Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A new set of Multilocus Intron Polymorphism Loci (MIPs) for identifying mammal species, subspecies and their first hybrid and backcross generations with their closely related taxa or domestic counterparts

Lara Marinangeli
lara.marinangeli@fmach.it
Fondazione Edmund Mach
Annalisa Scapolatiello
University of Padua
Elisa Boscari
University of Padua
Barbara Crestanello
Fondazione Edmund Mach
Giulio Galla
Fondazione Edmund Mach
Romolo Caniglia
Istituto Superiore per la Protezione e la Ricerca Ambientale
Elena Fabbri
Istituto Superiore per la Protezione e la Ricerca Ambientale
Federica Mattucci
Istituto Superiore per la Protezione e la Ricerca Ambientale
Nadia Mucci
Istituto Superiore per la Protezione e la Ricerca Ambientale
Jaroslav Piálek
Czech Academy of Sciences, Institute of Vertebrate Biology
Francesco Nonnis-Marzano
University of Parma
Heidi Christine Hauffe
Fondazione Edmund Mach
Leonardo Congiu
University of Padua

Article

Keywords:

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11318655/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Animal conservation efforts benefit from rapid, cost-effective genomic markers for species and hybrid identification across taxa. PCR-based protocols have dominated the study of gene flow and introgression in the last 30 years, allowing the molecular characterization of target DNA even from low-quality non-invasive samples. While whole genome sequencing remains beyond the budget of most conservation studies, the characterization of genome-wide single nucleotide polymorphisms (SNPs) has overcome the need for expensive development of marker sets yet provides thousands or tens of thousands of markers; however, SNP-chips or other SNP-based approaches like genotyping-by-sequencing (GBS), require large quantities of high-quality DNA from each sample. Marker sets that combine the advantages of a PCR step (requiring less good quality DNA) applicable to a wide range of species, remain elusive. Recently, Multilocus Intron Polymorphisms (MIPs) have been shown to be useful for defining species, hybrids, and population genomic structure in freshwater fish. Here, we propose a new panel of 187 MIPs for mammals, selected by aligning 121 genomes. Cross-amplification performance of the panel was tested across 15 species, subspecies and hybrids from six different mammalian families. To validate the capacity of this panel to distinguish different classes of hybrids and backcrosses, offspring from known laboratory crosses of Muridae, as well as from free-ranging and domestic taxa of Leporidae, Bovidae, Canidae, Felidae and Suidae were analysed. Results compared favourably with simulated genotypes representing different hybrid and backcross classes, generated from real parental genotypes. The main advantages and limitations of our proposed method are discussed, including its cross-amplification efficiency across different mammalian families and its limited ability to resolve complex patterns of historical introgression.

INTRODUCTION

Natural hybridization between closely related taxa is widely recognized as an evolutionary process that can generate novel adaptive variability (Hunnicutt and Schumer 2026). However, anthropogenic hybridization is globally increasing, driven by intentional admixture, translocations, habitat modification, and range expansion in response to human-mediated climate change (Grabenstein & Tailor 2018; McFarlane and Pemberton 2019). Such hybridization is now considered a major conservation concern, as it may compromise the genetic integrity of wild local populations through introgression from alien or domesticated lines (Hunnicutt and Schumer 2026). Therefore, to assess the impact of anthropogenic hybridization on the long-term viability of natural populations and plan appropriate interventions, the ability to accurately identify ‘pure’ species and interspecific hybrids is critical (McFarlane & Pemberton 2019; Ottenburghs 2021). Conservation genetics benefits from the availability of easily transferable, rapid, and cost-effective markers for identifying species, subspecies, and different levels of inter- and intra-specific admixture (Boscari et al., 2024; Broggini et al., 2024; Mozer et al., 2025). For species identification, the characterization of partial or whole mitochondrial DNA (mtDNA) sequences is still one of the most extensively applied tools (Hebert et al 2003). However, since mtDNA is typically maternally inherited in mammals, the mitochondrial genome can trace only admixture along the maternal line thus for defining hybrids it needs to be combined with information from nuclear markers. On the other hand, microsatellites (STRs), polymorphic, codominant markers characterized by a high mutation rate and found throughout the genome, have also been widely used, especially for genetic diversity studies, parentage analysis, and hybrid identification due to their capacity to detect fine-scale genetic structure (Webster & Reichart 2005; Leo & Millien 2016; Vieira et al., 2016), even with non-invasive samples (Wultsch et al., 2014; De Barba et al., 2023). However, due to cost constraints, the analysis of STRs often involves the amplification of a limited number of loci, which are also species-specific (or restricted to closely related taxa), limiting their application to multiple, phylogenetically distant taxa (Barbarà et al., 2007). With the recent advances of NGS sequencing technologies, the use of Single Nucleotide Polymorphisms (SNPs) for species definition has increased. Despite having lower mutation rates than STRs, these genomic markers are also widely distributed throughout the eukaryote genome, and approaches such as GBS-seq and RAD-seq have proven effective for SNP discovery and analysis in wild species (Narum et al., 2013). However, the uneven distribution of restriction sites across individuals leads to a high level of missing data and requires careful optimization and validation of protocols for each species and technological platform to avoid biasing the results. Moreover, in the absence of a PCR enrichment step, the method requires a relatively large amount of high-quality DNA to ensure sufficient coverage at the target loci, which is generally not available for non-invasive samples of threatened species.

The potential use of introns as molecular markers for both genomic and evolutionary studies is well documented (Irimia & Roy 2008), yet with few exceptions, their investigation has been limited to the amplification and sequencing of individual loci, which generally translates into high costs per identified SNP. Although introns cannot be considered completely neutral (Forcina et al., 2021), these regions are generally subject to weaker selective constraints than exons and therefore, tend to exhibit higher variability in terms of both length and sequence (Yeo et al., 2004). Recently, Boscari et al. (2024) developed the first panel of loci suitable for intron-targeted amplicon sequencing, named ‘Multilocus Intron Polymorphisms’ (MIPs). Using 48 annotated fish genomes, the authors developed a panel of 121 well-conserved exonic flanking primers suitable for the amplification of corresponding intronic sequences. Combining the advantages of

an amplification step with the simultaneous characterization of hundreds of highly polymorphic intronic regions throughout the genome, this approach proved to be a fast, affordable and robust method for distinguishing fish species and detecting episodes of interspecific hybridization, even for cases where the species genomes are not available (Boscari et al., 2024). To date, however, no MIP panel exists for species confirmation and hybrid identification in mammals. Therefore, we developed a panel of 187 MIPs, generated by aligning 121 publicly available mammalian genomes, spanning 24 Orders (out of 27) and 66 families (out of 167) (Burgin et al., 2025; https://www.mammaldiversity.org/). We validated this panel using tissue DNA samples from representative taxa of six mammalian families (Muridae, Leporidae, Bovidae, Canidae, Felidae, Suidae), known for their importance in evolutionary biology, documented anthropogenic hybridization and/or growing conservation interest (Ferrari 2002 (Master thesis); Vernesi et al., 2003; Caniglia et al., 2020; Mattucci et al., 2013, Mattucci et al., 2015; Piálek et al., 2025; Marinangeli et al., 2026). To assess the capacity of the mammalian MIP panel to discriminate among different admixture classes, we compared results obtained from real admixed genotypes with those obtained from simulated genotypes of different classes of hybrids and backcrosses.

We developed a mammalian MIP panel that can be successfully amplified across a broad range of mammalian taxa, while providing sufficient resolution to distinguish among families, species, subspecies, populations, first-generation hybrids and multiple levels of backcrossing within the taxa included in the study. Moreover, our results support the utility of the panel for hybridization and introgression studies across diverse mammalian families, with inferred ancestry proportions closely reflecting the degree of admixture with the parental lines.

METHODS

Identification of candidate intron markers

We selected 121 mammalian genomes (67 annotated) to represent as many mammalian families as possible, including species and subspecies of conservation interest, and/or featuring in studies of natural hybridization, and downloaded them from NCBI (https://www.ncbi.nlm.nih.gov) and Ensembl (Harrison et al., 2024) databases (Tab. S1). For each taxon, only the most recently published genome was used as a reference.

Using the General Transfer Format (GTF) annotation files for each genome, the exon-flanking regions of every annotated intron were extracted, and 100 bp before and after the intron junction were retained for further analyses. A mean of 157,621 sequences was retrieved from each annotated genome, for a total of 10,718,241 intron-flanking regions. Intra-locus sequencing clustering was performed on annotated genomes using CD-HIT version 4.8.1 (Limin et al., 2012) with the following settings: (i) sequence identity = 0.9, (ii) '-g' option = 1, allowing for a slower but more accurate clustering, and (iii) the word length = 5, producing a total of 6,713,515 clusters of intron-spanning regions. To focus on introns with the highest transferability potential, the CD-HIT clusters were then processed to retain only those including sequences of at least 15 species. A total of 32,638 clusters passed this filtering step and were considered for further analysis, for a total of 841,768 exonic-flanking sequences.

To include non-annotated and unmasked (raw) genomes, leading to the identification of previously undetected homologous sequences and those potentially known but masked in annotated genomes, for each cluster, the representative sequence provided by CD-HIT was aligned on the reference genomes of all species previously excluded. Alignment was implemented in BLAT v37.1 (Kent W. J., 2002) using highly stringent criteria; that is, only sequences that aligned with a percentage identity of at least 90% and a sequence coverage of 100% were retained. As a result, a total of 326,370 sequences were retrieved and added to the previously filtered clusters. Of these, 4,545 clusters were then selected based on the following criteria: i) the mean intron length was between 350 and 500 bp, to allow efficient sequencing on the technology selected (i.e. Illumina MiSeq with PE300); ii) the intron was present in > 15 genomes, to allow the broadest possible applicability of the method and iii) the intron was found in the genomes of mountain hare (*Lepus timidus)*, grey wolf (*Canis lupus)*, house mouse (*Mus musculus*) and brown bear (*Ursus arctos*), all species of particular conservation interest to the authors. Based on these criteria, 127 clusters were selected as candidate markers and consensus sequences of exon-flanking regions were used to design locus-specific primer pairs using Geneious Prime 2023.0.4 (Tab. S2a). In addition, given their importance in evolutionary biology, and conservation interest (Lacher et al., 2016; Michell et al., 2022; Hernández-Mazariegos et al., 2025), the same procedure was used to obtain MIP primers focused on the Superorder Glires, representing about 2700 mammalian species. Here, 61 genomes were selected (47 annotated), from 55 rodent and six lagomorph species, obtaining a total of 4,578 clusters from which 32 primer pairs were designed and added to the original (mammalian) set. Finally, to address the increasing calls by wildlife technicians for reliable wolf–dog hybrid detection, an additional set of 33 MIP primers focused on Canids was designed with the same procedure using only the available RefSeq genomes from the four *Canidae* species. The resulting 192 primer pairs were then anchored with Nextera overhang adapters (Forward: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3'; Reverse: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3').

Sample collection and primer validation

The performance of the mammalian MIP panel was tested on samples of representative taxa belonging to the six families: Muridae, Leporidae, Bovidae, Canidae, Felidae, Suidae (Table 1) for which the following samples were available:

Table 1

Samples used for the validation of the mammalian MIP panel. Samples are grouped by family and by category within each family. Hybrid classes are defined as follows: F1 = P1xP2; BC1 = F1xP2; BC2 = BC1xP2; BC3 = BC2xP2; where P1 and P2 represent the two parental taxa; F1, first generation hybrid and; BC, backcross. In the house mouse, 'a' and 'b' represent the two different directions of multiple backcrosses. Hybrid classes not reported were unknown. For each family, the previous molecular characterization for each taxon or hybrid category (STR only or STR and mtDNA) is also reported, together with the corresponding reference. Suidae samples originate from different areas, as follows: [a]Hungary; [b]Tuscany; [c]two different populations from central Italy (Castelporziano and Maremma); [d]two different farms in the Italian Regions of Calabria and Tuscany.

Family	Category	Nr.	Characterization	Reference
Muridae	*Mus musculus musculus* (P1)	6	laboratory crosses	Piálek et al., 2025
	F1	10		
	BC1a-b	6, 6		
	BC2a-b	6, 5		
	BC3a-b	6, 6		
	M. m. domesticus (P2)	6		
Leporidae	*Lepus europaeus* (P1)	9	STR + mtDNA	Pecchioli et al., 2006
	L. europaeus x L. timidus	7		
	L. timidus varronis (P2)	12		
Bovidae	*Capra hircus cretica*	4	STR + mtDNA	Crestanello et al., 2009;
	Rupicapra pyrenaica ornata	4		Pecchioli 2005
	R. p. pyrenaica	5		
	R. p. pyrenaica x R. r. rupicapra	4		
	R. r. rupicapra	5		
Canidae	*Canis lupus familiaris* (P1)	10	STR + mtDNA + ChrY + K-locus	Caniglia et al., 2020
	F1	1		
	BC1	4		
	BC2	2		
	BC3-BC4	2		
	C. l. italicus (P2)	10		
Felidae	*F. silvestris catus*	3	STR	Mattucci et al., 2013;
	F. s. silvestris x F. s. catus	4		Mattucci et al., 2015;
	F. s. silvestris	4		ISPRA Felis database
Suidae	*Sus scrofa scrofa*[a]	5	STR	Ferrari 2002 (Master thesis);
	S. s. scrofa x S. s. majori[b]	4		Vernesi et al., 2003
	S. s. majori[c]	7		
	S. s. domesticus[d]	5		

Muridae

Spleen samples were collected during dissection of captive-bred wild house mice, previously euthanized by cervical dislocation. For details on how the mice were housed and handled in the laboratory, we kindly refer to Piálek et al., 2025. 69 ethanol preserved samples were provided by JP at the Institute of Vertebrate Biology (IVB, Studenec, Czech Republic).The IVB facility is authorized for the use of experimental animals (licences 61974/2017-MZE-17214 and MZE-50144/2022–13143), as well as for the breeding and supply of experimental animals to third parties (62065/2017-MZE-17214 and MZE-50151/2022–13143). These licences are in compliance with the corresponding regulations and standards of the European Union, as specified in Council Directive 86/609/EEC. Spleen tissues from the inbred parental lines STRA, representing *M. m. domesticus* (P1) and BUSNA, representing *M. m. musculus* (P2) (Piálek et al., 2025), along with randomly selected hybrids (corresponding to F1) and individuals obtained from multiple backcrosses were available (Table 1).

Leporidae and Bovidae

Samples were obtained from animals killed during regular, legally authorized hunting activity. No animals were killed for the purposes of this study, and no permits were required for sample collection. DNA from 28 hares (*Lepus* spp.), 18 chamois (*Rupicapra* spp.) and four Cretan ibex (*Capra hircus cretica*), previously taxonomically characterised using panels of a few STR and mtDNA markers (Pecchioli 2005; Crestanello et al., 2009), were available at the Fondazione E. Mach (BC, HCH).

Canidae and Felidae

All samples were post-mortem obtained from animals found dead following road traffic accidents. Samples from wolves and wildcats were collected opportunistically as part of national multi-year monitoring activities managed by the Italian Institute for Environmental Protection and Research (ISPRA) (Gervasi et al. 2024; Mattucci et al. 2013). Tissue samples from domestic dogs and domestic cats were collected by veterinarians during necropsies of animals found dead. No animals were captured, killed, or subjected to experimental procedures for this study, and no animal research ethics committee approval was required for sample collection. DNA from 10 domestic dogs (*Canis lupus familiaris)*, 10 Italian grey (Appennine) wolves (*C. l. italicus)*, and nine wolf x dog hybrids with different levels of domestic ancestry, as well as three domestic cats (*Felis silvestris catus)*, four European wildcats (*F. s. silvestris)* and four putative wildcat x domestic cat hybrids, all previously characterised for subspecies and admixture category using STR markers (Caniglia et al., 2020; Mattucci et al., 2013; Mattucci et al., 2015), were provided by the Italian Institute for Environmental Protection and Research (RC, EF, FM, NM).

Suidae

Samples were collected from hunted or trapped pigs and wild boar in Italy and Hungary. No animals were killed for research purposes, and no permits were required for sample collection. DNA from five domestic pigs *(Sus scrofa domesticus)*, five wild boar *(Sus scrofa scrofa*), seven Italian boar (*S. s. majori*), and four wild boar-Italian boar hybrids previously characterised for subspecies and hybrid class using STR markers (Ferrari, 2002; Vernesi et al., 2003), were provided by the Department of Life Sciences and Biotechnology, University of Ferrara.

DNA was extracted from the house mouse tissue samples using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). For DNA extraction protocols of other species, please refer to the cited references above. All DNA samples were quantified using UV/VIS (Techan spark). To test the 192 primer pairs, PCR amplification of each intronic locus was performed separately on six DNA samples (two from each of three mammalian families Bovidae, Canidae and Leporidae) using two alternative annealing temperatures: 57°C and 60°C. PCR products were visualized using a QIAxcel DNA Screening cartridge in a QIAxcel Connect capillary electrophoresis instrument (QIAGEN, Hilden, Germany).

Primer multiplexing and library preparation

The 187 primer pairs that passed the amplification test were divided into eight multiplexes based on the best-performing annealing temperature and amplicon concentration estimated with QiAxcel from the single-locus PCR amplification assays (Tab. S2a): five multiplexes with generic mammalian loci (121 loci; MAM57a, MAM57b, MAM57c, MAM57d; MAM60); one focused on Canidae (25 loci; CAN57); one onLeporidae (23 loci; LEP57); and one with a mix of generic loci and those focused on Canidae and Leporidae (18 loci; MIX60). Suitability of each primer combination was confirmed using PrimerPooler V. 2.0 (Brown et al., 2017) to ensure representative amplification of targeted sequences. For each mammalian family, appropriate multiplexes were tested as listed in Table 2. All multiplexes were tested on Muridae for which we have a documented hybridization history. DNA amplification was performed with the Multiplex PCR Master Mix (Qiagen, Hilden, Germany) according to the manufacturer's protocol, and negative DNA amplification controls were added to detect contamination. For each Multiplex three independent amplification replicates per sample were performed, following Boscari et al.

(2024), to minimize PCR-induced errors. The cycling parameters were as follows: initial denaturation at 95°C for 15', followed by 35 cycles of denaturation at 94°C for 30", annealing at 57°C or 60°C for 90", as reported in Tab. S2a, and extension at 72°C for 60", with a final extension at 60°C for 30'. After amplification, for each sample, the three replicates of each multiplex were pooled and all multiplexes were then combined into a single library for sequencing. Amplicon sequencing with 30K reads/sample sequencing depth was performed on an Illumina MiSeq PE 300.

Table 2
Genotyping success of MIP loci for each of the six mammalian families. The multiplexes amplified per family are highlighted in grey. The number of successfully genotyped loci, expressed as both absolute values and percentages, is reported.

Family	Multiplexes amplified								Genotyped/ analysed loci	% Success	PIDsibs	Min Nr. of loci for individual discrimination
	MAM 57a	MAM 57b	MAM 57c	MAM 57d	MAM 60	LEP 57	CAN 57	MIX 60				
Muridae	5	5	5	5	5	5	5	5	65 / 187	35	5,10E-05	33
Leporidae									105 / 162	65	7,10E-23	16
Bovidae									84 /139	60	1,65E-14	29
Canidae									117 / 164	71	7,63E-11	15
Felidae									101 / 139	73	5,05E-12	7
Suidae									90 / 139	65	8,81E-06	15

Bioinformatic analysis for haplotype identification

Raw amplicon sequencing data were processed using the snakemake pipeline described in Scapolatiello et al. (2025). Briefly, raw fastq reads were processed using Cutadapt (Martin et al., 2011) to demultiplex, discard low-quality bases, and remove adapters, setting a minimum of 15 as the quality cutoff of the reads. Paired-end reads were merged with FLASH v2.2.00 (Magoč & Salzberg 2011), discarding loci where merged reads were less than 20% of the total and coverage was fewer than 30 sequences. The SeekDeep (Nj, Hathaway, 2018) pipeline was used to reconstruct the genotype, cluster allelic variants, estimate allele frequencies (with the qluster package), and filter out low-frequency alleles (with processClusters). The final output file retained only the most abundant alleles by frequency for each locus and sample, with a cumulative fixed abundance. Finally, the output was converted to a Genepop file format to perform subsequent analyses. The discriminatory power of the marker panel across mammalian families was initially evaluated by Principal Component Analysis (PCA) based on euclidean distance using the R package adegenet v.2.1.11 (Jombart 2008). PCA was performed on two datasets: a broader dataset comprising 147 loci that produced genotypes in at least five of the six mammalian families analyzed, and a more stringent dataset comprising 24 loci that successfully generated genotypes in all families included in the study. The same R package was used to calculate the Probability of Identity among siblings (PIDsibs) and the minimum number of loci necessary to discriminate among individuals. The genetic structure and individual assignment coefficients (q_i) were defined using Structure v.2.3.4 with 20 iterations, 1000 burnin, 10,000 MCMC and setting the Admixture model with correlated allele frequencies. Using the method Evanno et al. (2005) implemented in StructureSelector (Li & Liu, 2018) the number of clusters that better represented our data using K = 1–6 was defined.

To assess the capacity of the mammalian MIP panel to discriminate among different admixture classes, for families with at least five parental individuals with *qi* > 0.99, i.e. for each taxon pair of Muridae (*M. m. musculus* x *M. m. domesticus*), Leporidae (*L. europaeus* x *L. timidus*), Bovidae (*R. p. pyrenaica* x *R. r. rupicapra*), and Canidae (*C. l. familiaris* x *C. l. italicus*), parental genotypes were processed using a Python (v. 3.12.11) script executed in a local computational environment (Pandas module v 2.2.2, McKinney et al. 2010) to simulate virtual genotypes representing diverse hybrid and backcross classes, generated through successive backcrosses with one parental line. Gametes were generated from the parental genotypes, representing the potential genetic combinations transmissible to the progeny. A combination function was implemented to merge genetic values from two gametes. Using random selection, a fixed number of new F1 individuals were created by combining gametes from P1 and P2. Starting from the F1 generation, three consecutive backcross cycles

were performed to estimate the number of cycles after which the contribution of one parent was no longer detectable. Simulated genotypes were then analyzed with Structure using the same parameters described above. The resulting outputs were used to define assignment thresholds for classifying genotypes into the different hybrid and backcross categories. These thresholds were subsequently compared with *q* values obtained for the real individuals.

RESULTS

Cross-amplification and inter-family discrimination performance of MIPs

Five out of 192 primer pairs were discarded as they did not yield reliable amplification results on single-locus PCR tests. The remaining 187 primer pairs (97%) were included in the final MIP panel (Tab S2a). The application of the panel to samples of six mammalian families showed that a median of 95.5 loci were amplified for each family (range: 65–117; Table 2; S2b). PIDsibs values were consistently low across all tested families, ranging from 10^{-5} to 10^{-23} (Table 2). Moreover, for each family the minimum number of loci required to reliably discriminate among individuals was substantially lower than the number of loci successfully genotyped in this study (Table 2).

Principal component analysis (PCA) performed with a broader dataset comprising 147 loci across all mammalian families included in the study revealed a clear spatial separation among taxa (Fig. S1). Notably, Canidae, Leporidae, and Muridae were clearly and consistently distinguished. However, while Bovidae, Felidae, and Suidae showed partial overlap when examined alongside all other families, they nevertheless formed distinct and well-resolved clusters when analyses were restricted to these families alone (Fig. S2). Similar results were obtained when PCA were performed on a stringent dataset comprising 24 loci.

Admixed Class Identification

The analysis of Muridae simulated genotypes generated through different backcrosses with *M. m. domesticus* demonstrated good discriminatory power for the five admixed classes tested (Fig. 1a; Tab. S3a). As shown in Fig. 1a, parental lines could be distinguished from each other (P1 in blue and P2 in orange, $q_i > 0.99$; Tab. S3a) and from the five admixed categories (F1, BC1, BC2, BC3). In addition, admixed individuals consistently showed an increasing *qi*, to the parental line with an increasing number of backcrosses with that line (Tab. S3a). Comparable threshold values were observed for simulated and real genotypes, with F1 hybrids characterized by q_i values ranging from 0.40 to 0.60 and putative introgressed individuals by q_i values > 0.60. In both simulated and real genotypes, individuals assigned to the third backcross generation (BC3) were not clearly distinguished from the P2 parental group, with some individuals exhibiting q_i values of up to 0.99 to the parental cluster.

Analysis of Leporidae, Bovidae and Canidae simulated genotypes showed, as expected, a progressive increase in the genetic contribution of the selected parental line with increasing numbers of backcrosses, confirmed by their q_i values (Fig. S3; Tab. S3a). Notably, domestic/allochthonous ancestry of Leporidae and Bovidae (*Lepus* and *Rupicapra* species) was detectable up to the second backcross generation (BC2) with $qi < 0.16$ and $q_i < 0.14$, respectively, while in Canidae (subspecies of *C. lupus*) it was only clearly detectable up to the first backcross generation (BC1) with $q_i < 0.29$ (Tab. S3a).

Analyses of real Leporidae and Bovidae genotypes demonstrated that pure parental lines, previously defined using species-specific sets of STR and mtDNA markers, can be clearly distinguished based on our MIP panel (Fig. 2a-b; Tab. S3b), with each parental individual showing $qi = 1.00$. Comparable results were obtained for Canidae, with all individuals previously assigned to either the domestic (*C. l. familiaris*) or wild (*C. l. italicus*) parental lines using a set of 39 STR markers, being consistently assigned to the same parental lines based on MIPs with $qi > 0.99$ (Fig. 2c; Tab. S3b).

Based on the cluster assignment thresholds derived from the simulated genotypes, only one leporid individual was classified as an F1 hybrid ($0.40 < qi < 0.60$ according to the simulated data), whereas all remaining hybrids showed high assignment coefficients ($qi > 0.98$) to *L. timidus* (Tab. S3b). However, similarly high q_i to this parental cluster were also detected in many simulated genotypes from the second backcross (BC2) onward (Tab. S3a). Similar trends were reported for Bovidae, with all hybrids exhibiting $q_i = 1.00$ to *R. r. rupicapra*, a pattern that was also observed in several simulated genotypes from to the third backcross generation (BC3) (Tab. S3).

In Canidae, the only hybrid previously classified as an F1 individual using an STR panel ($q_i = 0.48$; Caniglia et al., 2020) showed a $q_i = 0.78$ to *C. l. italicus* when analyzed with MIPs (Tab. S3b), exceeding the assignment threshold to this cluster defined by simulated MIP genotypes ($0.35 < q_i < 0.65$) (Tab. S3a). All other admixed individuals showed progressively increasing q_i values to *C. l. italicus* as the

number of backcrosses with this parental line increased, following a pattern comparable to that observed in the simulated genotypes, ultimately reaching q_i = 1.00 in BC3 individuals (Tab. S3).

In Felidae, the assignments obtained with the MIP panel were consistent with previous analyses based on 29 STR markers conducted by Mattucci et al. (2013, 2015). All individuals previously assigned to the domestic cat (*F. s. catus*) were similarly classified using the MIP dataset (Fig. S4a; Tab. S3b), with high assignment probabilities ($q > 0.99$). Likewise, assignments of wildcat individuals were largely concordant between the two sets of markers: the assignment thresholds defined by Mattucci et al. (2013, 2015) from Structure analyses of simulated STR genotypes ($qi > 0.90$ for wildcats, $0.80 < qi < 0.90$ for putatively introgressed wildcats, and $qi < 0.80$ for recent hybrids) were comparable to the ones obtained in our study, with the only exception of a single *F. s. silvestris* individual (Tab S3b).

Finally, in agreement with STR and genome-wide studies (Vernesi et al., 2003; Scandura et al., 2022), Suidae samples from Hungary (*S. s. scrofa*) were clearly distinguishable from those from central Italy, i.e. Tuscany (hybrids between *S. s. scrofa* × *S. s. majori*), and Castelporziano and Maremma (*S. s. majori*) based on MIPs (Fig. S4b; Tab S3b). Also, a certain degree of introgression from Hungarian individuals into all three central Italian populations was confirmed. However, while STR-based analyses indicated that the Castelporziano and Maremma populations were genetically distinct from one other and from the Hungarian population (Vernesi et al., 2003), MIP-based results did not clearly separate these two populations. Results for domestic pigs also indicated a clear genetic differentiation between domesticated individuals from Calabria and those from central Italy (Fig. S4b; Tab S3b). Interestingly, individuals of Cinta Senese from central Italy showed a higher genetic contribution from the Hungarian wild boar, rather than from geographically closer wild boar populations.

DISCUSSION

Intertaxon hybridization, especially that induced by anthropogenic activities, is increasingly recognized as a major threat to biodiversity and can ultimately contribute to population decline or species extinction. This highlights the need for developing readily transferable, rapid, and cost-effective molecular markers that can be applied in conservation genetics to accurately identify species, subspecies, and different levels of inter- and intraspecific admixture. Here, we developed and validated the first panel of 187 MIPs for mammals by testing its performance across individuals from six mammalian families. Our results demonstrate the capacity of the MIP panel to discriminate among different families and species, subspecies, populations, first generations of hybrids and different levels of backcrossing within the taxa included in this study, with inferred ancestry proportions reflecting the backcrossing with a given parental line.

Good cross-amplification and efficient inter-family discrimination

The application of the panel across six mammalian families demonstrated successful amplification of at least 65 loci for each family, indicating that the panel has a good potential for cross-application to other mammalian taxa. This result highlighted a key advantage of MIPs: a mean cross-amplification capability of about 62% across mammalian families tested here. Moreover, the successfully amplified loci proved to be highly reliable for inter-family discrimination and the identification of introgression following hybridization. Indeed, using a dataset of 147 loci, PCA revealed distinct clustering of the taxa across all mammalian families examined. The observed clustering patterns indicate that the selected loci capture phylogenetically informative signals that are sufficiently robust to differentiate taxa at the family level. The consistently low PIDsibs values and the substantial excess of genotyped loci relative to the minimum required for individual discrimination indicate that the panel provides reliable and highly robust individual identification, even when closely related, across diverse mammalian taxa.

Identification of parental classes, first hybrid and backcross generations

The close similarity between the laboratory and in silico results for *M. musculus* supports the conclusion that the mammalian MIP panel can reliably distinguish different admixed classes, even when derived from two closely related parental lines, up to BC3 individuals. Domestic or allochthonous ancestry in Leporidae and Bovidae simulated genotypes could be detected up to the second backcross generation (BC2), whereas in Canidae it was consistently identifiable only up to the first backcross generation (BC1). This pattern suggests that the power of this panel to detect admixed ancestry may be affected by the genetic divergence between the parental taxa, ie. higher at the interspecific level (Leporidae, Bovidae) than at the subspecific level (Canidae). Consistently, previous studies showed that the power to detect hybridization and introgression depends on the degree of genetic differentiation represented in the marker panel, the genomic distribution of markers, and the analytical approaches used (Vähä & Primmer 2006; McFarlane & Pemberton 2019). Genome-wide marker sets such as MIPs are particularly effective at identifying recent hybridization events; however, older hybridization events are increasingly difficult to detect with increasing numbers of backcross generations. In our dataset, this is translated into reliable detection up to BC3 in interspecific crosses and BC2 in inter-subspecific crosses. When analysing real Leporidae and Bovidae genotypes,

many hybrids showed signatures of advanced hybridization, with most introgressed individuals being genetically indistinguishable from one of the parental lines, i.e. *L. timidus* and *R. r. rupicapra*, respectively, a pattern that confirms the one observed in BC2-BC3 simulated genotypes. This result was expected as most of these “admixed individuals” were categorized based on the discrepancy between STR and mtDNA markers, morphology or sampling location, indicating ancient multiple hybridization events. Similar results were obtained for Canidae; however, the individual assignment coefficient to *C. l. italicus* of introgressed individuals (BC1-BC4) derived from both simulated and real MIP genotypes, were higher than those previously obtained using the STR panel (Caniglia et al., 2020) suggesting that, although hybrids can be reliably detected, the MIP panel has a reduced ability to assign genotypes to specific backcross classes. This may be explained by a lower sensitivity of a broadly applicable marker set, such as MIPs, compared with STR panels specifically designed and selected for their high discriminatory power between wolves and dogs. However, in Felidae, the panel provided estimates of genetic introgression that are comparable with those previously defined using STRs (Mattucci et al., 2013; Mattucci et al., 2015), as ancestry coefficients inferred from MIPs were broadly consistent with those obtained using the STR panel. MIP genotypes data clearly differentiated Suidae individuals from Hungary (*S. s. scrofa*) from those in central Italy, mirroring patterns previously reported from STR and genome-wide datasets (Vernesi et al., 2003; Scandura et al., 2022). However, in contrast to previous findings obtained with STR markers, MIP-based results did not clearly distinguish the Castelporziano and Maremma populations, a result that may be attributable to the lower number of samples used in this study compared to the previous one.

Overall, the key advantage of MIPs lies in the cross-amplification efficiency across a broad range of phylogenetically distant mammalian families. This makes the panel particularly suitable for, e.g., forensic genetics involving multiple species, reducing the need for species-specific protocol development, ultimately reducing both time and costs. Moreover, the inclusion of a PCR amplification step is expected to enhance the applicability of this approach to low-quantity and degraded DNA samples, such as those obtained from non-invasive sampling or environmental DNA (eDNA). Importantly, the applicability of this panel to non-model species without the need for a reference genome further highlights the versatility and transferability of MIPs compared to other molecular techniques, such as SNPs-based approaches, which typically rely on the availability of a reference genome for the species under study. Furthermore, it should be noted that, to date, only the haplotypic information derived from the MIPs has been utilized, with each haplotype corresponding to an allele. However, these markers also provide an additional layer of information through their underlying sequence variation, which may offer valuable opportunities for future research. Potential applications include the development of diagnostic markers for direct PCR-based identification and the assessment of genetic divergence among distinct evolutionary lines. Nevertheless, some limitations should also be considered. In particular, the power of MIPs to detect hybridization is expected to decrease with increasing numbers of backcross generations, as successive introgression events progressively dilute the genomic contribution of one parental line. As a consequence, older hybridization events and highly introgressed individuals may become increasingly difficult to distinguish from pure parental individuals, potentially reducing the sensitivity of the approach for the detection of advanced hybridization and historical introgression.

CONCLUSIONS

The mammalian MIP panel presented here showed good cross-taxon amplification and performed well in distinguishing different species, subspecies, populations and different levels of hybrid classes with their closely related taxa or domestic counterparts. The MIP panel distinguishes parental lines from hybrid and introgressed individuals up to BC3 in Muridae, Leporidae and Bovidae, and up to BC2 in Canidae. These findings were partly consistent with previously published studies on the investigated taxa, confirming several known patterns of genetic differentiation and admixture, while also revealing some discrepancies. The power to detect historical admixture proved to be lower than that of taxon-specific marker panels (e.g. STRs), likely due to the fact that MIPs were designed as a broadly transferable marker system across mammals rather than being optimized for specific taxa. Therefore, the panel may be more effective for identifying recent hybridization events than for resolving complex patterns of historical introgression.

Future applications of this MIP panel to additional wild populations with larger sample sizes would be valuable for further assessing its effectiveness and applicability in conservation genetics. Moreover, future studies should investigate the applicability of the intronic panel to other mammalian species, to non-invasive (e.g. feces, hair) and eDNA samples.

Declarations

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

This study was partially carried out with funding to HCH at the Fondazione E. Mach (Project BIOALPEC) under the National Biodiversity Future Centre (NBFC) Project (code CN_00000033, Concession Decree No. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research, CUPD43C22001280006), funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.4 - Call for tender No. 3138 of 16 December 2021, rectified by Decree n.3175 of 18 December 2021 of Italian Ministry of University and Research funded by the European Union – NextGenerationEU. This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them. Mouse samples were obtained as part of the Czech Science Foundation project (206/08/0640).

Author Contribution

BC, RC, EF, FM, NM, JP and HCH provided the samples. LM and BC completed the laboratory analyses. Data analyses were provided by LM, AS and GG, with support from EB, BC, RC, EF, FM, NM, FNM, HCH and LC. All authors contributed to the writing, edited and approved the final manuscript.

Acknowledgement

We thank Giorgio Bertorelle and Silvia Fuselli (Population Genetics and Genomics Group, Department of Life Sciences and Biotechnology, University of Ferrara) for providing Suidae samples.

Data Availability

The raw amplicon-sequencing data has been deposited at NCBI Sequence Read Archive (SRA) under the BioProject ID PRJNA1541250. Reviewers can access BioProject and associated SRA metadata at https://dataview.ncbi.nlm.nih.gov/object/PRJNA1541250?reviewer=dsqf31ehuesp52agckja2gtm33

References

1. Barbará, T. et al. Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Mol. Ecol.* **16**, 3759–3767. https://doi.org/10.1111/j.1365-294X.2007.03439.x (2007).
2. Boscari, E. et al. MIPs: multi-locus intron polymorphisms in species identification and population genomics. *Sci. Rep.* **14**, 17870. https://doi.org/10.1038/s41598-024-68065-8 (2024).
3. Broggini, C. et al. Development and Validation of an SNP-Based OpenArray® Genotyping Panel for Discriminating Coturnix coturnix, Coturnix japonica and Their Hybrids. *Genes* **17**, 739. https://doi.org/10.3390/genes17070739 (2026).
4. Burgin, C. J. et al. How many mammal species are there now? Updates and trends in taxonomic, nomenclatural, and geographic knowledge. *Preprint at.* https://doi.org/10.1101/2025.02.27.640393 (2025).
5. Caniglia, R. et al. A standardized approach to empirically define reliable assignment thresholds and appropriate management categories in deeply introgressed populations. *Sci. Rep.* **10**, 2862. https://doi.org/10.1038/s41598-020-59521-2 (2020).
6. Crestanello, B. et al. The Genetic Impact of Translocations and Habitat Fragmentation in Chamois (*Rupicapra*) spp. *J. Hered.* **100**, 691–708. https://doi.org/10.1093/jhered/esp053 (2009).
7. De Barba, M. et al. Individual genotypes from environmental DNA: Fingerprinting snow tracks of three large carnivore species. *Mol. Ecol. Resour.* **24**, e13915. https://doi.org/10.1111/1755-0998.13915 (2024).
8. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* **14**, 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x (2005).
9. Ferrari, M. Tipizzazione genetica ed analisi dei processi di ibridazione e reintroduzioni nelle popolazioni di cinghiale (S*us scrofa*) dell'Italia centrale mediante lo studio di loci microsatelliti. Unpub. Master thesis, Università di Ferrara, Italy (2002).
10. Forcina, G., Camacho-Sanchez, M., Tuh, F. Y. Y., Moreno, S. & Leonard, J. A. Markers for genetic change. *Heliyon* **7**, e05583. https://doi.org/10.1016/j.heliyon.2020.e05583 (2021).

11. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565 (2012).
12. Gervasi, V. et al. Estimating distribution and abundance of wide-ranging species with integrated spatial models: Opportunities revealed by the first wolf assessment in south-central Italy. *Ecol. Evol.* **14**, e11285. https://doi.org/10.1002/ece3.11285 (2024).
13. Grabenstein, K. C., Taylor, S. A. & Breaking Barriers Causes, Consequences, and Experimental Utility of Human-Mediated Hybridization. *Trends Ecol. Evol.* **33**, 198–212. https://doi.org/10.1016/j.tree.2017.12.008 (2018).
14. Harrison, P. W. et al. Ensembl 2024. *Nucleic Acids Res.* **52**, D891–D899. https://doi.org/10.1093/nar/gkad1049 (2024).
15. Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* 270, 313–321 (2003). https://doi.org/10.1098/rspb.2002.2218
16. Hernández-Mazariegos, W. C., Palma, R. E. & Escobar, L. E. Rodents of Chile: a brief appraisal of their conservation status and ecological significance. *ZK* **1254**, 107–129. https://doi.org/10.3897/zookeys.1254.148057 (2025).
17. Hunnicutt, K. E. & Schumer, M. Hybridization in Animal Evolution. *Annu. Rev. Ecol. Evol. Syst.* https://doi.org/10.1146/annurev-ecolsys-102924-052935 (2026).
18. Irimia, M. & Roy, S. W. Spliceosomal introns as tools for genomic and evolutionary analysis. *Nucleic Acids Res.* **36**, 1703–1712. https://doi.org/10.1093/nar/gkn012 (2008).
19. Jombart, T. *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405 (2008). https://doi.org/10.1093/bioinformatics/btn129
20. Kent, W. J. BLAT - The BLAST -Like Alignment Tool. *Genome Res.* **12**, 656–664. https://doi.org/10.1016/j.ymeth.2009.06.009 (2002).
21. Lacher, T. E., Murphy, W. J., Rogan, J., Smith, A. T. & Upham, N. S. Evolution, phylogeny, ecology and conservation of the Glires clade: Lagomorpha and Rodentia. In: *Handbook of Mammals of the World - Lynx Editions* (ed. Wilson, D. E., Lacher, T.E., Jr., Mittermeier, R. A.) 15–26. (2016).
22. Leo, S. S. T. & Millien, V. Microsatellite markers reveal low frequency of natural hybridization between the white-footed mouse (*Peromyscus leucopus*) and deer mouse (*Peromyscus maniculatus*) in southern Quebec. *Can. Genome*. **60**, 454–463. https://doi.org/10.1139/gen-2016-0163 (2017).
23. Magoč, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507 (2011).
24. Marinangeli, L. et al. Sympatric *Lepus* spp. in the central Italian Alps host significantly different gut microbiotas. *Sci. Rep.* **16**, 15866. https://doi.org/10.1038/s41598-026-44592-4 (2026).
25. Mattucci, F. et al. Genetic structure of wildcat (*Felis silvestris*) populations in Italy. *Ecol. Evol.* **3**, 2443–2458. https://doi.org/10.1002/ece3.569 (2013).
26. Mattucci, F., Oliveira, R., Lyons, L. A., Alves, P. C. & Randi, E. European wildcat populations are subdivided into five main biogeographic groups: consequences of Pleistocene climate changes or recent anthropogenic fragmentation? *Ecol. Evol.* **6**, 3–22. https://doi.org/10.1002/ece3.1815 (2016).
27. McFarlane, S. E. & Pemberton, J. M. Detecting the True Extent of Introgression during Anthropogenic Hybridization. *Trends Ecol. Evol.* **34**, 315–326. https://doi.org/10.1016/j.tree.2018.12.013 (2019).
28. McKinney, W. *Data Structures for Statistical Computing in Python* in 56–61 (Austin, 2010). https://doi.org/10.25080/Majora-92bf1922-00a
29. Michell, C. T., Pohjoismäki, J. L. O., Spong, G. & Thulin, C. G. Mountain- and brown hare genetic polymorphisms to survey local adaptations and conservation status of the heath hare (*Lepus timidus sylvaticus*, Nilsson 1831). *Sci. Data*. **9**, 667. https://doi.org/10.1038/s41597-022-01794-5 (2022).
30. Mozer, A. et al. Determination of tortoises origin based on SNPSTR markers: a case study of melanistic *Testudo hermanni*. *Conserv. Genet. Resour.* **17**, 253–264. https://doi.org/10.1007/s12686-025-01401-w (2025).
31. Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R. & Hohenlohe, P. A. Genotyping-by-sequencing in ecological and conservation genomics. *Mol. Ecol.* **22**, 2841–2847. https://doi.org/10.1111/mec.12350 (2013).
32. Ottenburghs, J. The genic view of hybridization in the Anthropocene. *Evol. Appl.* **14**, 2342–2360. https://doi.org/10.1111/eva.13223 (2021).
33. Pecchioli, E. Variabilità e struttura genetica in *Lepus timidus*: aspetti micro– e macrogeografici, filogenetici e adattativi. Unpub. PhD thesis, Università di Ferrara, Italy (2005).

34. Pecchioli, E. et al. Progetto FAUNAGEN - Conservazione e gestione della fauna: un approccio genetico. Report Nr. 35, ISSN: 1827–8558 (2006).
35. Piálek, J. et al. Phenogenomic resources immortalized in a panel of wild-derived strains of five species of house mice. *Sci. Rep.* **15**, 12060. https://doi.org/10.1038/s41598-025-86505-x (2025).
36. Scandura, M. et al. Resilience to Historical Human Manipulations in the Genomic Variation of Italian Wild Boar Populations. *Front. Ecol. Evol.* **10**, 833081. https://doi.org/10.3389/fevo.2022.833081 (2022).
37. Scapolatiello, A., Boscari, E., Schiavon, L., Vitulo, N. & Congiu, L. Intronomics-MIP: a snakemake pipeline for analyzing multilocus intron polymorphisms in species identification and population genomics. *BMC Res. Notes*. **18**, 203. https://doi.org/10.1186/s13104-025-07264-6 (2025).
38. Vähä, J. & Primmer, C. R. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Mol. Ecol.* **15**, 63–72. https://doi.org/10.1111/j.1365-294X.2005.02773.x (2006).
39. Vernesi, C. et al. The genetic impact of demographic decline and reintroduction in the wild boar (*Sus scrofa*): A microsatellite analysis. *Mol. Ecol.* **12**, 585–595. https://doi.org/10.1046/j.1365-294x.2003.01763.x (2003).
40. Vieira, M. L. C., Santini, L., Diniz, A. L. & Munhoz, C. D. F. Microsatellite markers: what they mean and why they are so useful. *Genet. Mol. Biol.* **39**, 312–328. https://doi.org/10.1590/1678-4685-GMB-2016-0027 (2016).
41. Webster, M. S. & Reichart, L. Use of Microsatellites for Parentage and Kinship Analyses in Animals. in Methods in Enzymology vol. 395 222–238 (Elsevier, (2005).
42. Wultsch, C., Waits, L. P. & Kelly, M. J. Noninvasive individual and species identification of jaguars (*Panthera onca*), pumas (*Puma concolor*) and ocelots (*Leopardus pardalis*) in Belize, Central America using cross-species microsatellites and faecal DNA. *Mol. Ecol. Resour.* **14**, 1171–1182. https://doi.org/10.1111/1755-0998.12266 (2014).
43. Yeo, G., Hoon, S., Venkatesh, B. & Burge, C. B. Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15700–15705 (2004). https://doi.org/10.1073/pnas.0404901101

Figures

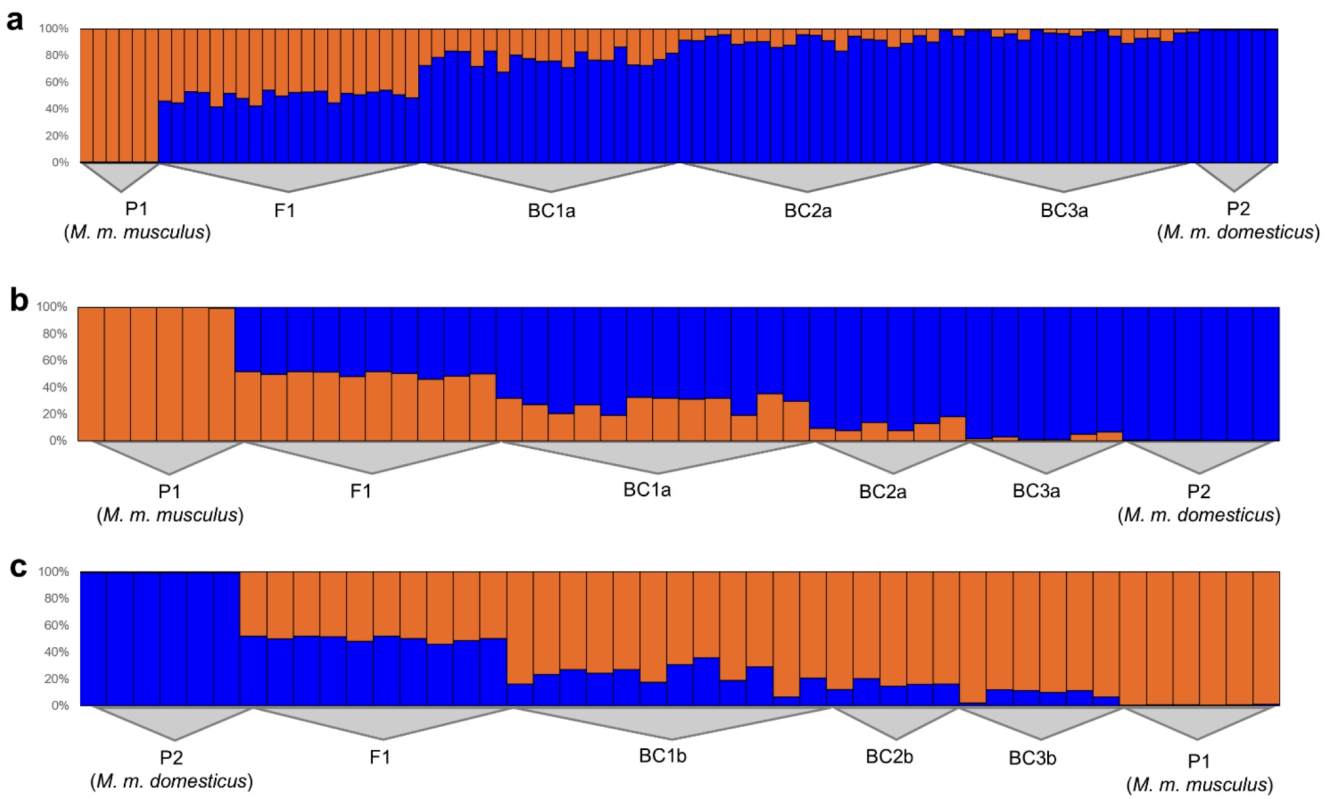


Figure 1

Barplots obtained with Structure for *Mus musculus* parental and hybrid individuals. According to the Evanno method, the most likely number of genetic populations inferred from our data is K=2. Each vertical bar represents a different individual, while the y-axis shows the proportions of inferred ancestry for the two clusters identified by the software. **a)** Results of in silico simulations of hybridization using MIP genotypes generated from *M. musculus* parental lines**.** Results obtained from captive-bred individuals are shown in panel **b)**, which includes hybrids generated through repeated backcrosses with the *M. m. domesticus* parental line (P2); and in panel **c)**, which includes hybrids resulting from repeated backcrosses with the *M. m. musculus*parental line (P1). Hybrid classes are defined as follows: P1=*M. m. musculus*; P2=*M. m. domesticus;* F1=P1xP2; BC1a=F1axP2; BC2a=BC1axP2; BC3a=BC2axP2; F1=P2xP1; BC1b=F1bxP1; BC2b=BC1bxP1; BC3b=BC2bxP1.

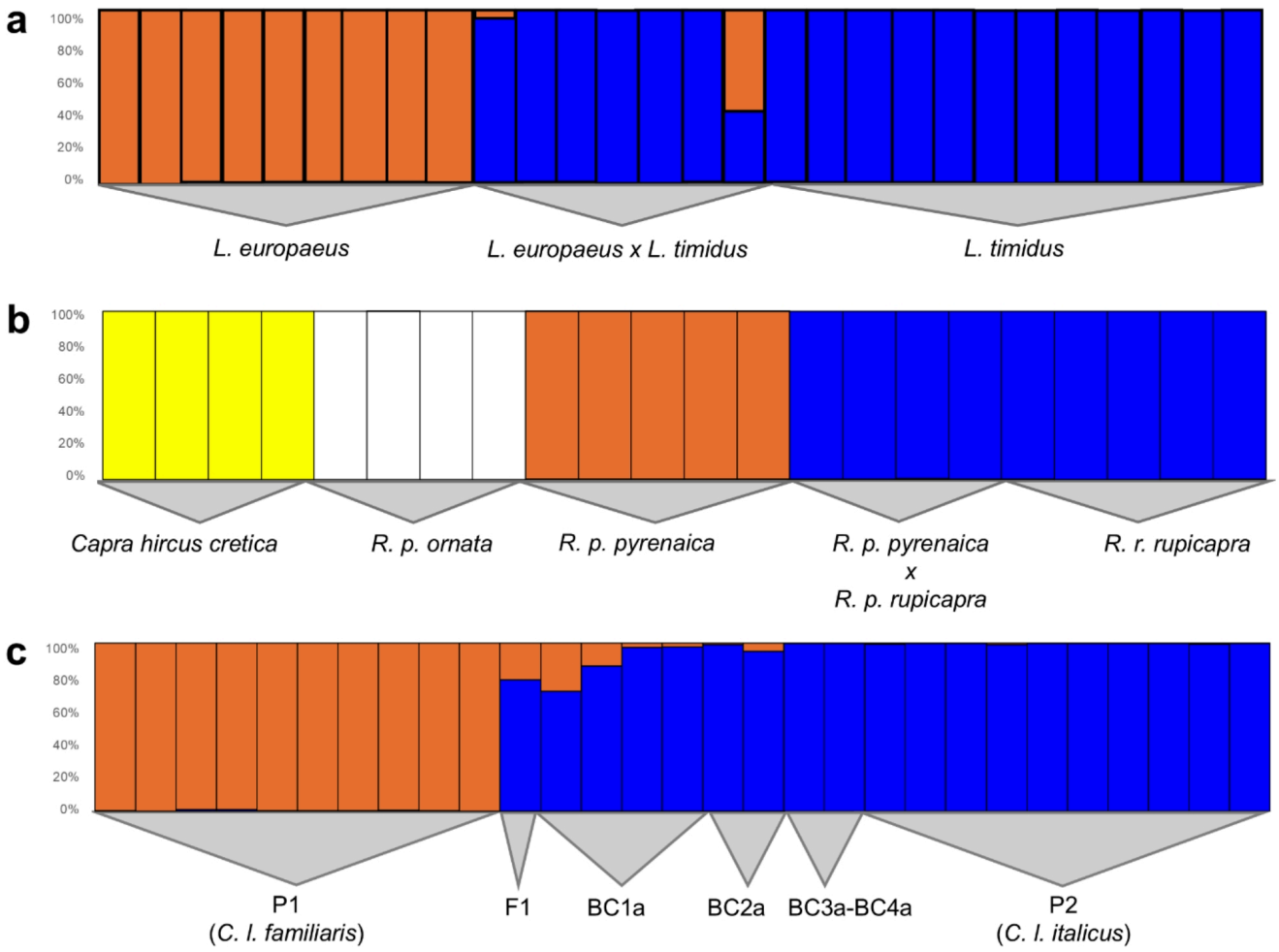


Figure 2

Results of Structure analysis performed on *Leporidae, Bovidae* and *Canidae* individuals. Results obtained for **a)** Leporidae, **b)** Bovidae and **c)** Canidae MIP genotypes generated in this study are shown. Hybrid classes are defined as follows: F1=P1xP2; BC1=F1xP2; BC2=BC1xP2; BC3-BC4=BC2xP2 where P1 and P2 represent the two parental species. Each panel shows the percent inferred ancestry obtained with the Structure software using K=2 (Leporidae and Canidae) and K=4 (Bovidae), i.e. the most likely according to the Evanno method. Each bar represents a different individual, while the y-axis shows the proportions of inferred ancestry for the two clusters identified by the software.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- MarinangelietalMIPsmammiferiSUPPLTablesOCT2026.xlsx
- MarinangelietalMIPsmammiferiSUPPLFiguresOCT2026.pptx