Title

Standardization of a SNP panel for parentage verification and identification in the domestic cat (Felis silvestris catus)

Journal

Animal Genetics, 52(5)

ISSN

0268-9146

Authors

Groot, M
Anderson, H
Bauer, H
et al.

Publication Date

2021-10-01

DOI

10.1111/age.13100

Peer reviewed

ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

FULL PAPER

doi: 10.1111/age.13100

Standardization of a SNP panel for parentage verification and identification in the domestic cat (*Felis silvestris catus*)

M. de Groot*, H. Anderson†, H. Bauer‡, C. Bauguil§, R. R. Bellone¶,, R. Brugidou§, R. M. Buckley††, P. Dovč‡‡, O. Forman†, R. A. Grahn¶, L. Kock§§, M. Longeri¶¶, S. Mouysset-Geniez§, J. Qiu§§, G. Sofronidis***, L. H. P. van der Goor††† and L. A. Lyons††**

*MolGen, Traverse 2, Veenendaal, Utrecht 3905NL, The Netherlands. †Wisdom Health, Helsinki 00290, Finland. ‡Laboklin GMBH & Co. KG, Bad Kissingen 97688, Germany. §Aveyron Labo, Rodez 12031, France. ¶Veterinary Genetics Laboratory, School of Veterinary Medicine, University of California, Davis, CA 95616, USA. **Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA 95616, USA. ††Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri, Columbia, MO 65211, USA. ‡‡Department of Animal Science, Biotechnical Faculty, University of Ljubljana, Ljubljana 1000, Slovenia. §§Neogen Genomics, Lincoln, NE 68504, USA. ¶¶Department of Veterinary Medicine, University of Milan, Milan 20133, Italy. ***Orivet Genetic Pet Care, Suite St. Kilda, Melbourne, Vic. 3182, Australia. †††VHLGenetics, Wageningen, NL 6708 PW, The Netherlands.

Summary

The domestic cat (*Felis silvestris catus*) is a valued companion animal throughout the world. Over 60 different cat breeds are accepted for competition by the cat fancy registries in different countries. Genetic markers, including short tandem repeats and SNPs, are available to evaluate and manage levels of inbreeding and genetic diversity, population and breed structure relationships, and individual identification for forensic and registration purposes. The International Society of Animal Genetics (ISAG) hosts the Applied Genetics in Companion Animals Workshop, which supports the standardization of genetic marker panels and genotyping for the identification of cats via comparison testing. SNP panels have been in development for many species, including the domestic cat. An ISAG approved core panel of SNPs for use in cat identification and parentage analyses is presented. SNPs ($n = 121$) were evaluated by different university-based and commercial laboratories using 20 DNA samples as part of the ISAG comparison testing procedures. Different SNP genotyping technologies were examined, including DNA arrays, genotyping-by-sequencing and mass spectroscopy, to select a robust and efficient panel of 101 SNPs as the ISAG core panel for cats. The SNPs are distributed across all chromosomes including two on the X chromosome and an XY pseudo-autosomal sexing marker (*zinc-finger XY*; *ZFXY*). A population study demonstrated that the markers have an average polymorphic information content of 0.354 and a power of exclusion greater than 0.9999. The SNP panel should keep testing affordable while also allowing for the development of additional panels to monitor health, phenotypic traits, hybrid cats and highly inbred cats.

Keywords breeds, DNA profile, genetic testing, single nucleotide polymorphism

Introduction

The domestic cat (*Felis catus*) has been a companionship animal for millennia. The transition of humans from hunter–gatherers to farmers produced a vacant ecological niche for cats, which has led to their 'domestication' and symbiotic relationship with man. In ancient Egypt, cats were treated as divine manifestations of the goddess Bastet, making frequent appearances in hieroglyphs and as votive offerings to the gods once mummified (Zivie & Lichtenberg 2005; Malek 2006). As technological advances in sailing improved, cats expanded from the Old World to inhabit all continents and most islands, except for Antarctica (Driscoll *et al.* 2007). Ever since, cats have never left the side of humans; they are present in one in four households within the Netherlands (Dibevo 2020) and one in three households in the USA (American Pet Products Association 2019).

Address for correspondence

L. A. Lyons, Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri – Columbia, E109 Vet Med Building, 1525 E. Rollins St, Columbia, MO 65211, USA.
E-mail: Lyonsla@Missouri.edu

Accepted for publication 20 May 2021

With active worldwide registries representing over 60 cat breeds today, the domestic cat remains immensely popular and the investment associated with the breeding, sale and purchase of purposely bred cats necessitates accurate individual and lineage validation (Cat Registries 2020). DNA profiling is the most accurate means of validating pedigrees and confirming a cat's identity.

Routine identification and parentage verification in many animal species are currently based on well-characterized short tandem repeat (STR) panels, such as in horses (Binns *et al.* 1995; Van de Goor *et al.* 2010, 2011), cattle (Heaton *et al.* 2002), domesticated dogs (Van Asch *et al.* 2009), domesticated cats (Lipinski *et al.* 2007) and even species of lesser economic importance like pigeons (de Groot & Van Haeringen 2017). They result in high rates of individual identification and parental exclusion. In addition to routine applications, DNA profiling also plays an important role in forensic casework (Menotti-Raymond *et al.* 1997; Kun *et al.* 2013; Wictum *et al.* 2013; Brooks *et al.* 2016) and phylogenetic research (Lipinski *et al.* 2008; Kurushima *et al.* 2013; Sauther *et al.* 2020).

Panels were developed based on STRs owing to their high discriminatory characteristics, which result in a high rate of parental or identification exclusion. Probability of exclusion (PE) values of greater than 0.99 in many species can be obtained by combining approximately 15 polymorphic STR markers in one multiplex PCR (de Groot & Van Haeringen 2017). A transition to SNPs for parentage verification and identification for production animals has attracted much interest (Heaton *et al.* 2002; Holl *et al.* 2017; Wu *et al.* 2019) and is underway for many species. Although SNPs have long been recognized as important genetic markers for identification, in general, a polymorphic STR has higher polymorphic information content (PIC) than a given SNP (Krawczak 1999). For SNPs, the decreased variability results in a lower resolution power, which means that more SNPs are needed to provide the same parentage discriminating power compared with an STR panel (Van Eenennaam *et al.* 2007). The improvements in fluorescence-based fragment analysis technologies, which are amenable to STR genotyping, also gave STRs an advantage for DNA profiling. However, the recent improvements in sequencing technologies and reduction in costs have returned the advantage to SNPs. Hundreds of SNPs can now be genotyped at roughly the same cost as STR genotyping, and SNP genotyping is more conducive to robotic automation, which can increase testing efficiency and accuracy with additional cost benefits.

An additional advantage of genotyping SNPs over STRs is that parentage and identification panels can be combined with disease and phenotypic variants to create more efficient and robust genotyping panels. Additionally, genotyping for SNPs is often easier, as genotyping for STRs is based on variation in fragment sizes, which takes expertise in discerning accurate calls. In addition, various technologies can be used for SNP genotyping, such as GBS, DNA arrays or mass spectroscopy, which utilize robust variant callers to discern the different nucleotides. Further, these technologies allow the combination of different species onto a single testing platform. Thus, commercial genotyping laboratories have more flexibility in the services they can provide since markers for a low-volume species, such as cat, can be included on a high-volume multiplex platform, such as one designed for use in cattle. Therefore, services can be provided more cost-effectively and efficiently.

The current panel for domestic cats, widely applied in parentage verification casework, is based on STRs (Lipinski *et al.* 2007). However, millions of SNPs have been identified in domestic cats and have been applied to cat breed studies (Kurushima *et al.* 2013), the development of a 63K DNA array (Gandolfi *et al.* 2018) and exploration of GBS in cats (Longeri *et al.* 2019). As part of the discussions at the Applied Genetics of Companion Animals workshop at the International Society of Animal Genetics (ISAG) meeting organized in 2016, a need to explore SNP markers for cat parentage and identification was recognized. A review of three independent SNP panels with genotyping data from different groups was presented at the ISAG workshop in 2017 in Dublin, Ireland, leading to the first cat SNP-based comparison test being conducted in Lleida, Spain in 2019. This paper presents the effort of the ISAG Applied Genetics of Companion Animals workshop to develop a standardized SNP panel for the domestic cat. An evaluation is presented of 120 SNPs for the domestic cat for use in parentage testing and individual identification from the ISAG 2019 cat comparison test. The statistical parameters of the SNPs are presented, as well as a comparison with the ISAG-recommended STR panel for domestic cats. Marker concordance is demonstrated via the ISAG comparison test.

Materials and methods

Candidate SNP selection

The participants in the cat STR comparison test for ISAG were polled for information regarding SNP usage in cats. Three groups had SNP data available: (i) cat SNP dataset A, a published 148 SNP panel used for cat ancestry and population diversity (Kurushima *et al.* 2013) with data from approximately 2000 cats; (ii) cat SNP dataset B, a 99 SNP panel developed by Orivet and Neogen from publicly available DNA array data (Gandolfi *et al.* 2018), including genotypes from approximately 200 cats; and (iii) cat SNP dataset C, a privately developed 18 SNP panel from public cat genomic data with private genotypes on approximately 2000 cats (G. Cothran, pers. comm.). For cat SNP dataset A, a subset of SNPs selected were also included on the cat 63K array (109 of 148) (Gandolfi *et al.* 2018) and were also mostly included in the SNP panel used in forensic applications (Brooks *et al.* 2016). These 109 SNPs and the SNPs of

datasets B and C were prioritized for selection in the SNP panel under development and described here. To focus on markers that would have a high discriminatory power and independent assortment, the inclusion criteria included SNPs with a heterozygosity greater than 0.30 (as determined from estimates provided). Based on cat reference assembly *Felis catus* 6.2 Montague *et al.* (2014), SNPs were selected to be at least 5 Mb apart and to represent both arms of non-acrocentric chromosomes.

ISAG 2019 cat SNP comparison test

The ISAG Applied Genetics of Companion Animals cat comparison test was conducted to determine the concordance of SNP genotypes as performed by different laboratories using different instrumentation and technologies. Twelve laboratories (denoted as Labs 1–12) completed the comparison test of the cat parentage and identification an SNP panel with 20 animals using the MassArray System for mass spectroscopy (Agena Biosciences), Illumina-based arrays and/or Ion Torrent G5 GBS (AgriSeq Thermo Fisher Scientific). For assay design, each laboratory was provided with SNP information and approximately 100 bp of sequence flanking the candidate SNPs (File S1). The ISAG duty laboratory (University of California Davis, Veterinary Genetics Laboratory) extracted and distributed 22 DNA samples, including two reference samples with provided genotypes, and 20 samples from random-bred or mixed-breed cats, all obtained in the USA. Approximately 1.5 μg of DNA was distributed to the 12 laboratories worldwide. SNP genotype data from one of the control cats (CCL-94, CRFK; ATCC® CCL-94™, Manassas, VA, USA) were provided by Neogen Genomics to standardize the nucleotide base calls and to support selection of the proper DNA strand for reporting. Laboratories provided genotyping data to ISAG for comparison in Excel worksheets. The consensus genotypes across all 12 laboratories were determined. The accuracy of each SNP and laboratory was measured as 'relative', which considered only genotyping errors, and 'absolute', which also included missing genotypes (blanks).

SNP panel population data

To assess the performance characteristics of the developed SNP panel, 3748 samples originating from 41 pedigree cats and a mixed-breed/random-bred population were commercially genotyped for 118 of the 120 SNPs at the Dr Van Haeringen Laboratorium BV, Wageningen, Netherlands. Two markers passed the design for GBS (chrB3.129823001 and chrD3.86169540), but had inconsistent call rates and performance; therefore, no further data was collected on these markers. The X-linked markers were also excluded from analyses. Global performance characteristics including observed homozygosity (H_O), expected homozygosity (H_E), PIC, average non-exclusion probability for one candidate parent, average non-exclusion probability for one candidate parent given the genotype of a known parent of the opposite sex, average non-exclusion probability for a candidate parent pair, average non-exclusion probability for identity of two unrelated individuals, average non-exclusion probability for identity of two siblings, estimated null allele frequency and HWE (Hartl & Clark 2006) were determined with the program CERVUS v3.1.3 (Marshall & Hodgson 1998; Kalinowski *et al.* 2007). In addition, expected heterozygosity and PIC were estimated for individual breeds.

Results

Candidate SNPs

The panel of SNPs selected for use in the cat parentage SNP panel and the 2019 cat ISAG comparison test included 120 SNPs, including four SNPs on the X chromosome and the *ZFXY* locus. The panel was composed of 31 SNPs from cat SNP dataset A, 79 SNPs from dataset B and 10 SNPs from dataset C (File S1).

The SNPs were originally selected at different times, by different investigators, using different assemblies of the cat genome. Dataset A was selected from an early genome build (Pontius *et al.* 2007) and later remapped to version 6.2 cat assembly (Montague *et al.* 2014). Dataset B was selected from the 63K Illumina array, where these SNPs were selected from cat assembly version 6.2 (Montague *et al.* 2014). Dataset C SNPs was selected from cat assembly version 6.2 but was not supplied with location identifiers. The SNPs selected for the cat parentage SNP panel were remapped to *Felis catus* 8.0 (Li *et al.* 2016) and the latest cat genome assembly *Felis catus* 9.0 (Buckley *et al.* 2020) using the flanking sequences (File S1). Four laboratories (University of Missouri, University of Milan, Mars and AgriSeq ThermoFisher Scientific) independently repositioned the SNPs and positions were concordant between all remapping efforts (File S1). However, four SNPs were not identified in the *Felis catus* 9.0 genome assembly (chrB3.49170524 and chrB3.143855324) or mapped to two different regions (chrC1.45295530 and chrE2.13480422). The average distance between SNPs was 22.9 Mb, ranging from 177 kb between two SNPs on chromosome D2 to 76 Mb for two SNPs on chromosome C2 (Table 1).

Cat SNP panel nomenclature

The positions of the SNPs will undoubtedly change with improvements in the assemblies of the cat genome. However, to allow cross-laboratory comparisons, a standard nomenclature for ISAG has been adopted that should not alter even with updates to new assemblies. The SNPs have therefore been named by chromosome followed by the number in the ISAG panel. The SNP names, the sequence with 200 bp flanking the SNP and the positions

Table 1 Distribution and density of International Society of Animal Genetics (ISAG) cat parentage panel SNPs

Chromosome	Size[1]	SNPs	Average distance[1]	Chromosome	Size[1]	SNPs	Average distance[1]
A1	242.5	10	22.20	D2	90.9	5	14.52
A2	172.6	10	15.62	D3	97.3	4	15.41
A3	144.0	6	18.13	D4	97.1	5	12.72
B1	208.7	8	23.19	E1	64.0	3	206.47
B2	156.7	4	23.71	E2	65.1	2[1]	31.09
B3	150.7	6	17.70	E3	44.9	2	4.64
B4	145.6	6	20.71	F1	72.6	4	10.10
C1	223.9	9	19.83	F2	86.0	5	11.29
C2	161.9	5	26.65	Xq	131.9	2	18.54
D1	119.0	4	19.33	Total		100	17.66

[1]Sizes and average distances between SNPs on the same chromosome in Mb. *ZFXY* is an additional marker. Distances are based on *Felis catus* 9.0 positions (https://www.ncbi.nlm.nih.gov/assembly/GCF_000181335.3#/st; File S1). Total length of the reference genome is 2 521 846 836 bp.

corresponding to the three builds of the reference assembly for the cat, including version 6.2 (Montague *et al.* 2014), version 8.0 (Li *et al.* 2016) and version 9.0 (Buckley *et al.* 2020), are presented in File S1.

ISAG 2019 cat SNP comparison test

The two SNPs that did not map to the cat version 9.0 assembly were not considered in the comparison test analyses. Considering the remaining 118 SNPs and the 20 cat samples utilized in the cat comparison test, each laboratory was expected to report 2360 genotypes (File S2). Considering 12 participating laboratories, 28 320 overall datapoints were expected. Seven laboratories provided between 2337 and 2357 genotypes. These seven laboratories all performed genotyping using an Illumina DNA array. Interestingly, in all seven cases, three genotypes were missing for the same cats and SNPs (chrX:157577155 for Cat 11 and chrB1:161403614 for Cats 10 and 18). One laboratory did not report data for SNP C1:45295530, one of the duplicate mapping markers. In comparison with the consensus genotypes, seven laboratories had nearly perfect genotyping results (Table 2). Four laboratories used the same array design, two laboratories each had an independent array design and two designs were unknown.

One laboratory used Agena MassArray for genotyping and had high relative accuracy (99.74%), with only six erroneous genotypes but with 61 missing datapoints (Table 2). Two SNPs (chrC2:2254710 and chrF2:67965848) had no reported data as they did not have a design for mass spectroscopy, implying 20 additional missing genotypes for each marker. The absolute genotyping accuracy was 97.16%.

Three laboratories explored GBS, the newest technology for SNP genotyping. No data were reported for six to nine different SNPs. Four SNPs (chrB1:69970470, chrB2:39410270, chrB4:156816042 and chrX:5996958) consistently had no reported genotypes across the three laboratories using this technology. Two laboratories did not report data for SNP chrA1:66285706. Seven additional SNPs were reported by only one laboratory. Therefore, the relative genotyping accuracies (98.02–99.06%) and the absolute genotyping accuracies (89.87–93.35%) were lower for the GBS technology.

Overall, the concordance between laboratories while using different genotyping platforms was over 97%. The

Table 2 ISAG cat comparison test 2019 summary

Labcode	Blank	Consensus genotypes	Errors	Relative genotyping accuracy	Absolute genotyping accuracy	Genotyping platform
Four laboratories	3	2357	0	100	99.87	Illumina[1,2]
One laboratory	3	2356	1	99.96	99.83	Illumina[1]
One laboratory	4	2356	0	100.00	99.83	Illumina[3]
One laboratory	7	2353	0	100.00	99.7	Illumina[4]
One laboratory	23	2337	0	100.00	99.03	Illumina[4]
One laboratory	61	2293	6	99.74	97.16	MassArraym
One laboratory	136	2203	21	99.06	93.35	Ion S5 XL GBS
One laboratory	190	2127	43	98.02	90.13	Ion S5 XL GBS
One laboratory	214	2121	25	98.84	89.87	Ion S5 XL GBS

[1]Same array design.
[2]Independent Illumina array design.
[3]Independent Illumina array design.
[4]Unknown array design.

Table 3 SNPs with lower accuracy values for the ISAG 2019 cat comparison test

SNP	Relative	Absolute	SNP	Relative	Absolute
B1.69970470	100.00	75	A2.33979426	99.54	90.42
B2.39410270	100.00	75	E2.13480422[1]	100.00	91.25
B4.156816042	100.00	75	B3.129823001	100.00	91.67
X.5996958	100.00	75	B3.44006038	100.00	91.67
A1.66285706	97.95	79.58	C1.45295530[1]	100.00	91.67
E3.15324152	91.51	80.83	C2.2254710	100.00	91.67
D3.86169540	99.52	87.08	F1.24753896	100.00	91.67
B1.161403614	92.92	87.5	F2.67965848	100.00	91.67
D1.96334367	90.99	88.33	D2.70137294	99.56	94.17
C2.153875641	91.49	89.58			

[1]Eliminated owing to poor mapping to cat genome assembly *Felis catus* version 9.0.

genotyping rates are presented in Table 2 and the SNPs with the lower genotyping accuracies are presented in Table 3. The complete absence of an SNP in the MassArray or GBS design, which caused the majority of the missing genotypes (87.15%), led to an absolute genotyping accuracy of less than 100%. The majority of errors owing to incorrect genotypes were identified in GBS-produced data. Seventeen SNPs (17 of 118; 14.41%) had an absolute genotyping accuracy below 100%, ranging from 94.17 to 75.00%, eight SNPs had relative accuracy genotyping rates below 99.60% and five SNPs had relative accuracy genotyping rates below 98.00% (Table 3).

Cat breed population statistics

The calculations of the population statistics were corrected for the eight markers (chrA1:66285706, chrB1:161403614, chrB2:39410270, chrB3:129823001, chrB4:156816042, chrC1:45295530, chrD3:386169540 and chrE2:13480422) that were dropped during initial analysis (poor GBS design) (File S3). Also, two markers (chrB1.69970470 and chrD3.86169540) were unintentionally left out of the GBS panel design and were therefore not included in the population analysis. The results per marker of the analysis using CERVUS version 3.1.3 can be found in File S3. For 98 autosomal markers combined, a PE1, PE2 and full parent combo value of greater than 0.9999 was observed. The average call rate for all cats was 0.9851. Two call rates less than 0.95 were observed but in two breeds represented by only one cat each.

The cat population survey dataset included 3748 cats, including 1228 random-bred cats and 2460 cats representing 41 breeds. Twenty-eight breeds were represented by five or more individuals, totaling 2435 cats. The relationships of the cats within the breeds were not determined; thus, H_E and PIC may be underestimated. The average H_O was 0.403, the average H_E was 0.462 and the average PIC was 0.354 across all cats. Potential null alleles were estimated in a range from 0.0046 for chrC1.211548622 to +0.2022 for chrA3.117471748. The power of exclusion for the panel was estimated as greater than 0.9999 (File S3). The range of the average H_E based on breeds with five or more individuals was as low as 0.298 for Burmese and as high as 0.456 for Siberians for well-established breeds. The range of the average PIC based on breeds with five or more individuals was a low of 0.222 for Burmese and a high of 0.347 for Siberians (Table 4). Additional breeds with low H_E and PIC included Birman ($H_E = 0.324$, PIC = 0.257) and Tonkinese ($H_E = 0.334$, PIC = 0.249). The genotypes for the complete dataset are presented as File S4.

Suggested ISAG core SNP panel for the domestic cat

The four SNPs not identified in the *Felis catus* 9.0 genome assembly (chrB3.49170524 and chrB3.143855324) or mapped to two different regions (chrC1.45295530 and chrE2.13480422) are suggested to have been eliminated from the core SNP panel. Four SNPs probably had poor design for GBS, including SNPs chrB1:69970470, chrB2:39410270, chrB4:156816042 and chrX:5996958 and potentially chrA1:66285706. Two SNPs (chrC2:2254710 and chrF2:67965848) failed the design for mass spectroscopy. The 10 SNPs with relative accuracy genotyping rates below 99.50%, of which five had relative accuracy genotyping rates below 98.00%, are also suggested to have been eliminated for the first cat core SNP panel. Two SNPs with individual PIC scores below 0.2 are among the 10 SNPs eliminated for poor relative accuracy genotyping rates. Therefore, the selected ISAG core panel for parentage and identification consists of 101 SNPs that include one sexing marker (*ZFXY*) and two markers on the X chromosome (File S1).

Discussion

The cat has 18 autosomes and sex chromosomes X and Y. Only two chromosomes (F1 and F2) are acrocentric, thus the cat genome is represented on 35 chromosome arms, the pseudo-autosomal region of the X chromosome and the Y chromosome. The cat genome is approximately 2.6 Gb, thus

Table 4 SNP panel diversity in 41 cat breeds and random bred cats

Breed	*N*	Average call rate	H_E	PIC	Breed	*N*	Average call rate	H_E	PIC
Siberian cat	56	0.9896	0.4555	0.3474	Egyptian Mau	6	0.9813	0.3524	0.2589
British Longhair	36	0.9804	0.4436	0.3386	Oriental Shorthair	14	0.9767	0.3523	0.2712
LaPerm	5	0.9939	0.4423	0.3116	Tonkinese	8	0.9681	0.3338	0.2485
Russian blue	6	0.9966	0.4397	0.3139	Birman	48	0.9915	0.3242	0.2574
Ragdoll	462	0.9869	0.4376	0.3393	Burmese	7	0.9679	0.2976	0.2217
Norwegian Forest Cat	88	0.9879	0.4371	0.3360	Balinese	4	0.9923	0.3069	0.2150
Highland Fold	14	0.9883	0.4362	0.3278	Cheetoh	4	0.9949	0.4136	0.2848
British Shorthair	275	0.9856	0.4360	0.3377	Selkirk Rex	4	1.0000	0.3684	0.2557
Scottish Fold	38	0.9919	0.4357	0.3345	Bambino	2	0.9949	0.3741	0.2186
Exotic Shorthair	13	0.9623	0.4341	0.3247	Ragamuffin	2	0.9796	0.3503	0.2014
Savannah	36	0.9830	0.4243	0.3264	Toyger	2	0.9898	0.3793	0.2250
Persian	29	0.9870	0.4184	0.3218	Deutsch langhair	1	1.0000	0.3980	0.1492
Nebelung	5	0.9673	0.4089	0.2893	Don Sphynx	1	0.9898	0.4286	0.1607
Thai	55	0.9844	0.4030	0.3124	European short hair	1	0.9490	0.3367	0.1263
Sphynx	60	0.9840	0.4003	0.3126	Korat	1	1.0000	0.2755	0.1033
Maine Coon	984	0.9902	0.4002	0.3148	Kurillian Bobtail	1	1.0000	0.4082	0.1531
Devon Rex	14	0.9883	0.3834	0.2928	Munchkin	1	1.0000	0.4490	0.1684
Somali	18	0.9875	0.3834	0.2941	Ukrainian levkoy	1	0.9184	0.3878	0.1454
Bengal	92	0.9862	0.3721	0.2933					
Abyssinian	28	0.9931	0.3657	0.2857	Random bred	1295	0.9903	0.4617	0.3538
Cornish Rex	7	0.9913	0.3620	0.2653	Breed only (28 $N \geq 5$)	2428	0.0848	0.3921	0.2977
Siamese	24	0.9915	0.3569	0.2790	Total ($N \geq 5 = 3723$)	3748	0.9875	0.4269	0.3257

H_E, Expected homozygosity; PIC, polymorphic information content.

an average spacing for unlinked markers would be approximately 50 Mb, implying that a perfect distribution would require approximately 52 SNPs if using Haldane's mapping function ($\theta = 0.01 = 1$ cM $= 1$ Mb; Ott 1999).

The final cat SNP DNA profiling panel of consists of 100 SNPs and an SNP (*ZFXY*) distinguishing the X and Y chromosomes. Twenty SNPs were eliminated from the original 121 candidate SNPs; however, laboratories may continue to use the dropped SNPs in secondary or private profiling panels. Various factors, e.g. multi-mapping to or absence from the newest *Felis catus* version 9.0 genome assembly, non-concordance and poor performance across genotyping platforms, led to a smaller, more cost-effective panel. However, with a decreased panel size, solving complex parentage cases (e.g. parentage in highly inbred populations) becomes increasingly difficult (McClure *et al.* 2015). Therefore, secondary SNP panels have been developed for many species as population data are acquired to select informative markers balanced with effective power to address complex parentage verification or forensic cases.

The distribution of the SNPs needs to be balanced with parentage exclusion and identification as not all SNPs have the same frequency in all cat breeds. The current cat SNP panel has markers present on all chromosomes, including two SNPs on the X chromosome and an SNP distinguishing the X and Y chromosomes for *ZFXY*. This, in combination with the PIC values in the selected markers, supports the high exclusion power to discriminate individual animals as well as the power to exclude parents in parentage cases.

Lipinski *et al.* (2007) evaluated 19 selected microsatellite markers with the mean number of alleles for all markers 4.3 in pedigree breeds and 6.5 in random-bred cats. Additionally, the mean PIC was 0.47 in pedigree breeds and 0.63 in random-bred cats. While the dynamics for SNP markers is not comparable with that of microsatellite markers (two alleles vs. potentially many alleles for STRs), the SNP markers selected for this panel are informative, with an average PIC of 0.35. Two markers with individual PIC scores below 0.2 (chrE3.15324152 and chrD1.96334367) are suggested to have been eliminated from the panel. Development of a secondary panel of SNPs with data from rare and highly inbred breeds and populations is encouraged. Cat breeds shown to have low genetic variation, including Burmese and Birman and the closely related Tonkinese (Lipinski *et al.* 2008; Kurushima *et al.* 2013), also had the lowest H_E and PIC values in this SNP panel. Therefore, the research community should identify SNPs with more variation and discriminatory power within these breeds.

Future revisions of the panel could include additional SNPs on Y chromosome and autosomes with high PIC values to balance the distribution across the genome, and also wild felid-specific SNPs for leopard cats (*Prionailurus bengalensis*), servals (*Leptailurus serval*) and jungle cats (*Felis chaus*), which are used to produce the hybrid cat breeds Bengal, Savannah and Chausie, respectively, and could be used to detect recent hybridizations. Importantly, the current panel needs to be supplemented with SNPs that increase power in parentage analysis of highly inbred

populations and breeds such as Burmese and Birman, which are considered to have the lowest diversity amongst cat breeds (Lipinski *et al.* 2008; Alhaddad *et al.* 2013; Gandolfi *et al.* 2018).

Although no specific technology should strongly influence the selection of markers, fortuitously, several laboratories capable of employing different technologies were available for the first cat SNP comparison test. Because of the abundance of SNPs and because significant investment in a particular set of SNPs had not yet been established, the cat workshop participants had the luxury of selecting SNPs that were robust across the current technologies, including Illumina iSelect Infinium arrays, Ion torrent GBS and MassArray (Agena). The historical rules of the ISAG CT indicate missing data to be considered an error. This was not clearly defined prior to data collection and some laboratories may have increased their call rates by repeating sample genotyping, thus a direct indicator of robustness of a technology cannot be fully determined. The success of GBS is influenced by the depth of sequence coverage, which has been evaluated for the cat SNP panel, demonstrating that 10× coverage is insufficient for robust genotyping for some loci (Longeri *et al.* 2019). Thus, the lower absolute call rates are not indicative of the absolute accuracy of the technology as probe design and sequence coverage need to be closely examined. Probe design was a potential concern for one SNP (chrB1.161403614), which repeatedly failed to be genotyped in different laboratories for the same two cat samples, with all laboratories using Illumina array technology. Direct Sanger sequencing of the flanking regions may indicate a polymorphism within the probe region.

The ISAG Cat DNA profiling panel currently includes 98 SNPs on 17 autosomes and two SNPs on the X chromosome. An SNP in the gene *ZFXY* differentiates the X and Y chromosomes. The average PIC for the panel is 0.354 and the PE is greater than 0.9999. As more SNP data become available for additional breeds and populations, development of a secondary panel is encouraged to rectify parentage concerns in highly inbred populations. For forensic applications, an SNP panel with inclusion of mitochondrial markers to support the analyses of degraded samples would have added value. The selected SNPs on this panel have been shown to yield accurate and robust genotypes using various technologies, including Illumina Infinium arrays, Thermo Fisher Ion G5 GBS and Agena MassArray, and should serve the feline genetics community well.

Acknowledgements

We appreciate the provision of cat DNA samples for the comparison test by the University of California, Davis Veterinary Genetics Laboratory (ISAG Cat CT Duty Lab 2019). We appreciate the assistance of Jeremy Holzner at ISAG/FASS in the coordination and administration of the cat comparison test.

Conflicts of interest

Participating commercial laboratories and SNP contributors have potential conflicts considering an interest in promoting the SNP markers that are well developed and efficient for their own technologies of choice, provided services and research purposes. However, the overall selection of markers was determined by the ISAG workshop committee and participants and selection did not favor any specific technology, commercial application or research group. Overall, the authors declare no other conflicts of interest.

Funding

Support for the project was provided in part by the University of Missouri, Gilbreath McLorn Endowment (LAL).

Data availability statement

All data are provided in the supplementary files. The cat genome assemblies are publicly available at NCBI: https://www.ncbi.nlm.nih.gov/assembly/organism/9685/latest/. Supplementary data available at figshare: https://figshare.com/s/ae1e71856eaaefc96b8f.

References

Alhaddad H., Khan R., Grahn R.A. *et al.* (2013) Extent of linkage disequilibrium in the domestic cat, *Felis silvestris catus*, and its breeds. *PLoS One* **8**, e53537.

American Pet Products Association, 2019–2020 APPA National Pet Owners Survey (2019). https://www.americanpetproducts.org/pubs_survey.asp. Accessed 08 October 2020.

Cat Registries (2020). Associazione Nazionale Felina Italiana. (https://www.anfitalia.it/site/); Australian Cat Federation (http://www.acf.asn.au/cats2.php); Cat Fanciers' Association Inc. (https://cfa.org/breeds/), Federation International Feline (http://fifeweb.org/wp/breeds/breeds_prf_stn.php); Governing Council of the Cat Fancy (https://www.gccfcats.org/Cat-Breeds); The International Cat Association (https://www.tica.org/breeds/browse-all-breeds); World Cat Federation (http://wcf-online.de/WCF-EN/wcf/wcf.html). Accessed 08 October 2020.

Binns M.M., Holmes N.G., Rolliman A. & Scott A.M. (1995) The identification of polymorphicmicrosatellite loci in the horse and their use in thoroughbred parentage testing. *British Veterinary Journal* **151**, 9–15.

Brooks A., Creighton E.K., Gandolfi B., Khan R., Grahn R.A. & Lyons L.A. (2016) SNP miniplexes for individual identification of random-bred domestic cats. *Journal of Forensic Sciences* **61**, 594–606.

Buckley R.M., Davis B.W., Brashear W.A. *et al.* (2020) A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. *PLoS Genetics* **16**, e1008926.

de Groot M. & van Haeringen W.A. (2017) An evaluation of the International Society for Animal Genetics recommended parentage and identification panel for the domestic pigeon (*Columba livia domestica*). *Animal Genetics* **48**, 431–5.

Dibevo, 2019 - 2020 National Pet Population Survey (2020). https://dibevo.nl/pers/nieuwe-cijfers-meer-dan-27-miljoen-huisdieren-in-nederland

Driscoll C.A., Menotti-Raymond M., Roca A.L. *et al.* (2007) The near eastern origin of cat domestication. *Science* **317**, 519–23.

Gandolfi B., Alhaddad H., Abdi M. *et al.* (2018) Applications and efficiencies of the first cat 63K DNA array. *Scientific Reports* **8**, 7024. Erratum in: Sci Rep. 2018 Jun 4;8(1):8746. Erratum in: Sci Rep. 2019;9(1):4664.

Hartl D.L. & Clark A.G. (2006) *Principles of Populations Genetics*, 4th edn. Oxford University Press, Oxford.

Heaton M.P., Harhay G.P., Bennett G.L., Stone R.T., Grosse W.M., Casas E., Keele J.W., Smith TPL, Chitko-McKown C.G. & Laegreid W.W. (2002) Selection and use of SNP markers for animal identification and paternity analysis in US beef cattle. *Mammalian Genome* **13**, 272–81.

Holl H.M., Vanhnasy J., Everts R.E., Hoefs-Martin K., Cook D., Brooks S.A., Carpenter M.L., Bustamante C.D. & Lafayette C. (2017) Single nucleotide polymorphisms for DNA typing in the domestic horse. *Animal Genetics* **48**, 669–76.

Kalinowski S.T., Taper M.L. & Marshal T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**, 1099–106.

Krawczak M. (1999) Informativity assessment for biallelic single nucleotide polymorphisms. *Electrophoresis* **20**, 1676–81.

Kun T., Lyons L.A., Sacks B.N., Ballard R.E., Lindquist C. & Wictum E.J. (2013) Development and validation of a canine MiniSTR panel for forensic casework. *Forensic Science International Genetics* **7**, 151–8.

Kurushima J.D., Lipinski M.J., Gandolfi B., Froenicke L., Grahn J.C., Grahn R.A. & Lyons L.A. (2013) Variation of cats under domestication: genetic assignment of domestic cats to breeds and worldwide random-bred populations. *Animal Genetics* **44**, 311–24.

Li G., Hillier L.W., Grahn R.A. *et al.* (2016) A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. *G3 (Bethesda)* **6**, 1607–16.

Lipinski M.J., Amigues Y., Blasi M. *et al.* (2007) An international parentage and identification panel for the domestic cat (*Felis catus*). *Animal Genetics* **38**, 371–7.

Lipinski M.J., Froenicke L., Baysac K.C. *et al.* (2008) The ascent of cat breeds: genetic evaluations of breeds and worldwide random-bred populations. *Genomics* **91**, 12–21.

Longeri M., Chiodi A., Brilli M., Piazza A., Lyons L.A., Sofronidis G., Cozzi M.C. & Bazzocchi C. (2019) Targeted genotyping by sequencing: a new way to genome profile the cat. *Animal Genetics* **50**, 718–25.

Malek J. (2006) *The Cat in Ancient Egypt*, 2nd edn. British Museum Press, London.

Marshall A. & Hodgson J. (1998) DNA chips: an array of possibilities. *Nature Biotechnology* **16**, 27–31.

McClure M.C., McCarthy J., Flynn P., Weld R., Keane M., O'connel K., Mullen M., Waters S. & Kearney J.F. (2015) SNP selection for nationwide parentage verification and identification in beef and dairy cattle. *ICAR Technical Series* **19**, 175–81.

Menotti-Raymond M.A., David V.A. & O'Brien S.J. (1997) Pet cat hair implicates murder suspect. *Nature* **386**, 774.

Montague M.J., Li G., Gandolfi B. *et al.* (2014) Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proceedings of the National Academy of Sciences USA*, **111**, 17230–5.

Ott J. (1999) *Analysis of Human Genetic Linkage*, 3rd edn. Johns Hopkins University Press, Baltimore, MD.

Pontius J.U., Mullikin J.C., Smith D.R. *et al.* (2007) Initial sequence and comparative analysis of the cat genome. *Genome Research* **17**, 1675–89.

Sauther M.L., Bertolini F., Dollar L.J. *et al.* (2020) Taxonomic identification of Madagascar's free-ranging "forest cats". *Conservation Genetics* **21**, 443–51.

Van Asch B., Alves C., Gusmão L., Pereira V., Pereira F. & Amorim A. (2009) A new autosomal STR nineplex for canine identification and parentage testing. *Electrophoresis* **30**, 417–23.

Van de Goor LHP, Panneman H. & Van Haeringen W.A. (2010) A proposal for standardization in forensic equine DNA typing:allele nomenclature for 17 equine-specific STR loci. *Animal Genetics* **41**, 122–7.

Van de Goor LHP, Van Haeringen W.A. & Lenstra J.A. (2011) Population studies of 17 equine STR for forensic and phylogenetic analysis. *Animal Genetics* **42**, 627–33.

Van Eenennaam A.L., Weaber R.L., Drake D.J., Penedo MCT, Quaas R.L., Garrick D.J. & Pollak E.J. (2007) DNA-based paternity analysis and genetic evaluation in a large, commercial cattle ranch setting. *Journal of Animal Science* **85**, 3159–69.

Wictum E., Kun T., Lindquist C., Malvick J., Vankan D. & Sacks B. (2013) Developmental validation of DogFiler, a novel multiplex for canine DNA profiling in forensic casework. *Forensic Science International Genetics*, **7**, 82–91.

Wu X., Xu J., Li H. *et al.* (2019) Evaluation of genotyping concordance for commercial bovine SNP chips using quality-assurance samples. *Animal Genetics* **50**, 367–71.

Zivie A. & Lichtenberg R. (2005) The cats of the goddess Bastet. In: *Divine Creatures: Animal Mummies in Ancient Egypt* (Ed. by S. Ikram), pp. 106–19. American University in Cairo Press, Cairo.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

File S1. ISAG cat SNP panel, sequences and chromosomal locations.

File S2. ISAG 2019 cat comparison test genotypes.

File S3. SNP panel cat breed population statistics.

File S4. SNP genotypes for all cats.