Aalborg Universitet

Applying genomic data in wildlife monitoring

Development guidelines for genotyping degraded samples with reduced single nucleotide polymorphism panels

Thaden, Alina; Nowak, Carsten; Tiesmeyer, Annika; Reiners, Tobias E.; Alves, Paulo C.; Lyons, Leslie A.; Mattucci, Federica; Randi, Ettore; Cragnolini, Margherita; Galián, José; Hegyeli, Zsolt; Kitchener, Andrew C.; Lambinet, Clotilde; Lucas, José M.; Mölich, Thomas; Ramos, Luana; Schockert, Vinciane; Cocchiararo, Berardino

Published in:
Molecular Ecology Resources

DOI (link to publication from Publisher):
10.1111/1755-0998.13136

Creative Commons License
CC BY-NC 4.0

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Thaden, A., Nowak, C., Tiesmeyer, A., Reiners, T. E., Alves, P. C., Lyons, L. A., Mattucci, F., Randi, E., Cragnolini, M., Galián, J., Hegyeli, Z., Kitchener, A. C., Lambinet, C., Lucas, J. M., Mölich, T., Ramos, L., Schockert, V., & Cocchiararo, B. (2020). Applying genomic data in wildlife monitoring: Development guidelines for genotyping degraded samples with reduced single nucleotide polymorphism panels. *Molecular Ecology Resources*, *20*(3), 662-680. https://doi.org/10.1111/1755-0998.13136

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Received: 5 September 2018 | Revised: 24 December 2019 | Accepted: 5 January 2020

DOI: 10.1111/1755-0998.13136

RESOURCE ARTICLE

MOLECULAR ECOLOGY RESOURCES WILEY

Applying genomic data in wildlife monitoring: Development guidelines for genotyping degraded samples with reduced single nucleotide polymorphism panels

Alina von Thaden[1,2] | Carsten Nowak[1,3] | Annika Tiesmeyer[1,2] | Tobias E. Reiners[1,2] | Paulo C. Alves[4,5,6] | Leslie A. Lyons[7] | Federica Mattucci[8] | Ettore Randi[9,10] | Margherita Cragnolini[11] | José Galián[12] | Zsolt Hegyeli[13] | Andrew C. Kitchener[14,15] | Clotilde Lambinet[16] | José M. Lucas[12] | Thomas Mölich[17] | Luana Ramos[4,5] | Vinciane Schockert[16] | Berardino Cocchiararo[1,3]

[1]Conservation Genetics Group, Senckenberg Research Institute and Natural History Museum Frankfurt, Gelnhausen, Germany

[2]Institute for Ecology, Evolution and Diversity, Johann Wolfgang Goethe-University, Frankfurt am Main, Germany

[3]LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Frankfurt am Main, Germany

[4]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO – Laboratório Associado, Universidade do Porto, Vairão, Portugal

[5]Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal

[6]Wildlife Biology Program, University of Montana, Missoula, MT, USA

[7]Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri, Columbia, MO, USA

[8]Area per la Genetica della Conservazione, ISPRA, Ozzano dell'Emilia, Bologna, Italy

[9]Department BIGEA, University of Bologna, Bologna, Italy

[10]Department 18/Section of Environmental Engineering, Aalborg University, Aalborg, Denmark

[11]Institut für Spezielle Zoologie und Evolutionsbiologie, Biologisch-Pharmazeutische Fakultät, Friedrich-Schiller-Universität Jena, Jena, Germany

[12]Departamento de Zoología y Antropología Física, Facultad de Veterinaria, Universidad de Murcia, Murcia, Spain

[13]"Milvus Group" Bird and Nature Protection Association, Tîrgu Mureş, Romania

[14]Department of Natural Sciences, National Museums Scotland, Edinburgh, UK

[15]Institute of Geography, School of Geosciences, University of Edinburgh, Edinburgh, UK

[16]Department of Science and Environmental Management, University of Liège, Liège, Belgium

[17]Landesverband Thüringen e.V., BUND für Umwelt und Naturschutz Deutschland (BUND), Erfurt, Germany

Correspondence
Alina von Thaden and Berardino Cocchiararo, Conservation Genetics Group, Senckenberg Research Institute and Natural History Museum Frankfurt, Gelnhausen, Germany.
Email: alina.vonthaden@senckenberg.de (A.T.); berardino.cocchiararo@senckenberg.de (B.C.)

Funding information
Karl und Marie Schack-Stiftung; Negaunee Foundation; University of California, Davis, Center for Companion Animal Health, School of Veterinary Medicine, Grant/

Abstract

The genomic era has led to an unprecedented increase in the availability of genome-wide data for a broad range of taxa. Wildlife management strives to make use of these vast resources to enable refined genetic assessments that enhance biodiversity conservation. However, as new genomic platforms emerge, problems remain in adapting the usually complex approaches for genotyping of noninvasively collected wildlife samples. Here, we provide practical guidelines for the standardized development of reduced single nucleotide polymorphism (SNP) panels applicable for microfluidic genotyping of degraded DNA samples, such as faeces or hairs. We demonstrate how

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2020 The Authors. *Molecular Ecology Resources* published by John Wiley & Sons Ltd

Award Number: 2011-30-F/M; Office of Research Infrastructure Programs (ORIP), Grant/Award Number: R24OD010928; Friedrich-Schiller-University of Jena; Cat Health Network, Grant/Award Number: D12FE-505; National Institutes of Health - National Center for Research Resources (NCRR), Grant/Award Number: R24 RR016094R24; Leibniz Association, Grant/Award Number: SAW-2011-SGN-3

microfluidic SNP panels can be optimized to efficiently monitor European wildcat (*Felis silvestris* S.) populations. We show how panels can be set up in a modular fashion to accommodate informative markers for relevant population genetics questions, such as individual identification, hybridization assessment and the detection of population structure. We discuss various aspects regarding the implementation of reduced SNP panels and provide a framework that will allow both molecular ecologists and practitioners to help bridge the gap between genomics and applied wildlife conservation.

KEYWORDS

conservation biology, degraded DNA, *Felis silvestris*, genetic noninvasive sampling, reduced SNP panel, wildlife genomics

1 | INTRODUCTION

The emergence of massive parallel sequencing technologies during the last decade has paved the way for a new era in conservation genetics (Allendorf, Hohenlohe, & Luikart, 2010; Desalle & Amato, 2004). Traditional methodologies, such as Sanger sequencing and microsatellite genotyping, are currently complemented by more genome-wide approaches, which involve both opportunities and challenges (McMahon, Teeling, & Höglund, 2014; Shafer et al., 2015; Taylor & Gemmell, 2016). The transition from genetic to genomic data promises unprecedentedly detailed insights into conservation-relevant processes and patterns, such as inbreeding, genome-wide allelic diversity, or historical and recent introgression (Allendorf et al., 2010; Frankham, 2010).

However, at present, the implementation of conservation genomic inferences on a broad scale and, more importantly, their provision and access for applied conservation practitioners remains at an exploratory or early stage (Garner et al., 2016; Shafer et al., 2016). The current gap between genomic approaches from basic research and applicable solutions for conservation managers has been stressed by many researchers (Britt, Haworth, Johnson, Martchenko, & Shafer, 2018; Hogg, Taylor, Fox, & Grueber, 2018; Shafer et al., 2015; Taylor, Dussex, & van Heezik, 2017). Difficulties occur owing to poor incentives for researchers to engage in applied conservation, insufficient communication between scientists and practitioners, and increased complexity of genomic analyses, demanding advanced bioinformatic expertise and access to high-performance computer clusters (Holderegger et al., 2019; Shafer et al., 2015; Taylor et al., 2017). To bridge this gap, conceptual and practical frameworks must be established, user-friendly analytical pipelines need to be developed, and successful case studies should be disseminated to practitioners (Benestan et al., 2016; Shafer et al., 2015).

Another constraint is the limited applicability of genomic approaches to noninvasively collected and forensic samples, as access to tissue material is often restricted when dealing with endangered and elusive taxa (Carroll et al., 2018; McCormack, Hird, Zellmer, Carstens, & Brumfield, 2013; McMahon et al., 2014; Russello, Waterhouse, Etter, & Johnson, 2015; Shafer et al., 2015). Various approaches have recently been developed to overcome this limitation, including the enrichment of targeted genomic regions (GT-seq, high-throughput sequencing of microsatellites, RNA bait capture) or the genotyping of preselected single nucleotide polymorphism (SNP) panels on array-based platforms (MassARRAY, Amplifluor and Fluidigm; for a comprehensive overview see Carroll et al., 2018).

Currently, the application of reduced SNP panels (hereafter referred to as rSPs) is becoming increasingly popular for standardized assessments of species of conservation concern (Fitak, Naidu, Thompson, & Culver, 2016; Henriques et al., 2018; Kraus et al., 2015; Li, Wei, Ma, & Chen, 2018). SNP markers are generally valued for their even distribution throughout the genome, ease of identification from increasingly available next-generation sequencing data, unambiguous genotype calling, and suitability for parallel detection of large marker numbers covering significant parts of the genome (Garvin, Saitoh, & Gharrett, 2010; Helyar et al., 2011; Morin, Luikart, Wayne, & the SNP workshop group, 2004). Although SNP genotyping may suffer some constraints (such as effects of ascertainment bias and increased error rates for some high-density SNP arrays; e.g., Hoffman et al., 2012; Lepoittevin et al., 2015; Quinto-Cortés, Woerner, Watkins, & Hammer, 2018; Wray et al., 2013), the facilitated harmonization of SNP markers across laboratories for collaborative applications and their suitability for genetic noninvasive sampling (gNIS) make rSPs a highly promising method for applied conservation genomic efforts (deGroot et al., 2016; von Thaden et al., 2017). Small panels of ~100 SNPs have already been used to estimate population size and landscape relatedness of brown bears in Sweden (Norman et al., 2017; Spitzer, Norman, Schneider, & Spong, 2016), to identify genetic stocks for management of endangered fish species (Baetscher, Hasselman, Reid, Palkovacs, & Garza, 2017; Starks, Clemento, & Garza, 2015) and to monitor introgressive hybridization in European wildcats (*Felis silvestris* S.) in the Swiss Jura (Nussberger, Currat, Quilodran, Ponta, & Keller, 2018; Nussberger, Wandeler, Weber, & Keller, 2014b). A recent study has shown that microfluidic genotyping of wildlife samples using rSPs has several advantages compared to traditional microsatellite genotyping, such as reduced replication needs and lower costs, when analysing large sample sizes (von Thaden et al., 2017). These properties make rSPs

a promising genotyping alternative for studies relying on noninvasively collected or degraded samples; however, the respective SNP assays need to be carefully validated early in their development in order to prevent failure of assays when applied to samples with limited DNA quality (von Thaden et al., 2017).

To facilitate this, we present here comprehensive guidelines on how to efficiently develop and optimize rSPs using microfluidic arrays for noninvasively collected wildlife samples. As a case study, we report the development of an rSP for standardized genetic assessments of the European wildcat. This elusive felid is an excellent example of an endangered species in need of continuous monitoring that is heavily reliant on gNIS (Beaumont et al., 2001; Mattucci et al., 2013; Say, Devillard, Léger, Pontier, & Ruette, 2012; Steyer et al., 2016; Steyer, Simon, Kraus, Haase, & Nowak, 2013).

In the present study, we demonstrate how microfluidic rSPs can: (a) be developed, optimized and applied efficiently with regard to costs and hands-on time and (b) be set up in a modular fashion to accommodate informative markers for population genetics questions, such as hybridization assessment or inference of population structure. We discuss various aspects of the implementation of rSPs and provide a framework that will allow both molecular ecologists and practitioners to help bridge the gap between genomics and applied wildlife conservation.

2 | MATERIAL AND METHODS

2.1 | Identification of SNP loci and SNPtype™ assay design

Informative SNP loci for the European wildcat were identified from 62,897 SNPs included in the cat 63K DNA array (Illumina Infinium iSelect 63K Cat DNA Array, Illumina Inc.; Gandolfi et al., 2018) by genotyping a selection of domestic cats (*Felis catus*, *n* = 45), European wildcats (*n* = 100) and known hybrids (*n* = 41) from across the entire European distribution range (Mattucci, 2014; Mattucci, Oliveira, Lyons, Alves, & Randi, 2016). The 192 loci in the present study were filtered from the resulting data set through quality and linkage disequilibrium pruning using PLINK (Mattucci, 2014; Purcell et al., 2007). The selected pool of 192 SNPs encompassed 142 loci for individual identification (maximized heterozygosity in wildcat populations; "ID" markers) and 50 loci for hybridization assessment (F_{ST}> 0.8 between domestic and wildcats; "HYB" markers). SNPtype assays were designed using the web-based D3 assay design tool (Fluidigm Corp.) based on sequences enclosing the respective target SNPs within at least 150-bp distance. The assays consist of one specific target amplification (STA) primer, two allele-specific primers (ASPs) and one reverse locus-specific primer (LSP).

2.2 | 96.96 Dynamic Array™ SNP genotyping

SNP genotyping was performed on 96.96 Dynamic Arrays (Fluidigm) with integrated fluidic circuits (IFCs, Wang et al., 2009). Dynamic Arrays harbour distinct nano-PCR wells in which each SNPtype assay and sample is combined separately. The genotyping PCR generates allele-specific fluorescence that is measured in an endpoint reader (EP1 reader, Fluidigm). Fluorescence data are subsequently analysed with the SNP Genotyping Analysis Software (version 4.1.2, Fluidigm) which generates allelic discrimination cluster plots (Figure S1) to determine the SNP genotype for each sample at each locus. Samples with DNA of low concentrations or poor quality undergo a pre-amplification PCR (STA) before genotyping. In STA PCR the 96 target regions containing the SNPs are enriched in a single 96-multiplex reaction to ensure that sufficient amounts of template are available. During the testing procedures, we modified the manufacturer's original STA protocol as well as the combination of multiplexed SNP assays (details below). The genotyping PCR was conducted according to the manufacturer's protocol. The fluorescence signal was measured after the standard genotyping PCR protocol (SNPtype 96 × 96 version 1) and after four additional cycles to ensure sufficient fluorescence strength for samples with low DNA concentrations. All experiments were run along with four NTCs (no template controls) per array to monitor for potential contamination.

2.3 | Sample sets and DNA extraction

2.3.1 | Reference samples

To identify the best performing SNPtype assays under wet laboratory conditions and most informative SNP markers for individual discrimination, population assignment and hybrid screening, we assembled a reference set of 92 high-quality wildcat tissue extracts originating from several European populations, concentrated buccal swab extracts from random-bred domestic cats, as well as two samples of African wildcats (*Felis lybica*; Table 1; "High-DNA"). To evaluate the panel's applicability for low-DNA samples, we used a second reference set (Table 1; "Low-DNA") consisting of hair samples (*n* = 92) collected noninvasively either during genetic monitoring for European wildcats (Steyer et al., 2016), or by plucking hairs from reference individual domestic cats. A subset of the reference individuals (domestic cats, *n* = 24; wildcats, *n* = 8) were represented simultaneously by high- and low-DNA samples (see the testing scheme below). Additionally, both hair and scat samples were available for 13 of the reference domestic cats, which allowed comparison of the genotyping performance for the two major types of noninvasive samples used in wildlife and conservation studies.

2.3.2 | Nonreference samples

To test the panel's applicability in genetic monitoring programmes, we genotyped additional noninvasively collected hair samples (*n* = 30) obtained during routine German wildcat monitoring (Table 1; "Monitoring"). These samples were randomly selected and comprised a putative broad range of DNA quantities and qualities. We further tested noninvasive samples from various other species (*n* = 20; named "cross-species" in the following), because gNIS may regularly involve

TABLE 1 Sample sets used for the testing scheme

Test phase	Development and optimization:		Assessment of applicability for:		
Test purposes	Test 1, Test 2, Test 3, DNA dilution series, STA dilution series	Test 3, STA cycles, sample types	Genetic monitoring programmes based on gNIS	Samples not previously included in rSP development	Susceptibility of markers to nontarget species
Sample sets	Reference samples:		Nonreference samples:		
	High-DNA	Low-DNA	Monitoring	European populations	Cross-species
	n	*n*	*n*	*n*	*n*
Domestic cats	24	24	–	–	–
Germany	24	24	–	–	–
European wildcats	66	68	–	180	–
Germany	32	68	30	25	–
Austria	2	–	–	–	–
Belgium	–	–	–	28	–
Bulgaria	4	–	–	–	–
Greece	4	–	–	–	–
Italy	4	–	–	26	–
Luxembourg	4	–	–	–	–
Portugal	4	–	–	11	–
Romania	4	–	–	28	–
Scotland	4	–	–	–	–
Spain	4	–	–	20	–
African wildcats[a]	2	–	–	–	–
Iran	1	–	–	–	–
Morocco	1	–	–	–	–
Other species[b]	–	–	–	–	*20*
Total genotyped *n*	92	92	30	180[c]	20
Sample types	Concentrated saliva (DC), tissue (WC)	Hair	Hair	Concentrated saliva (DC), tissue (WC)	Hair, saliva, scats
Results	Figures: 2, 4[d], 5[d], S1, S3, S8; Tables: S2–S5, S10	Figures: S2, S3; Tables: S2–S4, S7–S9	Figure 2; Tables: S2, S4, S6	Figures: 2, 4–6, Figure S4–S7, S9; Tables: S2, S4, S6, S12	Figure 3; Tables: S2, S4

Abbreviations: STA, specific target amplification; gNIS, genetic noninvasive sampling; rSP, reduced SNP panel; *n*, number of individuals; DC, domestic cats; WC, European wildcats.

[a]Two samples of African wildcats (*F. lybica*) were initially included to test for polymorphisms of the markers in other wildcat species.

[b]Including samples of 18 different European mammal nontarget species.

[c]Some of these samples were excluded during analyses steps due to high missing data (n = 3), duplicate individuals (n = 4), or assignment values for wildcat clusters of $q^{(i)} \leq 0.85$ in STRUCTURE (n = 35). See text for more details.

[d]Only a selection of domestic cat samples, respectively.

samples from nontarget species (Table 1; "Cross-species"). Finally, we genotyped tissue samples (n = 180) from different populations of European wildcats to confirm the applicability of the marker set for samples that were not previously included in the development phases (i.e., were not reference samples; Table 1; "European populations").

We extracted DNA from tissue samples using the DNeasy Blood & Tissue Kit (Qiagen), the buccal swabs and hair samples using the QIAamp DNA Investigator Kit (Qiagen) and scats using the QIAamp DNA Stool Mini Kit (Qiagen). For DNA extraction from hairs, we used a minimum of 10 hairs with roots per sample (Goossens, Waits, & Taberlet, 1998; Steyer et al., 2013). To create high-DNA reference sample extracts from domestic cats, we pooled DNA extracts from five buccal swabs per individual, employing the clean-up of genomic DNA protocol (QIAamp DNA Micro Kit, Qiagen) to obtain DNA concentrations of >60 ng/µl. The pooled buccal extracts and all tissue extracts for the high-DNA

reference sample plate were quantified using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific) and adjusted to a DNA concentration of ~60 ng/µl for the first tests. No animals were harmed or killed for this study and all samples were collected in compliance with the respective local and national laws (details in Supporting Methods).

2.4 | Testing scheme

To develop an efficient 96 SNP panel for standardized genetic assessments of European wildcats from the initial pool of 192 SNPtype assays, we followed a sequential testing scheme (Figure 1; Table S1).

2.4.1 | Test 1: Initial evaluation of in silico designed SNPtype assays under wet laboratory conditions

We first tested all 192 assays partitioned into two genotyping PCR runs using the high-DNA reference sample set (92 samples, 60 ng/µl DNA, Table 1). This allowed us to evaluate the technical marker performance while avoiding bias due to diverging sample performance (e.g., resulting from varying purity and integrity of DNA extracts). To exclude assays showing poor performance (e.g. exhibiting insufficient fluorescence or producing ambiguous genotype clusters), STA pre-amplification was not employed at this step. SNP markers that (a) produced the clearest genotype clusters (i.e., the two homozygotes and the heterozygotes unambiguously assigned to three distinct clusters; compare Figure S1) and (b) showed either high heterozygosity within European wildcat samples (n = 103 ID markers) or (c) high F_{ST} between domestic cats and European wildcats (n = 35 HYB markers) were selected for the next wet laboratory testing phase, resulting in 138 candidate SNP markers (72% of the initial pool of 192 SNPs).

Furthermore, because the selection of the initial pool of 192 SNPs was based on samples from another project (Mattucci, 2014; Mattucci et al., 2016), and respective genotypes were generated on a different platform (Illumina Infinium iSelect DNA Array, Illumina Inc.), we verified them again based on the reference samples in this study (Table 1) and the genotyping platform applied here (96.96 Dynamic Arrays, Fluidigm). To do so, we assessed the informative value of these 138 preselected markers for hybrid detection and population structure using the Bayesian clustering algorithm implemented in NEWHYBRIDS version 1.1 beta (Anderson & Thompson, 2002) and STRUCTURE (Pritchard, Stephens, & Donnelly, 2000). We chose subsets of 32, 24, 16 and eight of the pool of 35 HYB markers by sequentially excluding those with heterozygotes (i.e., markers that were not fixed for domestic or wildcats). These subsets were tested for their capability of distinguishing domestic cats from European wildcats, in order to estimate the number of HYB assays required for the final rSP. Subsequently, we tested if subsets of 96, 88 and 84 of the selected 103 ID markers would

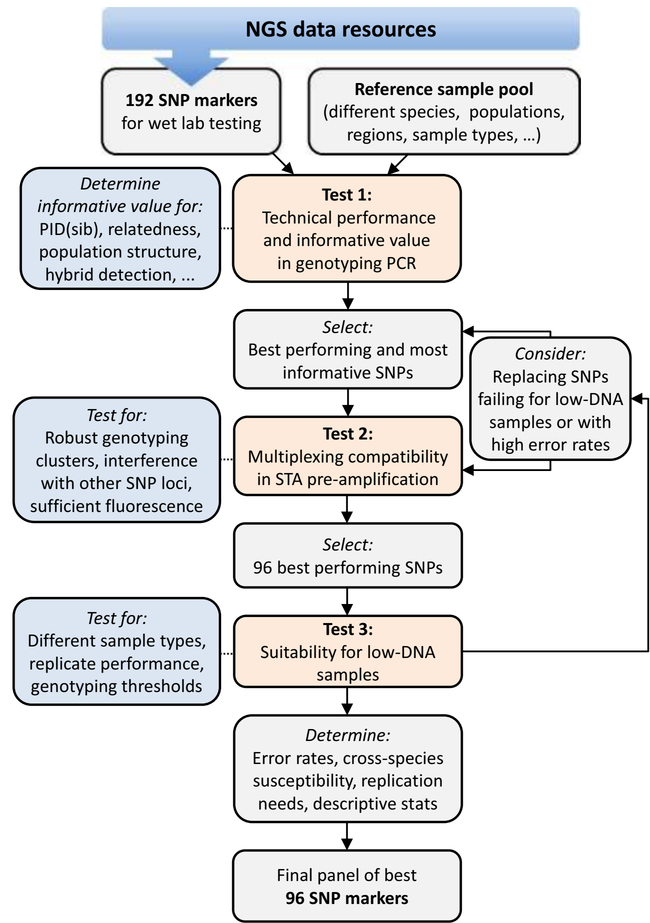


FIGURE 1 Overview of testing scheme. Flow diagram of testing and selection steps recommended for assembling a novel microfluidic SNP marker panel. Further information about the various stages is provided in the text and Supporting Information

suffice to detect population structures among the samples in our reference set, provided that at least eight HYB assays would be included in the final rSP.

2.4.2 | Test 2: Testing multiplexing compatibility in STA pre-amplification

The pre-amplification of target regions (STA) is particularly crucial for successful genotyping of low-DNA or degraded samples. We tested the multiplexing compatibility of the 138 candidate SNPs that had passed Test 1 in a high-multiplexed approach (138plex). A 5 ng/µl dilution of the high-DNA reference sample set was used as template because STA pre-amplification is advisable for samples with low DNA concentration or low copy number (Kraus et al., 2015). This allowed us to identify assays that fail or show diminished performance in genotyping PCR due to ineffective target enrichment (e.g., because of primer interference during multiplexing or inaccurate in silico STA primer design). For subsequent tests, we consequently applied 96 of the most promising markers from the 138 candidate assays.

2.4.3 | Test 3: Marker suitability for low-DNA samples

To determine the minimum DNA concentration threshold for reliable SNP genotyping and to identify unstable markers that fail at low DNA concentrations, we further diluted the high-DNA reference sample set to concentrations of 1, 0.2 and 0.1 ng/μl, respectively. We genotyped these three dilutions using the 96 markers selected after Test 2. Additionally, we tested the influence of several parameters (see below and Table S2) of the STA protocol for their benefit on genotyping performance of low-DNA samples (on the basis of results in von Thaden et al., 2017). All tests concerning STA protocol parameters were conducted based on the diluted reference sample set with 0.2 ng/μl DNA, as we identified this concentration to be the detection limit (according to the findings in Nussberger, Wandeler, & Camenisch, 2014a). We tested (a) different dilutions of STA PCR products as template for the subsequent genotyping PCR and (b) raising the cycle numbers in STA reactions from 14 to 18 and 28 cycles. Finally, to verify the suitability of the 96 markers for genotyping of noninvasively collected samples, we genotyped our reference set of domestic and wildcat hair samples (Table 1; Low-DNA).

2.4.4 | Substituting failing markers for the final 96 panel

During several rounds of optimization, we substituted 11 SNP markers that had failed in genotyping of low-DNA samples (Test 3). After the first replacement of assays, one of the newly included failed to amplify and was again substituted with another assay. We continued this procedure of replacing single failing assays (four rounds in total) until we ended up with a combination of 96 markers that showed stable performance for low-DNA samples. As substitutes we used nine of the remaining 42 SNP candidate assays that had previously passed Tests 1 and 2. Additionally, we included two assays targeted to the SRY gene for sex determination (called "SRY" SNPs in the following), taken from the wildcat rSP of Nussberger, Wandeler, and Camenisch (2014a). All SNP genotyping runs in this test flow were performed on noninvasively collected hair samples (Table 1; "Low-DNA"). The final panel of 96 SNPs consisted of 84 ID, 10 HYB and 2 SRY SNPs.

2.5 | Statistical data analyses

2.5.1 | Validation of genotypes and error estimation

The evaluation of tests during the development of the rSP relied on the genotyping of high- and low-DNA reference sample sets (Table 1). For these reference individuals, we assumed the genotypes generated from high-DNA extracts (60 ng/μl) as true genotypes (compare Foerster et al., 2017). Throughout the testing scheme, corresponding low-DNA and noninvasive samples of reference individuals were replicated between runs and compared to the high-DNA genotypes. Thus, the same reference individuals were used to evaluate different treatments or protocols to exclude sample bias. Furthermore, the comparison of replicated low-DNA to respective high-DNA genotypes of reference samples allowed us to identify the best genotyping protocol and derive the most adequate consensus-building rules to minimize genotype errors (definition below).

Because the genotypes of nonreference samples were previously unknown (not included in wet laboratory optimization tests), these samples were duplicated (tissue) or triplicated (hair). To build consensus genotypes of replicated nonreference samples, we assumed (following the inferred consensus-building rules) the most common allele to be true; also, if two replicates showed consistent homozygous genotypes while a third replicate showed a heterozygous genotype. If one of the replicates was heterozygous and the other two replicates showed opposing homozygous genotypes, we assumed the genotype to be heterozygous.

Consensus genotypes of nonreference samples were considered as belonging to the same individual based on consistent multilocus genotypes. We allowed for up to six mismatches (5.8% dissimilarity) between two consensus genotypes assigned to the same individual identity (Nussberger, Wandeler, Weber, et al., 2014b). Other relevant information, such as microsatellite genotypes, sampling date, locality, mitochondrial DNA (mtDNA) haplotype and sex, were also taken into account, if available. Samples with SNP call rates <80% were excluded from further analyses.

Error rates were calculated for two types of genotyping errors, namely allelic dropouts (ADO) and false alleles (FA; compare Broquet & Petit, 2004; Taberlet, Waits, & Luikart, 1999). The ADO rate was defined as the sum of false homozygous scores (as compared to the corresponding individual's high-DNA or consensus genotype) divided by the sum of the individual's heterozygous loci. The FA rate was defined as the sum of false heterozygous and false opposite homozygous scores divided by the sum of the individual's homozygous loci.

Errors were calculated for separate sample sets, or subsets thereof (Table S3), to test different sample properties (e.g., DNA concentration or sample type) and genotyping protocol treatments (e.g., STA cycles or STA PCR product dilutions). Differences in error rates and missing data were tested for statistical significance using the *kruskal.test()* and *wilcox.test()* in R (R Development Core Team, 2009) after ascertaining nonnormal distribution for most of the data sets using the *shapiro.test()*. Furthermore, we determined amplification of putative nontarget noninvasive samples (18 cross-species; Table 1) that are likely to be detected on valerian-treated hair traps or were processed in the same laboratory in the course of other projects. We defined the call rate (amplification success) as 1 minus the number of loci with no genotype score ("no call") divided by the sum of all heterozygous, homozygous and no call scores for that locus or sample.

2.5.2 | Probabilities of identity and kinship analysis

To assess the power of the final 84 ID markers, we estimated the probabilities of identity (PID) and probabilities of identity among siblings

(PIDsib), according to Waits, Luikart, and Taberlet (2001) using GENALEX version 6.501 (Peakall & Smouse, 2012), and ranking the loci for highest expected heterozygosity (H_E). We further tested the ability of the ID markers to assign kinships for related individuals using the software ML-RELATE (Kalinowski, Wagner, & Taper, 2006). We calculated kinship assignments for the reference group of domestic cats (n = 24) and compared the results with the known kinship information (family pedigree) of a subset of five reference domestic cats.

2.5.3 | Hybridization assessment

We used the subset of 10 HYB markers to detect possible hybrids and domestic cats within our sample of 173 nonreference individuals from European wildcat populations. Identification of the parental groups as well as potential admixed individuals was based on the Bayesian clustering methods implemented in STRUCTURE (Pritchard et al., 2000) and NEWHYBRIDS version 1.1 beta (Anderson & Thompson, 2002), respectively. In STRUCTURE, we performed 500,000 Markov chain Monte Carlo runs after a burn-in of 250,000 steps for 10 replicated runs of K = 1–5 under the admixture model with correlated allele frequencies. The most likely K value was selected based on the Evanno method (Evanno, Regnaut, & Goudet, 2005) implemented in STRUCTURE HARVESTER (Earl & vonHoldt, 2012) after combining the replicates using the LARGEKGREEDY algorithm of CLUMPP (Jakobsson & Rosenberg, 2007). NEWHYBRIDS was used to estimate the posterior probabilities of the individuals belonging to each of six genealogical classes, including parental and hybrid categories, with an initial burn-in of 100,000, followed by 200,000 sweeps under uniform prior (additional testing of HYB SNP subsets see Supporting Methods). Subsequently, all identified domestic cats and putative hybrids were excluded from further analyses (assignment value for wildcat clusters $q^{(i)} \leq 0.85$).

2.5.4 | Analyses of genetic differentiation and population structure

We used a Discriminant Analysis of Principal Components (DAPC) in the ADEGENET version 2.1.1 R package (Jombart, 2008) to test the capability of the 84 ID markers to differentiate between European wildcat populations. Groups for DAPC were assigned according to country of origin. The variable contributions of the 84 loci in discriminating the populations were assessed and weights of each locus were illustrated in a loading plot. Population genetic clusters were estimated with STRUCTURE (Pritchard et al., 2000), using the same settings as in hybridization assessment, but for a larger range of K = 1–10. Genetic differentiation between the populations was further estimated based on pairwise F_{ST} values (Weir & Cockerham, 1984) calculated in ARLEQUIN version 3.5 (Excoffier & Lischer, 2010) and based on 5,000 permutations. Additionally, we calculated descriptive population genetics statistics and deviations from Hardy–Weinberg equilibrium (HWE), as well as linkage disequilibrium (LD), for the final 96 SNP panel (Supporting Methods).

3 | RESULTS

3.1 | Setting-up a 96 SNP panel optimized for degraded samples

The initial verification of the wet laboratory performance (Test 1; Figure 1) of 192 designed SNPtype assays allowed us to immediately exclude 28 assays (15%), as these failed to amplify or produced no interpretable genotype clusters under optimal genotyping conditions (high-DNA reference set, no pre-amplification influence). For the remaining 164 assays, we found apparent differences in the qualities of the allelic discrimination cluster plots. To avoid genotyping errors deriving from scoring mistakes, we excluded an additional 26 assays (14%) that showed ambiguous clustering formation (compare Figure S1). The residual 138 assays (103 ID and 35 HYB) were screened for their informative values and tested for their multiplexing compatibility (Test 2), where 135 assays (98%) showed robust multiplexing performance. At this stage, we set up a first panel of 96 SNPs, including the best 88 ID and eight HYB markers regarding clarity of cluster plots and informative values (i.e., maximized heterozygosity and F_{ST}, respectively). Genotyping of noninvasive samples and the diluted reference samples (Test 3; compare Table 1) revealed diminished performance (e.g., amplification failure or ambiguous clusters; Figure S1) for 5–17 assays (5%–18%) of the first 96 SNP panel with low-DNA samples (Figure 2). In the following tests, we substituted 11 of these assays by adding two HYB and seven ID assays from the 138 selected candidate SNPs, as well as two Y-linked assays (SRY SNPs; from Nussberger, Wandeler, & Camenisch, 2014a). The resulting final SNP panel consisted of 84 ID markers for individual and population identification, 10 HYB markers for hybridization detection and two SRY markers for sex determination.

3.2 | Validation of genotyping performance

3.2.1 | Effects of genotyping protocol modifications on error rates and missing data

DNA template concentrations of 5 or 1 ng/µl resulted in genotyping error rates of <1% (Figure 2). Lower concentrations increased mean ADO rates to 3%–8% and mean FA rates to 1%–2% with the diluted reference sample set (p < .001, Table S4). Similarly, the rate of missing data continuously increased with decreasing DNA concentration (p < .01 for 5–0.2 ng/µl, Table S4), up to a mean of 20% missing data per run for a concentration of 0.1 ng/µl DNA. Tissue sample extracts with DNA concentrations of 2–98 ng/µl showed negligible error rates of <0.1% (ADO and FA) and missing data of <1%. The randomly selected nonreference hair samples obtained during routine wildcat monitoring in Germany (Table 1) showed mean ADO rates of 11% and mean FA rates of 1%, while the total mean of missing data was 6% (with p < .001 between invasive and noninvasive samples, Table S5). The type of noninvasively collected samples (hairs, scats) marginally affected missing data and error rates, with scat samples generally

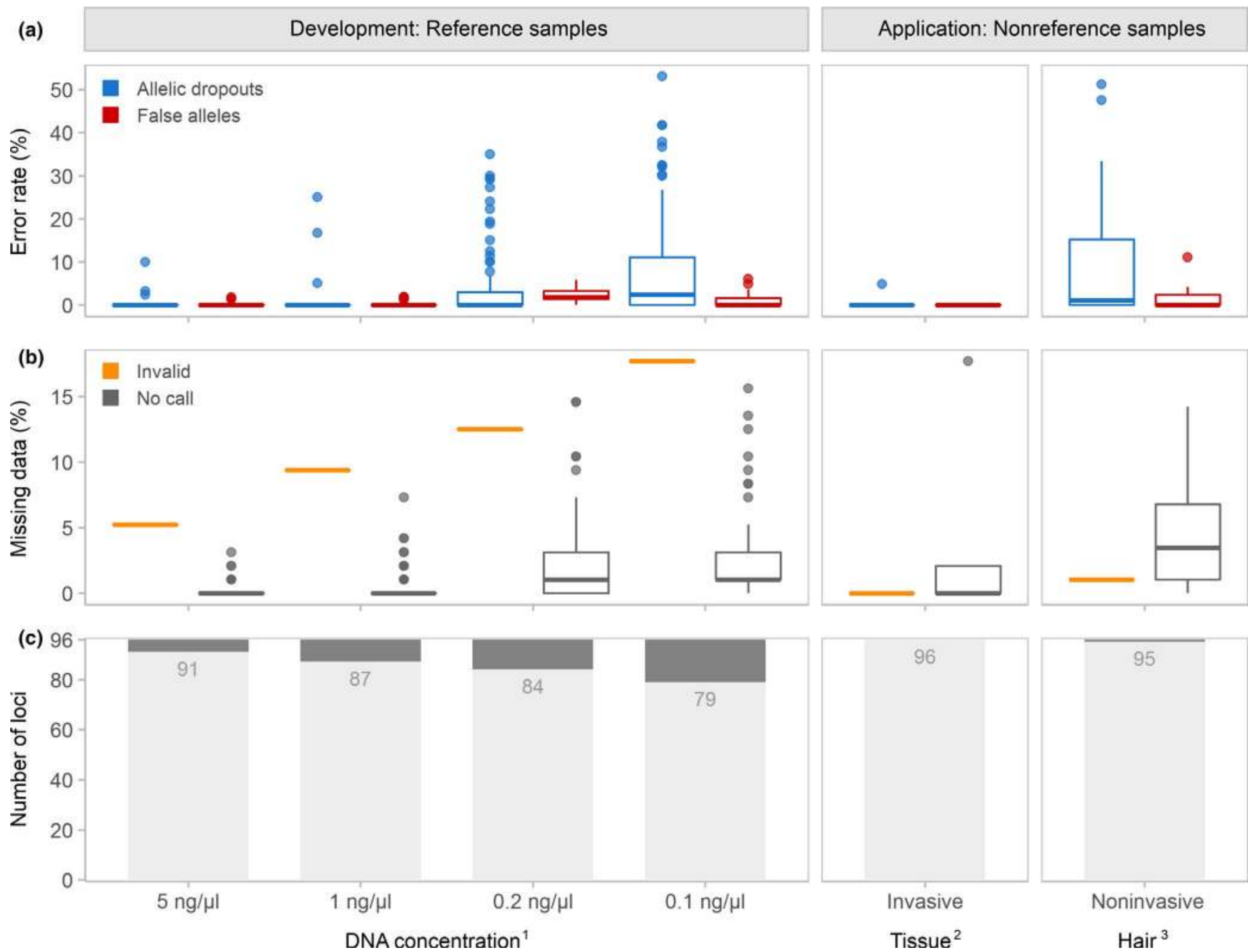


FIGURE 2 Genotyping performance across dilution series and sample types. Boxplots display the counts of errors or missing data over sample sets across all scorable SNP loci. The bars within boxplots represent medians, while data points show single samples that appear as outliers. (a) Error rates: allelic dropouts (blue) and false alleles (red). (b) Rates of missing data resulting from invalidation of whole SNP loci or whole sample genotypes during scoring (orange) or lack of genotype signal for single samples at single loci, i.e. no calls (grey). (c) Number of loci that produced scorable genotype data (light grey, numbers) or failed in genotyping (dark grey). [1]*n* = 92 reference samples (high-DNA, Table 1) with known genotypes; [2]*n* = 90 duplicated tissue samples from various European wildcat populations; [3]*n* = 30 triplicated hair samples from genetic wildcat monitoring in Germany; a single failing marker was later successfully substituted for the final panel

performing slightly better than hairs, albeit not significantly (mean ADO: 4% vs. 15%, mean FA: 0% vs. 0.7%, respectively; Figure S2 and Table S6). Sex identification showed disagreements between the two SRY markers in 2% of the genotypes (eight out of 354 tissue samples).

When adjusting parameters of the STA pre-amplification protocol, we found that increasing the cycle numbers to 18 or 28 cycles had a positive effect on the number of analysable SNP markers, as well as fewer missing data and lower or comparable error rates (Figure S3 and Tables S7–S9). With 28 cycles in STA we were able to analyse all 96 SNP loci in our final SNP panel, while the rate for missing data was 3%, ADO rate 8% and FA rate 1% for hair samples from reference individuals. Increasing the dilution ratios of STA PCR products from 1:10 to 1:20 or 1:40 led to an increase in errors and missing data (mostly $p < .05$, Table S9), while a lesser dilution of 1:5 resulted in lower FA rates on the one hand ($p < .001$, Table S9), but more missing data (invalidated loci) due to a higher prevalence of ambiguous genotype clusters (Figure S3). Thus, in our experience, an STA product dilution ratio of 1:10 gave the best genotyping performance.

3.2.2 | Species specificity

Cross-species testing revealed various degrees of unspecific amplification for all 18 nontarget species (Figure 3). Call rates were below 20%, except for Eurasian lynx (*Lynx lynx*; 95% call rate), which represented the only other felid species tested. However, genotyping of three lynx individuals showed almost consistent genotypes (95% identity), indicating low polymorphic content of the 96 SNP markers for this species. While 21 of the 96 loci showed amplification of ≤20% for all nontarget species, three assays (GTA0099295, GTA0099301 and GTA0099323) amplified for ≥60% of the tested cross-species (all

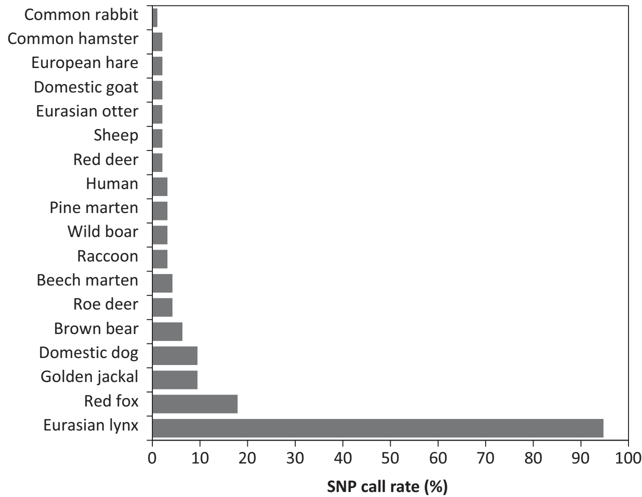


FIGURE 3 SNP call rates (amplification success) for a selection of 18 European mammal nontarget species

homozygous calls, hardly different alleles between nontarget species). The two African wildcat samples solely included in the high-DNA reference set amplified as well as European wildcat samples and showed domestic cat-like genotypes.

3.3 | Assessing applicability of the final SNP panel

After excluding three samples with missing data >20%, we used non-reference samples (n = 177) from different populations of European wildcats (Table 1) to assess the performance of the marker set in individual discrimination, hybridization assessment and population structure.

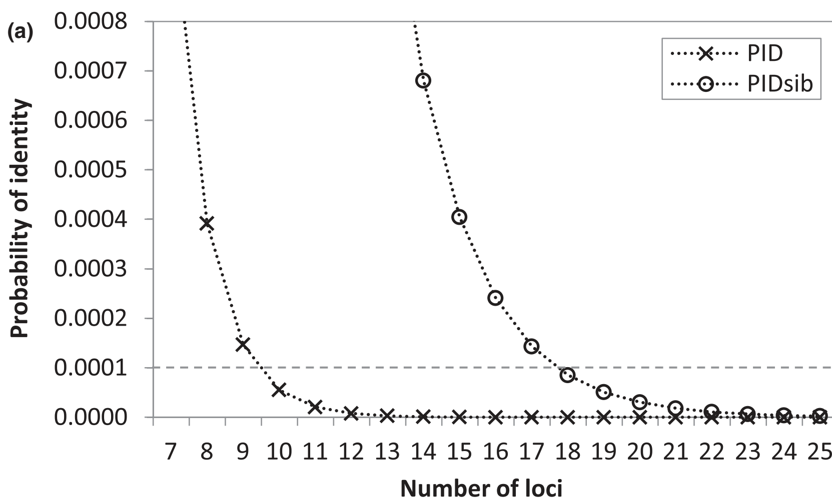
3.3.1 | Individual discrimination, kinship analysis and descriptive statistics

In our European-wide sample set, we identified three individuals from Portugal that were represented by multiple tissue samples (due to repeated sampling of the same carcasses), reducing the data set to 173 individuals. An estimated probability of identity of < 0.0001 (less than 1 in 10,000) was reached with a combination of 10 SNPs for PID or 18 SNPs for PIDsib, when using the most heterozygous ID SNPs (Figure 4a), consistent with previous findings for different species (grey wolf, brown bear) and rSPs, respectively (von Thaden et al., 2017). Kinship analysis of the 24 reference domestic cats allowed us to reconstruct known relationships correctly based on ID SNP genotypes (compare example in Figure 4b).

Measures of allele and genotype frequencies confirmed Mendelian inheritance and high information content of the loci (Table S10 and Supporting Results). HWE calculations based on the 84 ID markers revealed sporadic deviations of some loci at the population level (Figure S4). We found no clear evidence for the presence of linked loci (Figure S5), except for the population with the smallest sample size (n = 11, Portugal; Figure S6). All identified domestic cat and putative hybrid individuals were excluded from these analyses (see below).

3.3.2 | Admixture analysis and population structure

Genetic admixture analyses of the European-wide individuals (n = 173) using the 10 HYB markers selected for high F_{ST} between domestic and wildcats indicated several domestic cats (GB, n = 1; ES, n = 1; BE, n = 1) as well as wildcats with an admixed genotype (GB, n = 10; PT, n = 5; ES, n = 8; RO, n = 2; IT, n = 3; BE, n = 1; DE, n = 2; Figure 5). Results from



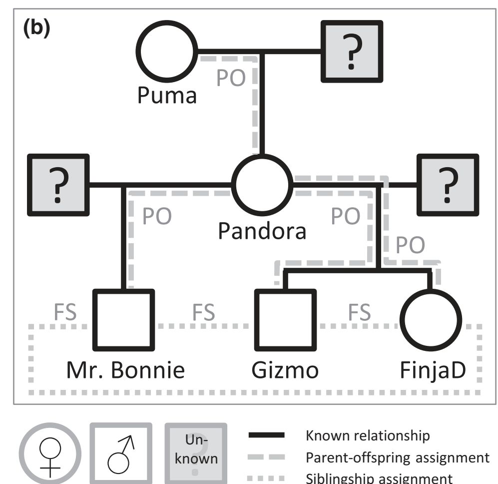


FIGURE 4 Power of the ID SNP panel to distinguish individuals (a) and reconstruct kinships (b). (a) Relationship between the number of genotyped SNP loci and probability of identity (PID) and probability of identity between siblings (PIDsib). Loci were ranked according to highest heterozygosity (H_E). A cut-off of 0.0001 was used because it is considered as sufficiently low for most applications involving natural populations (Waits et al., 2001). (b) Assignments of parentage or siblingship as calculated with ML-RELATE (Kalinowski et al., 2006) compared to known pedigrees of a domestic cat family. Circles represent females and squares represent males. Shaded symbols represent individuals not known or sampled. Assignments for single parent–offspring relationships are highlighted with grey dashed lines, and sibling relationships with dotted lines. PO, parent–offspring; FS, full siblingship

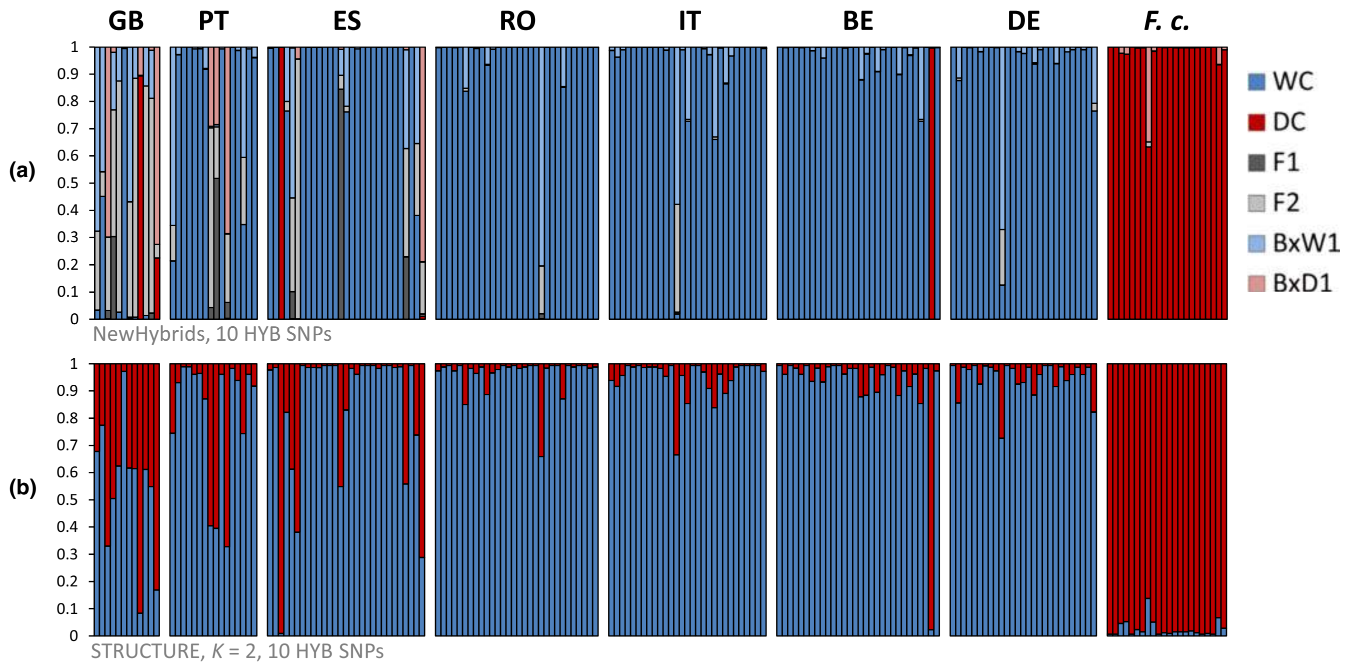


FIGURE 5 Identification of parental and hybrid individuals of wildcat and domestic cat using NEWHYBRIDS (a) and STRUCTURE (b). n = 173 wildcat individuals sampled from seven European wildcat populations and n = 22 random-bred reference domestic cats were analysed with 10 HYB SNPs selected to maximize F_{ST} between both parental groups. GB, United Kingdom; PT, Portugal; ES, Spain; RO, Romania; IT, Italy; BE, Belgium; DE, Germany; *F.c.*, reference domestic cats. WC, European wildcat; DC, domestic cat; F1, domestic × wildcat; F2, $F_1 \times F_1$; BxW1, F_1 × wildcat; BxD1, F_1 × domestic cat

Bayesian clustering methods implemented in NEWHYBRIDS and STRUCTURE were congruent. All domestic cats and admixed individuals with assignment values for wildcat clusters of $q^{(i)} \leq 0.85$ (n = 34) and the single nonadmixed Scottish wildcat were subsequently excluded from the European-wide data set.

Analyses of genetic population structure with DAPC and STRUCTURE based on 84 ID SNPs revealed five clusters largely corresponding to the different regions of European wildcat distribution (Iberian Peninsula [yellow and orange; PT, ES], Western German-Belgian cluster [red; DE, BE], Central German cluster [green; DE], Carpathians [blue; RO], Italy [magenta; IT]; Figure 6). Both methods detected similar patterns, although the population samples of Portugal and Spain were detected as one cluster in STRUCTURE but not in DAPC (Figure 6). A loading plot of marker weights revealed that 22 SNP markers have similarly high informative values for population discrimination, demonstrating that the differentiation is due to allele frequency differences at 26% of the loci (Figure S7). Differentiation of the populations was corroborated by pairwise F_{ST} values ranging from 0.03 to 0.371 (Table S11).

4 | DISCUSSION

4.1 | De novo design, development and optimization of SNP marker panels

The genomics era has been accompanied by an ever increasing availability of whole genomes generated from high-throughput sequencing for a plethora of species. Declining sequencing costs, as well as journals demanding full data accessibility for publications, further promote this trend. Hence, comprehensive genomic data (e.g., whole genomes, restriction site associated DNA sequencing, expressed sequence tags, SNP arrays of several 100k, genotyping by sequencing data) for mining of informative SNP loci for various research questions have become readily available.

An initial pool of candidate SNP loci for assay design may be filtered from genomic data according to the respective application purposes, such as high heterozygosities for individual identification (Kraus et al., 2015; Stetz et al., 2016), maximized F_{ST} between species for hybridization or introgression assessments (Henriques et al., 2018; Nussberger, Greminger, Grossen, Keller, & Wandeler, 2013; Pritchard et al., 2016), and high minor allele frequencies for relatedness estimations (Andrews et al., 2018; Baetscher et al., 2017; Zhao et al., 2018), or for sex determination (Katzner et al., 2017; Norman, Street, & Spong, 2013; Nussberger, Wandeler, & Camenisch, 2014a).

If appropriately designed SNP assays have already been verified by previous studies, assays can also easily be ordered and recombined in a novel panel, as we have shown in this study with Y-linked assays for sex determination (developed by Nussberger, Wandeler, & Camenisch, 2014a, on the basis of Luo et al., 2007). SNP genotyping with the Fluidigm system allows the development of marker panels in a modular fashion, where subsets of loci can be integrated for assessing different biological and demographic parameters, such as sex, individual identification and hybridization assessment (examples of modular panels in: DeWoody et al., 2017; Doyle et al., 2016;

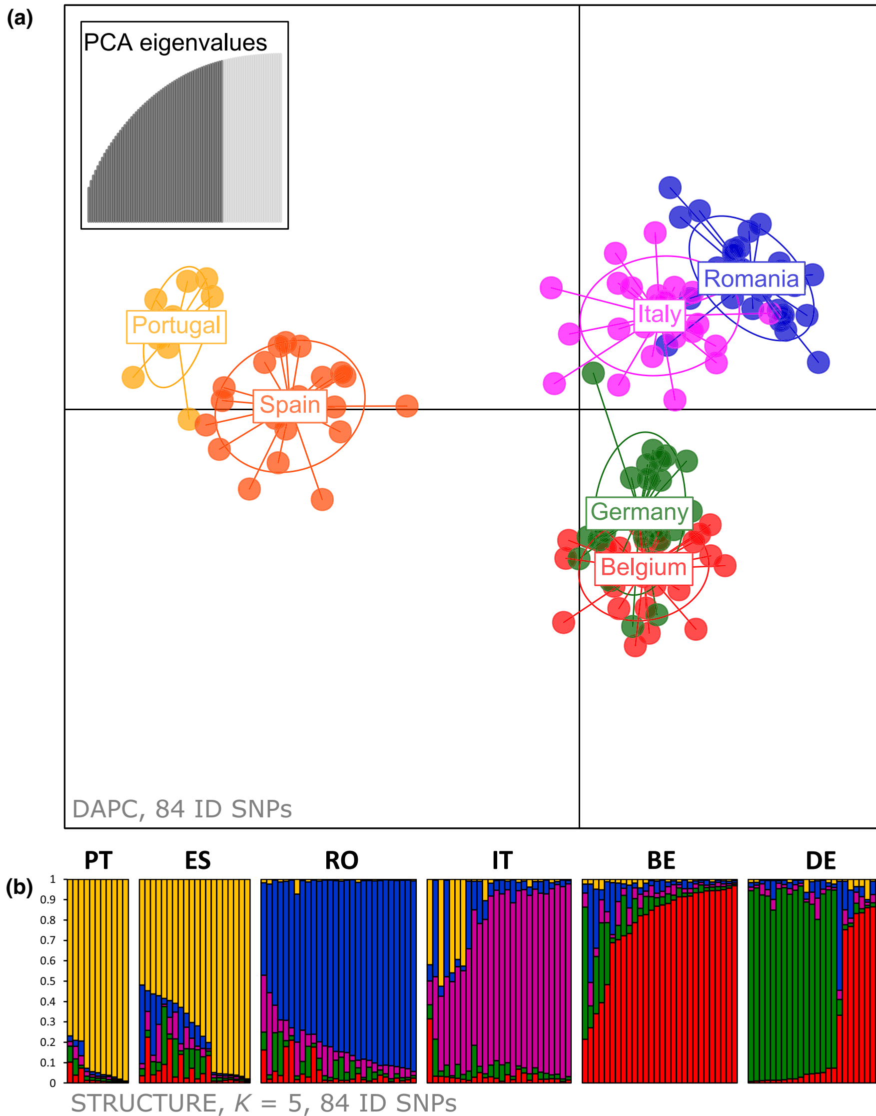


FIGURE 6 Identification of population structuring using Discriminant Analysis of Principal Components (a) and Bayesian clustering (b). Analyses based on 84 ID SNPs selected for individual identification and tested for n = 138 individuals from six European wildcat populations. (a) Groups were assigned according to geographical sample origins. (b) STRUCTURE plot for the most likely K as calculated with the Evanno method. PT, Portugal; ES, Spain; RO, Romania; IT, Italy; BE, Belgium; DE, Germany

Nussberger, Wandeler, & Camenisch, 2014a; Spitzer et al., 2016). The opportunity to generate multifaceted genetic or genomic information about a population is especially favourable for conservation monitoring, where scientists and practitioners alike aim to assess the status of a threatened species or its populations (Carroll et al., 2018; Flanagan, Forester, Latch, Aitken, & Hoban, 2018; Leroy et al., 2018).

The development of reliable genetic markers has always required sophisticated workflows and the use of designated reference samples. However, a practical guide for the straightforward development of rSPs, using microfluidic arrays for noninvasively collected samples, has been missing. Following our sequential testing scheme, we were able to develop and optimize a novel 96 SNP panel, which allows the assessment of several relevant demographic and ecological questions, even when using noninvasively collected hair and faecal samples. If SNP marker panels are not strictly optimized for stable performance with noninvasive samples early in wet laboratory development, many loci may have to be omitted in subsequent analyses when applied to low-DNA or degraded samples. In our experience, between 10% and 15% of the SNP assays fail when applied to samples of low DNA quality and/or quantity (von Thaden et al., 2017). Our proposed workflow for SNP panel development and optimization (Figure 1) minimizes the failure of markers at this critical step through prior testing of general technical performance and high-multiplexing compatibility. First, by filtering for SNP markers that are stable in genotyping PCR and have high informative value for the desired application (Test 1), unsuitable loci are immediately excluded from further wet laboratory testing. This strategy prevents laborious and unprofitable optimization procedures for a priori weakly performing markers and ultimately saves a lot of hands-on time and laboratory costs. Second, well-performing markers are tested for their high-multiplexing compatibility by including all candidate loci (≥96) in one multiplex (Test 2). This filtering step allows identification of markers that are principally suitable for common multiplexing and thus provides a pool of candidate markers (96+) for further development.

Despite these precautions, single SNP markers may fail in genotyping samples with low or degraded DNA. Failure of apparently

well-performing markers could result from sensitivity towards inhibitors, enhanced formation of primer dimers in the absence of high-quality DNA or suboptimal design of assay primers (e.g., binding site mismatches). Thus, comprehensive testing of the 96 SNP shortlist of best performing markers (from Test 2) for different noninvasive sample types (e.g., faeces, saliva, hairs, feathers) or other degraded samples (e.g., historical and museum samples) is fundamental (Test 3). Importantly, at this step, reference samples should include low-DNA samples from the same individuals previously tested with high DNA, to enable accurate evaluation of the performance of the loci and allow identification of artefacts. If needed, single malfunctioning loci or markers with highest error rates may be replaced by surplus candidate loci.

Introducing novel genomic methodologies implies an initial monetary investment as well as hands-on time of the laboratory staff for establishing the new systems (Figure 7). Researchers who plan to get started with microfluidic SNP genotyping should expect development costs of ~ 1,400€ (plus tax) for oligos per 96 SNP markers (sufficient for 75 genotyping runs, including STA pre-amplification) and have access to a Fluidigm genotyping system (EP1 or Biomark; Fluidigm Corp.). Costs per 96.96 microfluidic array, including additional chemistry, average 600€ (plus tax). If the required reference sample sets are ready to use, an SNP panel can be established within a few weeks upon arrival of the assays. Following our own guidelines, we were able to set up a new marker panel for standardized hybridization assessments of grey wolves and domestic dogs on the basis of six genotyping runs (our unpublished data).

Currently, the microfluidic array genotyping technology of Fluidigm provides platforms for 12, 24, 48 and 96 SNP markers for 12, 192, 48 and 96 samples per run, respectively. Furthermore, if the required analysis depends on a higher number of SNP markers, additional marker panels can be easily run sequentially.

4.2 | Genotyping performance in gNIS

Genetic noninvasive sampling is increasingly recognized as an efficient and cost-effective tool for monitoring rare and elusive species (Carroll et al., 2018; Ferreira et al., 2018; Waits & Paetkau, 2005). Making genomic methods accessible for genotyping of samples with low DNA quantity or quality, such as noninvasively collected wildlife samples, is one of the main prerequisites for bridging the conservation genomics gap (Carroll et al., 2018; McCormack et al., 2013; McMahon et al., 2014; Shafer et al., 2015). By implementing some specific protocol adjustments, microfluidic technology enables robust SNP genotyping of degraded DNA samples.

In our experience, adjusting the manufacturer's protocol for STA pre-amplification results in the greatest benefits. In particular, raising the number of STA cycles and increasing the volume of DNA extract reduced error rates and missing data (Figure S3 and von Thaden et al., 2017). Furthermore, a dilution of STA PCR products of 1:10 instead of the originally recommended 1:100 could be confirmed as most appropriate for our proposed genotyping protocol (Figure S3 and Table S2). Basically, all of these adjustments aim to increase the amount of template DNA available for the subsequent genotyping PCR. Providing a minimum amount of template is critical for gaining reliable genotype data, as is evident from our dilution series (Figure 2), where errors and missing data increase significantly for DNA concentrations of <1 ng/µl. This is in line with the findings of Bayerl et al. (2018) who showed that higher DNA quantities correlate with increased PCR efficiency as

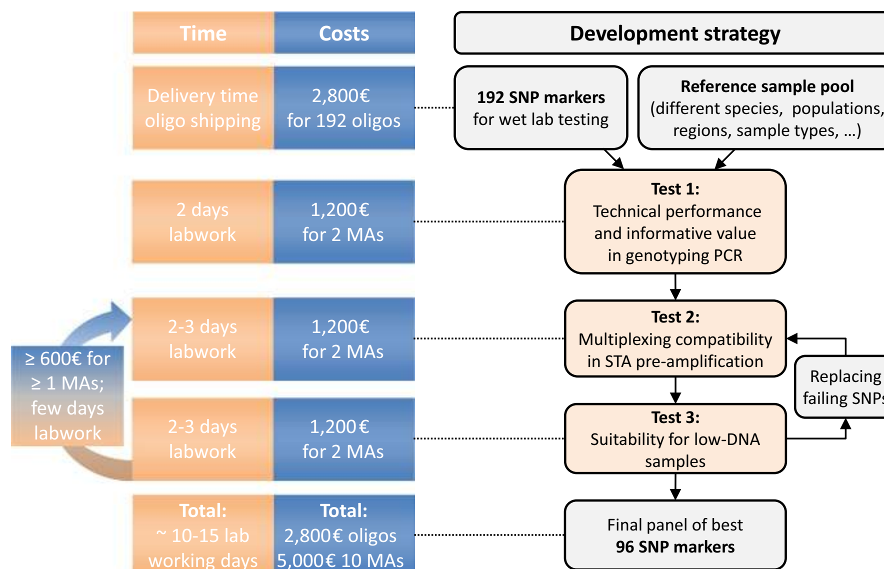


FIGURE 7 Approximate investment of hands-on time and start-up costs for establishing a new SNP marker panel following our proposed strategy. MA, microfluidic array. Labwork time estimates for one trained staff member, excluding analyses and experiment design. Cost estimates based on Fluidigm Corp. list prices in Germany (June 2018), without taxes. Costs for 192 oligos include STA pre-amplification and are sufficient for 75 runs; costs for MAs include additional chemistry. Both time and cost estimates are subject to variation depending on training and/or availability of laboratory staff as well as regional list prices for oligos and MA consumables

well as decreasing error rates for SNP genotyping of otter faeces. The minimal amount of genomic DNA required for successful SNP genotyping is estimated to be ~200 pg for hair (Nussberger, Wandeler, & Camenisch, 2014a) or 25 pg/µl for faecal samples (Bayerl et al., 2018; Hausknecht, Bayerl, Gula, & Kuehn, 2010). However, depending on the type of noninvasive sample (faeces, hairs, feathers, etc.), and the environmental conditions it is exposed to in the field, the level of DNA degradation or presence of inhibitors in the extracts may differ and thus the minimal amounts of genomic DNA may vary as well. For example, genotypes of scats showed slightly lower error rates than hair samples from the same reference individuals in this study (Figure S2). Ultimately, when setting up novel rSPs for other species and sample types, we recommend making a small pilot experiment to define a threshold for a favourable minimum DNA concentration, balancing low genotyping error rates against exclusion of valuable collected samples for the respective research question of the study.

In general, the performance of noninvasively collected or degraded samples will depend upon the methods used for DNA extraction and the degree of standardization for these procedures. Moreover, most molecular ecology laboratories will preselect noninvasive samples for their DNA quality or quantity before submitting them for multilocus genotyping. First, DNA concentrations of samples may be assessed beforehand, for example by RT-PCR quantification, where corresponding Ct values may be used as a proxy for genotyping quality (Morin, Chambers, Boesch, & Vigilant, 2001; for microfluidic technology see Nussberger, Wandeler, & Camenisch, 2014a; von Thaden et al., 2017). Second, samples may have undergone species identification based on mtDNA sequencing or have even been typed with microsatellite panels before SNP genotyping.

Several studies have applied SNP marker panels to noninvasively collected wildlife samples (Bayerl et al., 2018; Doyle et al., 2016; Giangregorio, Norman, Davoli, & Spong, 2018; Katzner et al., 2017; Kraus et al., 2015; Nussberger, Wandeler, & Camenisch, 2014a; Ruegg et al., 2014; Spitzer et al., 2016; Steyer, Tiesmeyer, Muñoz-Fuentes, & Nowak, 2018). The authors used different protocol adjustments and sample preselections, as the implementation of rSPs to noninvasive samples is still at an early stage. Some of the authors increased STA cycle numbers from 14 to 18 (von Thaden et al., 2017), 25 (Bayerl et al., 2018), 28 (this study) or even up to 35 (Giangregorio et al., 2018), while others raised the number of cycles in the subsequent genotyping PCR from the original 34 to 42 (Kraus et al., 2015) or 46 (Nussberger, Wandeler, & Camenisch, 2014a). The authors of the last study, who were among the first to apply microfluidic SNP genotyping to noninvasive samples, made extensive adjustments to the manufacturer's protocol to enable genotyping from single hairs of wildcats. Depending on their previous RT-PCR assessments, they raised the STA reaction volume up to 21 µl, increasing the DNA template volume up to 10 µl, and lowered STA primer concentrations to 250 nm to increase specificity. While we also raised the STA reaction volume slightly (from 5 to 8 µl), to accommodate 3.2 µl of DNA extract, we found that adjusting the primer concentrations from 500 to 250 nm did not improve performance (data not shown). As with other genetic markers, we expect that every laboratory will come up with their own adjustments to serve their individual application's needs. However, based on our comprehensive error calculations using reference samples, we are confident that increasing the number of cycles in STA does not lead to increased generation of artefacts or erroneous genotype signals, and is thus generally advisable for these types of samples.

Genotyping error rates of ADO (11%) and FA (1%) found in this study for the randomly selected subsets of noninvasively collected wildcat samples seem relatively high when compared to those reported for other microfluidic rSPs from Nussberger, Wandeler, and Camenisch (2014a); 1.6% genotyping error rate per locus; based on four reference individuals), Kraus et al. (2015; ~1% errors for dilution series of tissue and blood samples), Spitzer et al. (2016; 0.38%; based on preselected samples that had worked with microsatellite genotyping) or Doyle et al. (2016; 0.4%; after sample preselection). However, all of these studies estimated the genotyping errors after removing low-quality samples during the laboratory procedures. In our case, we estimated the genotyping errors for a randomly selected pool of samples originating from routine wildcat monitoring, where qualities and quantities of samples vary greatly. When comparing our reported SNP error rates for noninvasive samples to those obtained with traditional microsatellites on the same kinds of samples, error rates are significantly lower for the rSPs (ADO 11% vs. 19%–23% and FA 1% vs. 2%; Hartmann, Steyer, Kraus, Segelbacher, & Nowak, 2013; Steyer et al., 2016).

To account for genotyping errors, noninvasively collected samples with very low amounts of DNA are usually replicated in a multiple-tube approach (Pompanon, Bonin, Bellemain, & Taberlet, 2005; Taberlet et al., 1999; Waits & Paetkau, 2005). A previous study recommended classifying the replication needs of noninvasive samples based on their SNP call rates when using 18 cycles in STA pre-amplification (von Thaden et al., 2017), as samples with a call rate of 100% showed practically no disagreements between replicates. However, when increasing the cycles to 28 as tested here, the SNP call rates were generally much higher, including samples with very low DNA quantity or quality, with call rates of >95%. Thus, the relationship between SNP call rate and sample quality may need to be carefully re-evaluated when adjusting the STA protocol. Until then, we recommend a minimum of three replicates for low-DNA samples to detect potential genotyping errors. When low performance is obvious, more replicates may be needed to secure the genotype.

Furthermore, the consensus-building rules for replicated samples need to be empirically re-evaluated due to the differing characteristics of the bi-allelic SNPs, such as intrinsically higher rates of homozygotes as compared to multi-allelic markers. In the present study, the applied rules vary slightly from the recommendations for traditional microsatellite-based genotyping (Morin et al., 2001; Taberlet et al., 1996), which imply (a) extensive replication of homozygous genotypes as well as (b) solely assuming a heterozygous consensus genotype if observed at least twice among the replicates of a sample. To identify the most appropriate consensus-building rules for a given marker set, a pilot experiment based on extensive replication of reference samples comprising high and low DNA sample types is advisable, as we show here.

When genotyping wildlife samples, nontarget species may amplify and produce erroneous genotype calls. In our novel SNP panel for European wildcats, this was the case for all 18 tested cross-species. However, the vast majority (17 species) showed call rates below 20% and would immediately be invalidated due to these low amplification successes. Only samples of the Eurasian lynx produced call rates of over 90%, but showed merely three dissimilarities between samples of different individuals and very low observed heterozygosity (0.03). Thus, if lynx samples were to be genotyped without prior mtDNA-based species identification, they would most probably be identified as belonging to lynx based on their characteristic genotype or low heterozygosity. In contrast to nontarget species, transferability of the SNP assays to closely related species—such as the African wildcat samples briefly tested here—may prove beneficial. However, the applicability and informative value of an rSP for genotyping a closely related species needs to be thoroughly tested and assessed before its implementation in research or monitoring programmes, respectively (Giangregorio et al., 2018).

4.3 | Applications for standardized genomic assessments of natural populations

Several studies have shown that rSPs provide sufficient genomic information to answer questions on individual identification, introgression or relatedness (Baetscher et al., 2017; Holl et al., 2017; Katzner et al., 2017; Muñoz et al., 2015). Here, we have demonstrated how rSPs can be applied in conservation to gain information about several population demographic parameters in one genotyping run.

To distinguish individuals, as few as 18 of the most heterozygous SNPs were sufficient to reach PIDsib < 0.0001. Thus, when considering the 84 ID SNP markers in our panel, even closely related individuals should be distinguishable with high certainty. However, when we quantified mismatches between genotypes of known related domestic cat individuals, we found that siblings differed for only 11 of the 84 ID SNPs (13%), while unrelated domestic cats differed for as few as 30 ID SNPs (36%). Although our SNP panel was optimized to distinguish wildcats rather than domestic cat individuals, a minimum number of 40 SNP markers seems advisable for individual identification. With regard to possible genotyping errors or missing data for some loci, these numbers of SNP markers should yield enough polymorphisms to reliably distinguish individuals. Additionally, if available, mtDNA haplotypes, microsatellite genotypes, life history data and other metadata should be taken into account. Nevertheless, even without provision of metadata, we were able to reconstruct known kinships based on ID SNP genotypes. This may allow for further genealogical analyses that rely on these kinds of data.

Population structure was evident from DAPC and STRUCTURE and confirmed the findings from Mattucci et al. (2016), who used a panel of 31 microsatellites to discern biogeographical groups from the distribution range of the European wildcat. Pairwise F_{ST} values (Table S11) corroborated the genetic differentiation (mean = 0.147) and resembled values from Mattucci et al. (2016; mean = 0.108). Thus, although ID SNP markers were selected mainly for individual identification, an assignment of population origins is possible at least to some extent. Given the initial SNP identification process, it is not surprising that overall low values for HWE deviations and LD were found within wildcat populations. The few observed deviations from Hardy–Weinberg or linkage equilibrium for some of the loci are probably due to a sampling effect (Wahlund, 1928), since our nonreference population samples contained samples collected across several years as well as partly from genetically distinct natural populations (see also Supporting Results and Discussion). This assumption is supported by the DAPC and STRUCTURE analyses, which show that some sampled populations are composed of different genetic lineages (e.g., the Central and Western population in Germany, see Steyer et al., 2016).

The inclusion of 10 highly informative HYB SNP markers allowed us to quickly identify possible hybrids in our data set. The incidence of admixed individuals was higher in the samples from Scotland, Portugal and Spain compared to the other population samples, because these samples were taken from another study focusing on hybridization of wildcats in Europe (A. Tiesmeyer et al., unpublished data) and do not represent a random sample of individuals from these populations, although levels of hybridization in Scotland are known to be high (Beaumont et al., 2001). While many more SNP markers are necessary to distinguish between different hybrid categories with high certainties (compare Figure S8 and e.g., Mattucci et al., 2019; Nussberger, Wandeler, Weber, et al., 2014b; Oliveira et al., 2015; Steyer et al., 2018), the rapid detection of conspicuous genotypes will be of considerable usefulness in high-throughput monitoring. Events of hybridization will easily be detected in the course of monitoring routines, and, after closer evaluation of potential introgression with the above-mentioned SNP panels, appropriate conservation measures may be implemented such as the removal of hybrid individuals from the wild or their nonconsideration for breeding programmes (Kilshaw et al., 2016; Senn et al., 2019).

4.4 | Main conclusions and perspectives

The introduction of novel genomic tools to wildlife conservation brings both benefits and drawbacks. The biggest advantages of rSPs are the straightforward marker development described here, cost-effective high-throughput application using microfluidic arrays, and the opportunity to harmonize marker panels across different laboratories for joint conservation efforts and research (deGroot et al., 2016; Puckett, 2017). Short marker sizes (on average < 100 bp) and the very low volumes of required DNA extracts (<10 µl for triplicates) make the approach suitable for noninvasively collected samples that typically feature low quantities of mostly degraded DNA.

Nevertheless, implementing rSPs involves certain limitations. Compared to whole genome sequencing, the filtering steps to select informative SNPs for the respective research question may introduce ascertainment bias (Albrechtsen, Nielsen, & Nielsen, 2010; Malomane et al., 2018). As rare SNPs are often under-represented in reduced panels, values for expected heterozygosities as well as

fixation indices may be over- or underestimated (Malomane et al., 2018). Additionally, some of the SNPs may not be informative when marker panels are optimized for single populations only (compare Giangregorio et al., 2018). To minimize these biases, it is advisable to include as many samples as possible from the species' distribution range for SNP discovery and take appropriate measures when pruning the data for informative SNPs (e.g., by following an LD-based data pruning concept, as proposed by Malomane et al., 2018).

Another constraint may be the lower allelic richness of SNP markers as compared to multi-allelic microsatellites (Schopen, Bovenhuis, Visker, & van Arendonk, 2008). However, this is usually accounted for by including a selected set of highly informative SNPs suited for the respective research questions or by simply raising the number of investigated SNPs in order to reach the required statistical power (Morin, Martien, & Taylor, 2009). Generally, as with any marker type, it is of great importance to test the suitability and appropriateness of the marker panel and assess accuracy and efficiency for the respective research questions or monitoring purposes (Guichoux et al., 2013; Hoban, Gaggiotti, & Bertorelle, 2013; Landguth et al., 2012; Morin et al., 2009; examples for other species in, e.g., Galaverni et al., 2017; Tokarska et al., 2009). A more practical limitation for conservation management may be that single and/or urgent samples cannot be run alone economically, because an entire array of 96 samples needs to be run at once. Second, mixed samples containing DNA from several individuals (e.g., a family of cats marking a valerian-treated hair trap) or xenobiotic species (e.g., saliva from red fox scavenging on a wolf kill) may be hard to detect unless their genotypes show unusually high rates of heterozygosity and may further lead to increased error rates. While Steyer et al. (2016) found evidence for only 1.3% mixed DNA traces in noninvasively collected wildcat hair samples, other cases may lead to higher rates (e.g., Alpers, Taylor, Sunnucks, Bellman, & Sherwin, 2003; Roon, Thomas, Kendall, & Waits, 2005; Ruibal, Peakall, Claridge, Murray, & Firestone, 2010). Thus, the combined use of rSPs with, for example, mitochondrial species identification or comparison with existing microsatellite profiles is highly recommended for certain applications (Pun, Albrecht, Castella, & Fumagalli, 2009; von Thaden et al., 2017). Furthermore, the generation of thresholds for individual heterozygosity values may be used to sort out admixed samples, after careful evaluation of this approach for the specific research study at hand. Finally, provided that the respective SNP genotyping platform allows for it, the inclusion of tri-allelic SNP markers may facilitate identification of mixed DNA samples (Westen et al., 2009).

The rapid evolution of next-generation sequencing techniques brings forth a wealth of novel genomic approaches that may be embraced for conservation genomics by molecular ecologists and management practitioners alike. Here, we have provided practical guidelines and a framework for how to establish rSPs for genotyping of wildlife samples using microfluidic arrays. The technology has potential to become the method of choice for genomic monitoring of endangered species, due to ease of establishment, high standardization potentials, and reduced costs and hands-on time. However, as new sequencing methods evolve, microfluidic arrays may be complemented by other sophisticated monitoring methods (Meek & Larson, 2019; Pavey, 2015). In this respect, the guidelines for genomic marker development and testing presented in this study may be useful for a variety of marker systems prior to beginning a long-term monitoring or research project.

ACKNOWLEDGEMENTS

Funding for this study was generated through the wildlife genetics analysis service at the Senckenberg Conservation Genetics Section, Gelnhausen, Germany. Additional funds came from the Leibniz Association (SAW-2011-SGN-3) and the Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz of the German Federal State of Hesse. A.V.T. received generous funding from the Karl und Marie Schack-Stiftung; L.A.L. received funding from the National Institutes of Health – National Center for Research Resources (NCRR) grant R24 RR016094R24, now the Office of Research Infrastructure Programs (ORIP) grant R24OD010928, Cat Health Network (D12FE-505) and the University of California, Davis, Center for Companion Animal Health, School of Veterinary Medicine (2011–30-F/M); and M.C. received funding from the Friedrich-Schiller-University of Jena. We are grateful to Violeta Muñoz-Fuentes for initial assistance in SNP selection. We thank the Bund für Umwelt und Naturschutz Germany (BUND), the Tierschutz- und Wildgehegeverein im Tierzentrum e.V. (Gelnhausen, Germany), Z. Boratyński, L. Dorigo, A. Dutsov, M. S. Fischer, C. Frosch, L. Khalatbari, L. Lapini, F. Müller, R. Oliveira, A. Sforzi, K. Steyer, C. Übl, D. Youlatos and their collaborators as well as several private pet owners for sample provision. A.C.K. thanks the Negaunee Foundation for their generous support of a preparator who sampled cats from Scotland.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

A.V.T., B.C. and C.N. designed the study. P.C.A., L.A.L., F.M. and E.R. identified the pool of candidate SNPs for the panel development. M.C., J.G., Z.H., A.C.K., C.L., J.M.L., T.M., L.R. and V.S. provided European wildcat samples. A.V.T., B.C., A.T., L.R. and J.M.L. carried out sample extractions and microsatellite typing. A.V.T. and B.C. carried out SNP genotyping and all further laboratory procedures. A.V.T., B.C. and T.E.R. analysed the data. A.V.T., B.C. and C.N. wrote the manuscript. All authors edited and approved the final manuscript.

DATA AVAILABILITY STATEMENT

Sequences for the 96 SNPtype™ assay primers, SNP genotyping data and sample information for this study are available in the Supporting Excel file.

ORCID

Alina von Thaden https://orcid.org/0000-0001-9060-0912
Berardino Cocchiararo https://orcid.org/0000-0002-8427-7861

REFERENCES

Albrechtsen, A., Nielsen, F. C., & Nielsen, R. (2010). Ascertainment biases in SNP chips affect measures of population divergence. *Molecular Biology and Evolution*, *27*, 2534–2547. https://doi.org/10.1093/molbev/msq148

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews. Genetics*, *11*, 697–709. https://doi.org/10.1038/nrg2844

Alpers, D. L., Taylor, A. C., Sunnucks, P., Bellman, S. A., & Sherwin, W. B. (2003). Pooling hair samples to increase DNA yield for PCR. *Conservation Genetics*, *4*, 779–788. https://doi.org/10.1023/B:COGE.0000006119.77842.4b

Anderson, E. C., & Thompson, E. A. (2002). A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, *160*, 1217–1229.

Andrews, K. R., Adams, J. R., Cassirer, E. F., Plowright, R. K., Gardner, C., Dwire, M., … Waits, L. P. (2018). A bioinformatic pipeline for identifying informative SNP panels for parentage assignment from RADseq data. *Molecular Ecology Resources*, *18*(6), 1263–1281. https://doi.org/10.1111/1755-0998.12910

Baetscher, D. S., Hasselman, D. J., Reid, K., Palkovacs, E. P., & Garza, J. C. (2017). Discovery and characterization of single nucleotide polymorphisms in two anadromous alosine fishes of conservation concern. *Ecology and Evolution*, *7*(17), 6638–6648. https://doi.org/10.1002/ece3.3215

Bayerl, H., Kraus, R. H. S., Nowak, C., Foerster, D. W., Fickel, J., & Kuehn, R. (2018). Fast and cost-effective single nucleotide polymorphism (SNP) detection in the absence of a reference genome using semi-deep next generation random amplicon sequencing (RAMseq). *Molecular Ecology Resources*, *18*, 107–117. https://doi.org/10.1111/1755-0998.12717

Beaumont, M. A., Barratt, E. M., Gottelli, D., Kitchener, A. C., Daniels, M. J., Pritchard, J. K., & Bruford, M. W. (2001). Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology*, *10*, 319–336. https://doi.org/10.1046/j.1365-294X.2001.01196.x

Benestan, L. M., Ferchaud, A.-L., Hohenlohe, P. A., Garner, B. A., Naylor, G. J. P., Baums, I. B., … Luikart, G. (2016). Conservation genomics of natural and managed populations: Building a conceptual and practical framework. *Molecular Ecology*, *25*, 2967–2977. https://doi.org/10.1111/mec.13647

Britt, M., Haworth, S. E., Johnson, J. B., Martchenko, D., & Shafer, A. B. A. (2018). The importance of non-academic coauthors in bridging the conservation genetics gap. *Biological Conservation*, *218*, 118–123. https://doi.org/10.1016/j.biocon.2017.12.019

Broquet, T., & Petit, E. (2004). Quantifying genotyping errors in noninvasive population genetics. *Molecular Ecology*, *13*, 3601–3608. https://doi.org/10.1111/j.1365-294X.2004.02352.x

Carroll, E. L., Bruford, M. W., DeWoody, J. A., Leroy, G., Strand, A., Waits, L., & Wang, J. (2018). Genetic and genomic monitoring with minimally invasive sampling methods. *Evolutionary Applications*, *11*, 1094–1119. https://doi.org/10.1111/eva.12600

deGroot, G. A., Nowak, C., Skrbinšek, T., Andersen, L. W., Aspi, J., Fumagalli, L., … Muñoz-Fuentes, V. (2016). Decades of population genetic research reveal the need for harmonization of molecular markers: The grey wolf *Canis lupus* as a case study. *Mammal Review*, *46*, 44–59. https://doi.org/10.1111/mam.12052

Desalle, R., & Amato, G. (2004). The expansion of conservation genetics. *Nature Reviews Genetics*, *5*, 702–712. https://doi.org/10.1038/nrg1425

DeWoody, J. A., Fernandez, N. B., Brüniche-Olsen, A., Antonides, J. D., Doyle, J. M., San Miguel, P., … Bickham, J. W. (2017). Characterization of the gray whale *Eschrichtius robustus* genome and a genotyping array based on single-nucleotide polymorphisms in candidate genes. *The Biological Bulletin*, *232*, 186–197. https://doi.org/10.1086/693483

Doyle, J. M., Katzner, T. E., Roemer, G. W., Cain, J. W., Millsap, B. A., McIntyre, C. L., … Andrew DeWoody, J. (2016). Genetic structure and viability selection in the golden eagle (*Aquila chrysaetos*), a vagile raptor with a Holarctic distribution. *Conservation Genetics*, *17*, 1307–1322. https://doi.org/10.1007/s10592-016-0863-0

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

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

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Ferreira, C. M., Sabino-Marques, H., Barbosa, S., Costa, P., Encarnação, C., Alpizar-Jara, R., … Alves, P. C. (2018). Genetic non-invasive sampling (gNIS) as a cost-effective tool for monitoring elusive small mammals. *European Journal of Wildlife Research*, *64*, 46. https://doi.org/10.1007/s10344-018-1188-8

Fitak, R. R., Naidu, A., Thompson, R. W., & Culver, M. (2016). A new panel of SNP markers for the individual identification of North American pumas. *Journal of Fish and Wildlife Management*, *7*, 13–27. https://doi.org/10.3996/112014-JFWM-080

Flanagan, S. P., Forester, B. R., Latch, E. K., Aitken, S. N., & Hoban, S. (2018). Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications*, *11*(7), 1035–1052. https://doi.org/10.1111/eva.12569

Foerster, D., Bull, J. K., Lenz, D., Autenrieth, M., Paijmans, J. L. A., Kraus, R. H. S., Fickel, J. (2017). Targeted re-sequencing of coding DNA sequences for SNP discovery in non-model species. bioRxiv.org: the preprint server for biology, 163659. https://doi.org/10.1101/163659

Frankham, R. (2010). Where are we in conservation genetics and where do we need to go? *Conservation Genetics*, *11*, 661–663. https://doi.org/10.1007/s10592-009-0010-2

Galaverni, M., Caniglia, R., Pagani, L., Fabbri, E., Boattini, A., & Randi, E. (2017). Disentangling timing of admixture, patterns of introgression, and phenotypic indicators in a hybridizing wolf population. *Molecular Biology and Evolution*, *34*, 2324–2339. https://doi.org/10.1093/molbev/msx169

Gandolfi, B., Alhaddad, H., Abdi, M., Bach, L. H., Creighton, E. K., Davis, B. W., … Lyons, L. A. (2018). Applications and efficiencies of the first cat 63K DNA array. *Scientific Reports*, *8*, 7024. https://doi.org/10.1038/s41598-018-25438-0

Garner, B. A., Hand, B. K., Amish, S. J., Bernatchez, L., Foster, J. T., Miller, K. M., … Luikart, G. (2016). Genomics in conservation: Case studies and bridging the gap between data and application. *Trends in Ecology & Evolution*, *31*, 81–83. https://doi.org/10.1016/j.tree.2015.10.009

Garvin, M. R., Saitoh, K., & Gharrett, A. J. (2010). Application of single nucleotide polymorphisms to non-model species: A technical review. *Molecular Ecology Resources*, *10*, 915–934. https://doi.org/10.1111/j.1755-0998.2010.02891.x

Giangregorio, P., Norman, A. J., Davoli, F., & Spong, G. (2018). Testing a new SNP-chip on the Alpine and Apennine brown bear (*Ursus arctos*) populations using non-invasive samples. *Conservation Genetics Resources*, *11*(3), 355–363. https://doi.org/10.1007/s12686-018-1017-0

Goossens, B., Waits, L. P., & Taberlet, P. (1998). Plucked hair samples as a source of DNA: Reliability of dinucleotide microsatellite genotyping. *Molecular Ecology*, *7*, 1237–1241. https://doi.org/10.1046/j.1365-294x.1998.00407.x

Guichoux, E., Garnier-Géré, P., Lagache, L., Lang, T., Boury, C., & Petit, R. J. (2013). Outlier loci highlight the direction of introgression in oaks. *Molecular Ecology*, *22*, 450–462. https://doi.org/10.1111/mec.12125

Hartmann, S. A., Steyer, K., Kraus, R. H. S., Segelbacher, G., & Nowak, C. (2013). Potential barriers to gene flow in the endangered European wildcat (*Felis silvestris*). *Conservation Genetics*, *14*, 413–426. https://doi.org/10.1007/s10592-013-0468-9

Hausknecht, R., Bayerl, H., Gula, R., & Kuehn, R. (2010). Application of quantitative real-time polymerase chain reaction for noninvasive genetic monitoring. *The Journal of Wildlife Management*, *74*, 1904–1910. https://doi.org/10.2193/2009-421

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

Henriques, D., Parejo, M., Vignal, A., Wragg, D., Wallberg, A., Webster, M. T., & Pinto, M. A. (2018). Developing reduced SNP assays from whole-genome sequence data to estimate introgression in an organism with complex genetic patterns, the Iberian honeybee (*Apis mellifera iberiensis*). *Evolutionary Applications*, *11*(8), 1270–1282. https://doi.org/10.1111/eva.12623

Hoban, S. M., Gaggiotti, O. E., & Bertorelle, G. (2013). The number of markers and samples needed for detecting bottlenecks under realistic scenarios, with and without recovery: A simulation-based study. *Molecular Ecology*, *22*, 3444–3450. https://doi.org/10.1111/mec.12258

Hoffman, J. I., Tucker, R., Bridgett, S. J., Clark, M. S., Forcada, J., & Slate, J. (2012). Rates of assay success and genotyping error when single nucleotide polymorphism genotyping in non-model organisms: A case study in the Antarctic fur seal. *Molecular Ecology Resources*, *12*, 861–872. https://doi.org/10.1111/j.1755-0998.2012.03158.x

Hogg, C. J., Taylor, H. R., Fox, S., Grueber, C. E. (2018). Response to Britt et al. 2018 "The importance of non-academic co-authors in bridging the conservation genetics gap" Biological Conservation 218, 118–123. *Biological Conservation*, *222*, 287–288. https://doi.org/10.1016/j.biocon.2018.04.035

Holderegger, R., Balkenhol, N., Bolliger, J., Engler, J. O., Gugerli, F., Hochkirch, A., ... Zachos, F. E. (2019). Conservation genetics: Linking science with practice. *Molecular Ecology*, *28*, 3848–3856. https://doi.org/10.1111/mec.15202

Holl, H. M., Vanhnasy, J., Everts, R. E., Hoefs-Martin, K., Cook, D., Brooks, S. A., ... Lafayette, C. (2017). Single nucleotide polymorphisms for DNA typing in the domestic horse. *Animal Genetics*, *48*, 669–676. https://doi.org/10.1111/age.12608

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics (Oxford, England)*, *23*, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jombart, T. (2008). *adegenet*: A R package for the multivariate analysis of genetic markers. *Bioinformatics (Oxford, England)*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kalinowski, S. T., Wagner, A. P., & Taper, M. L. (2006). ML-RELATE: A computer program for maximum likelihood estimation of relatedness and relationship. *Molecular Ecology Notes*, *6*, 576–579. https://doi.org/10.1111/j.1471-8286.2006.01256.x

Katzner, T. E., Nelson, D. M., Braham, M. A., Doyle, J. M., Fernandez, N. B., Duerr, A. E., ... DeWoody, J. A. (2017). Golden Eagle fatalities and the continental-scale consequences of local wind-energy generation. *Conservation Biology*, *31*, 406–415. https://doi.org/10.1111/cobi.12836

Kilshaw, K., Montgomery, R. A., Campbell, R. D., Hetherington, D. A., Johnson, P. J., Kitchener, A. C., ... Millspaugh, J. J. (2016). Mapping the spatial configuration of hybridization risk for an endangered population of the European wildcat (*Felis silvestris silvestris*) in Scotland. *Mammal Research*, *61*, 1–11. https://doi.org/10.1007/s13364-015-0253-x

Kraus, R. H. S., vonHoldt, B., Cocchiararo, B., Harms, V., Bayerl, H., Kühn, R., ... Nowak, C. (2015). A single-nucleotide polymorphism-based approach for rapid and cost-effective genetic wolf monitoring in Europe based on noninvasively collected samples. *Molecular Ecology Resources*, *15*, 295–305. https://doi.org/10.1111/1755-0998.12307

Landguth, E. L., Fedy, B. C., Oyler-McCance, S. J., Garey, A. L., Emel, S. L., Mumma, M., ... Cushman, S. A. (2012). Effects of sample size, number of markers, and allelic richness on the detection of spatial genetic pattern. *Molecular Ecology Resources*, *12*, 276–284. https://doi.org/10.1111/j.1755-0998.2011.03077.x

Lepoittevin, C., Bodénès, C., Chancerel, E., Villate, L., Lang, T., Lesur, I., ... Kremer, A. (2015). Single-nucleotide polymorphism discovery and validation in high-density SNP array for genetic analysis in European white oaks. *Molecular Ecology Resources*, *15*, 1446–1459. https://doi.org/10.1111/1755-0998.12407

Leroy, G., Carroll, E. L., Bruford, M. W., DeWoody, J. A., Strand, A., Waits, L., & Wang, J. (2018). Next-generation metrics for monitoring genetic erosion within populations of conservation concern. *Evolutionary Applications*, *11*(7), 1066–1083. https://doi.org/10.1111/eva.12564

Li, L., Wei, L., Ma, J., & Chen, J. (2018). Development and characterization of 21 SNP markers in *Manis javanica* based on high-throughput sequencing. *Conservation Genetics Resources*, *10*(3), 351–353. https://doi.org/10.1007/s12686-017-0822-1

Luo, S.-J., Johnson, W. E., David, V. A., Menotti-Raymond, M., Stanyon, R., Cai, Q. X., ... O'Brien, S. J. (2007). Development of Y chromosome intraspecific polymorphic markers in the Felidae. *The Journal of Heredity*, *98*, 400–413. https://doi.org/10.1093/jhered/esm063

Malomane, D. K., Reimer, C., Weigend, S., Weigend, A., Sharifi, A. R., & Simianer, H. (2018). Efficiency of different strategies to mitigate ascertainment bias when using SNP panels in diversity studies. *BMC Genomics*, *19*, 22. https://doi.org/10.1186/s12864-017-4416-9

Mattucci, F. (2014). *Conservation genetics of European wildcat (*Felis silvestris silvestris*): A wide and integrating analysis protocol for admixture inferences and population structure* (Unpublished doctoral dissertation). University of Bologna. Retrieved from http://amsdottorato.unibo.it/6459/1/Phd_thesis_MattucciF_170314.pdf

Mattucci, F., Galaverni, M., Lyons, L. A., Alves, P. C., Randi, E., Velli, E., ... Caniglia, R. (2019). Genomic approaches to identify hybrids and estimate admixture times in European wildcat populations. *Scientific Reports*, *9*, 1–15. https://doi.org/10.1038/s41598-019-48002-w

Mattucci, F., Oliveira, R., Bizzarri, L., Vercillo, F., Anile, S., Ragni, B., ... Randi, E. (2013). Genetic structure of wildcat (*Felis silvestris*) populations in Italy. *Ecology and Evolution*, *3*, 2443–2458. https://doi.org/10.1002/ece3.569

Mattucci, F., Oliveira, R., Lyons, L. A., Alves, P. C., & Randi, E. (2016). European wildcat populations are subdivided into five main biogeographic groups: Consequences of Pleistocene climate changes or recent anthropogenic fragmentation? *Ecology and Evolution*, *6*, 3–22. https://doi.org/10.1002/ece3.1815

McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., & Brumfield, R. T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, *66*, 526–538. https://doi.org/10.1016/j.ympev.2011.12.007

McMahon, B. J., Teeling, E. C., & Höglund, J. (2014). How and why should we implement genomics into conservation? *Evolutionary Applications*, *7*, 999–1007. https://doi.org/10.1111/eva.12193

Meek, M. H., & Larson, W. A. (2019). The future is now: Amplicon sequencing and sequence capture usher in the conservation genomics era. *Molecular Ecology Resources*, *19*, 795–803. https://doi.org/10.1111/1755-0998.12998

Morin, P. A., Chambers, K. E., Boesch, C., & Vigilant, L. (2001). Quantitative polymerase chain reaction analysis of DNA from noninvasive

samples for accurate microsatellite genotyping of wild chimpanzees (*Pan troglodytes verus*). *Molecular Ecology*, *10*, 1835–1844. https://doi.org/10.1046/j.0962-1083.2001.01308.x

Morin, P. A., Luikart, G., Wayne, R. K., & the SNP workshop group (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, *19*, 208–216. https://doi.org/10.1016/j.tree.2004.01.009

Morin, P. A., Martien, K. K., & Taylor, B. L. (2009). Assessing statistical power of SNPs for population structure and conservation studies. *Molecular Ecology Resources*, *9*, 66–73. https://doi.org/10.1111/j.1755-0998.2008.02392.x

Muñoz, I., Henriques, D., Johnston, J. S., Chavez-Galarza, J., Kryger, P., & Pinto, M. A. (2015). Reduced SNP panels for genetic identification and introgression analysis in the dark honey bee (*Apis mellifera mellifera*). *PLoS One*, *10*, e0124365. https://doi.org/10.1371/journal.pone.0124365

Norman, A. J., Street, N. R., & Spong, G. (2013). De novo SNP discovery in the Scandinavian brown bear (*Ursus arctos*). *PLoS One*, *8*, e81012. https://doi.org/10.1371/journal.pone.0081012

Norman, A. J., Stronen, A. V., Fuglstad, G.-A., Ruiz-Gonzalez, A., Kindberg, J., Street, N. R., & Spong, G. (2017). Landscape relatedness: Detecting contemporary fine-scale spatial structure in wild populations. *Landscape Ecology*, *32*, 181–194. https://doi.org/10.1007/s10980-016-0434-2

Nussberger, B., Currat, M., Quilodran, C. S., Ponta, N., & Keller, L. F. (2018). Range expansion as an explanation for introgression in European wildcats. *Biological Conservation*, *218*, 49–56. https://doi.org/10.1016/j.biocon.2017.12.009

Nussberger, B., Greminger, M. P., Grossen, C., Keller, L. F., & Wandeler, P. (2013). Development of SNP markers identifying European wildcats, domestic cats, and their admixed progeny. *Molecular Ecology Resources*, *13*, 447–460. https://doi.org/10.1111/1755-0998.12075

Nussberger, B., Wandeler, P., & Camenisch, G. (2014a). A SNP chip to detect introgression in wildcats allows accurate genotyping of single hairs. *European Journal of Wildlife Research*, *60*, 405–410. https://doi.org/10.1007/s10344-014-0806-3

Nussberger, B., Wandeler, P., Weber, D., & Keller, L. F. (2014b). Monitoring introgression in European wildcats in the Swiss Jura. *Conservation Genetics*, *15*, 1219–1230. https://doi.org/10.1007/s10592-014-0613-0

Oliveira, R., Randi, E., Mattucci, F., Kurushima, J. D., Lyons, L. A., & Alves, P. C. (2015). Toward a genome-wide approach for detecting hybrids: Informative SNPs to detect introgression between domestic cats and European wildcats (*Felis silvestris*). *Heredity*, *115*, 195–205. https://doi.org/10.1038/hdy.2015.25

Pavey, S. A. (2015). High-throughput SNPs for all: Genotyping-in-thousands. *Molecular Ecology Resources*, *15*, 685–687. https://doi.org/10.1111/1755-0998.12405

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel: Population genetic software for teaching and research – an update. *Bioinformatics (Oxford, England)*, *28*, 2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Pompanon, F., Bonin, A., Bellemain, E., & Taberlet, P. (2005). Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics*, *6*, 847–859. https://doi.org/10.1038/nrg1707

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Pritchard, V. L., Erkinaro, J., Kent, M. P., Niemelä, E., Orell, P., Lien, S., & Primmer, C. R. (2016). Single nucleotide polymorphisms to discriminate different classes of hybrid between wild Atlantic salmon and aquaculture escapees. *Evolutionary Applications*, *9*, 1017–1031. https://doi.org/10.1111/eva.12407

Puckett, E. E. (2017). Variability in total project and per sample genotyping costs under varying study designs including with microsatellites or SNPs to answer conservation genetic questions. *Conservation Genetics Resources*, *9*, 289–304. https://doi.org/10.1007/s12686-016-0643-7

Pun, K.-M., Albrecht, C., Castella, V., & Fumagalli, L. (2009). Species identification in mammals from mixed biological samples based on mitochondrial DNA control region length polymorphism. *Electrophoresis*, *30*, 1008–1014. https://doi.org/10.1002/elps.200800365

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, *81*, 559–575. https://doi.org/10.1086/519795

Quinto-Cortés, C. D., Woerner, A. E., Watkins, J. C., & Hammer, M. F. (2018). Modeling SNP array ascertainment with Approximate Bayesian Computation for demographic inference. *Scientific Reports*, *8*, 10209. https://doi.org/10.1038/s41598-018-28539-y

R Development Core Team (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Roon, D. A., Thomas, M. E., Kendall, K. C., & Waits, L. P. (2005). Evaluating mixed samples as a source of error in non-invasive genetic studies using microsatellites. *Molecular Ecology*, *14*, 195–201. https://doi.org/10.1111/j.1365-294X.2004.02385.x

Ruegg, K. C., Anderson, E. C., Paxton, K. L., Apkenas, V., Lao, S., Siegel, R. B., … Smith, T. B. (2014). Mapping migration in a songbird using high-resolution genetic markers. *Molecular Ecology*, *23*, 5726–5739. https://doi.org/10.1111/mec.12977

Ruibal, M., Peakall, R., Claridge, A., Murray, A., & Firestone, K. (2010). Advancement to hair-sampling surveys of a medium-sized mammal: DNA-based individual identification and population estimation of a rare Australian marsupial, the spotted-tailed quoll (*Dasyurus maculatus*). *Wildlife Research*, *37*, 27–38. https://doi.org/10.1071/WR09087

Russello, M. A., Waterhouse, M. D., Etter, P. D., & Johnson, E. A. (2015). From promise to practice: Pairing non-invasive sampling with genomics in conservation. *PeerJ*, *3*, e1106. https://doi.org/10.7717/peerj.1106

Say, L., Devillard, S., Léger, F., Pontier, D., & Ruette, S. (2012). Distribution and spatial genetic structure of European wildcat in France. *Animal Conservation*, *15*, 18–27. https://doi.org/10.1111/j.1469-1795.2011.00478.x

Schopen, G. C. B., Bovenhuis, H., Visker, M. H. P. W., & van Arendonk, J. A. M. (2008). Comparison of information content for microsatellites and SNPs in poultry and cattle. *Animal Genetics*, *39*, 451–453. https://doi.org/10.1111/j.1365-2052.2008.01736.x

Senn, H., Ghazali, M., Kaden, J., Barclay, D., Harrower, B., Campbell, R. D., … Kitchener, A. C. (2019). Distinguishing the victim from the threat: SNP-based methods reveal the extent of introgressive hybridisation between wildcats and domestic cats in Scotland and inform future in-situ and ex-situ management options for species restoration. *Evolutionary Applications*, *12*, 399–414. https://doi.org/10.1111/eva.12720

Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., … Zieliński, P. (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution*, *30*, 78–87. https://doi.org/10.1016/j.tree.2014.11.009

Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Colling, G., Dalén, L., … Zieliński, P. (2016). *Trends in Ecology & Evolution*, *31*, 83–84. https://doi.org/10.1016/j.tree.2015.11.010

Spitzer, R., Norman, A. J., Schneider, M., Spong, G. (2016). Estimating population size using single-nucleotide polymorphism-based pedigree data. *Ecology and Evolution*, *6*, 3174–3184. https://doi.org/10.1002/ece3.2076

Starks, H. A., Clemento, A. J., & Garza, J. C. (2015). Discovery and characterization of single nucleotide polymorphisms in coho salmon, *Oncorhynchus kisutch*. *Molecular Ecology Resources*, *16*, 277–287. https://doi.org/10.1111/1755-0998.12430

Stetz, J. B., Smith, S., Sawaya, M. A., Ramsey, A. B., Amish, S. J., Schwartz, M. K., & Luikart, G. (2016). Discovery of 20,000 RAD–SNPs and development of a 52-SNP array for monitoring river otters. *Conservation Genetics Resources*, *8*, 299–302. https://doi.org/10.1007/s12686-016-0558-3

Steyer, K., Kraus, R. H. S., Mölich, T., Anders, O., Cocchiararo, B., Frosch, C., ... Nowak, C. (2016). Large-scale genetic census of an elusive carnivore, the European wildcat (*Felis s. silvestris*). *Conservation Genetics*, *17*, 1183–1199. https://doi.org/10.1007/s10592-016-0853-2

Steyer, K., Simon, O., Kraus, R. H. S., Haase, P., & Nowak, C. (2013). Hair trapping with valerian-treated lure sticks as a tool for genetic wildcat monitoring in low-density habitats. *European Journal of Wildlife Research*, *59*, 39–46. https://doi.org/10.1007/s10344-012-0644-0

Steyer, K., Tiesmeyer, A., Muñoz-Fuentes, V., & Nowak, C. (2018). Low rates of hybridization between European wildcats and domestic cats in a human-dominated landscape. *Ecology and Evolution*, *16*, 613. https://doi.org/10.1002/ece3.3650

Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., ... Bouvet, J. (1996). Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, *24*, 3189–3194. https://doi.org/10.1093/nar/24.16.3189

Taberlet, P., Waits, L. P., & Luikart, G. (1999). Noninvasive genetic sampling: Look before you leap. *Trends in Ecology & Evolution*, *14*, 323–327. https://doi.org/10.1016/S0169-5347(99)01637-7

Taylor, H. R., Dussex, N., & van Heezik, Y. (2017). Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Global Ecology and Conservation*, *10*, 231–242. https://doi.org/10.1016/j.gecco.2017.04.001

Taylor, H. R., & Gemmell, N. J. (2016). Emerging Technologies to Conserve Biodiversity: Further Opportunities via Genomics. Response to Pimm et al. *Trends in Ecology & Evolution*, *31*(3), 171–172. https://doi.org/10.1016/j.tree.2016.01.002

Tokarska, M., Marshall, T., Kowalczyk, R., Wójcik, J. M., Pertoldi, C., Kristensen, T. N., ... Bendixen, C. (2009). Effectiveness of microsatellite and SNP markers for parentage and identity analysis in species with low genetic diversity: The case of European bison. *Heredity*, *103*, 326–332. https://doi.org/10.1038/hdy.2009.73

von Thaden, A., Cocchiararo, B., Jarausch, A., Jüngling, H., Karamanlidis, A. A., Tiesmeyer, A., ... Muñoz-Fuentes, V. (2017). Assessing SNP genotyping of noninvasively collected wildlife samples using microfluidic arrays. *Scientific Reports*, *7*, 83. https://doi.org/10.1038/s41598-017-10647-w

Wahlund, S. (1928). Composition of populations and correlation appearances viewed in relation to the studies of inheritance. *Hereditas*, *11*, 65–106.

Waits, L. P., Luikart, G., & Taberlet, P. (2001). Estimating the probability of identity among genotypes in natural populations: Cautions and guidelines. *Molecular Ecology*, *10*, 249–256. https://doi.org/10.1046/j.1365-294X.2001.01185.x

Waits, L. P., & Paetkau, D. (2005). Noninvasive genetic sampling tools for wildlife biologists: A review of applications and recommendations for accurate data collection. *The Journal of Wildlife Management*, *69*, 1419–1433. https://doi.org/10.2193/0022-541X(2005)69[1419:NGSTFW]2.0.CO;2

Wang, J., Lin, M., Crenshaw, A., Hutchinson, A., Hicks, B., Yeager, M., ... Ramakrishnan, R. (2009). High-throughput single nucleotide polymorphism genotyping using nanofluidic Dynamic Arrays. *BMC Genomics*, *10*, 561. https://doi.org/10.1186/1471-2164-10-561

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-Statistics for the analysis of population structure. *Evolution*, *38*, 1358. https://doi.org/10.2307/2408641

Westen, A. A., Matai, A. S., Laros, J. F. J., Meiland, H. C., Jasper, M., de Leeuw, W. J. F., ... Sijen, T. (2009). Tri-allelic SNP markers enable analysis of mixed and degraded DNA samples. *Forensic Science International: Genetics*, *3*, 233–241. https://doi.org/10.1016/j.fsigen.2009.02.003

Wray, N. R., Yang, J., Hayes, B. J., Price, A. L., Goddard, M. E., & Visscher, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nature Reviews. Genetics*, *14*, 507–515. https://doi.org/10.1038/nrg3457

Zhao, H., Li, C., Hargrove, J. S., Bowen, B. R., Thongda, W., Zhang, D., ... Peatman, E. (2018). SNP marker panels for parentage assignment and traceability in the Florida bass (*Micropterus floridanus*). *Aquaculture*, *485*, 30–38. https://doi.org/10.1016/j.aquaculture.2017.11.014

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: vonThaden A, Nowak C, Tiesmeyer A, et al. Applying genomic data in wildlife monitoring: Development guidelines for genotyping degraded samples with reduced single nucleotide polymorphism panels. *Mol Ecol Resour*. 2020;20:662–680. https://doi.org/10.1111/1755-0998.13136