:1196–205.
94. Paradis E, Claude J, Strimmer K. A{PE}: analyses of phylogenetics and evolution in {R} language. *Bioinformatics.* 2004;20:289–90.
95. Paradis E. Pegas: an {R} package for population genetics with an integrated–modular approach. *Bioinformatics.* 2010;26:419–20.
96. Schliep KP. Phangorn: phylogenetic analysis in R. *Bioinformatics.* 2011;27:592–3.
97. Doytchinova IA, Flower DR. In silico identification of supertypes for class II MHCs. *J Immunol.* 2005;174:7085–95.
98. Sandberg M, Eriksson L, Jonsson J, Sjöström M, Wold S. New chemical descriptors relevant for the design of biologically active peptides. A multivariate characterization of 87 amino acids. *J Med Chem.* 1998;41:2481–91.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

