

## **Application of a mitochondrial DNA control region frequency database for UK domestic cats**

Barbara Ottolini<sup>1\*</sup>, Gurdeep Matharu Lall<sup>1</sup>, Federico Sacchini<sup>2\*</sup>, Mark A. Jobling<sup>†1</sup>, Jon H. Wetton<sup>†1,3</sup>

<sup>1</sup> Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH UK

<sup>2</sup> IDEXX Laboratories Ltd., Grange House, Sandbeck Way, Wetherby, West Yorkshire, LS22 7DN UK

<sup>3</sup> School of History, University of Leicester, University Road, Leicester LE1 7RH UK

<sup>1\*</sup> Present address: Department of Cancer Studies, University of Leicester, Leicester Royal Infirmary, Leicester LE1 5WW

<sup>2\*</sup> Present address: IDEXX Laboratories Italia S.r.l., Via Guglielmo Silva, 36–20149 Milano (MI), Italy

<sup>†</sup> Correspondence can be addressed to either author:

Dr Jon H. Wetton, Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK

Tel.: +44 (0)116 252 3377 Fax: +44 (0)116 252 3378

Email: [jw418@le.ac.uk](mailto:jw418@le.ac.uk)

Prof Mark A. Jobling, Department of Genetics, University of Leicester, University Road, Leicester LE1 7RH, UK

Tel.: +44 (0)116 252 3427 Fax: +44 (0)116 252 3378

Email: [maj4@le.ac.uk](mailto:maj4@le.ac.uk)

**Running head:** UK cat mitochondrial DNA database

## **Abstract**

DNA variation in 402 bp of the mitochondrial control region flanked by repeat sequences RS2 and RS3 was evaluated by Sanger sequencing in 152 English domestic cats, in order to determine the significance of matching DNA sequences between hairs found with a victim's body and the suspect's pet cat. Whilst 95% of English cats possessed one of the twelve globally widespread mitotypes, four new variants were observed, the most common of which (2% frequency) was shared with the evidential samples. No significant difference in mitotype frequency was seen between 32 individuals from the locality of the crime and 120 additional cats from the rest of England, suggesting a lack of local population structure. However, significant differences were observed in comparison with frequencies in other countries, including the closely neighbouring Netherlands, highlighting the importance of appropriate genetic databases when determining the evidential significance of mitochondrial DNA evidence.

**Keywords:** domestic cat; mitochondrial DNA; Sanger sequencing; haplotype diversity; mitotype; DNA database

## Introduction

As all pet owners will appreciate, their pet's fur is frequently found on clothing and home furnishings. Consequently the recovery of pet hairs whilst taping evidential items for fibres can help establish the likelihood that transfer occurred through contact with the clothing of a person or at a place with which a particular animal is closely associated. Although even an individual pet's coat hairs may vary in colour and morphology these characteristics can be invaluable in selecting hairs for DNA analysis which, via the shared possession of a rare genetic variant, can provide stronger evidence of origin from a common source than appearance alone. Indeed there have been instances when nuclear short tandem repeat (STR) [1] or mitochondrial DNA (mtDNA) matches [2] with cat hairs have provided vital evidence in forensic casework. The majority of transferred cat hairs have been naturally shed, and lack the large root bulb required for reliable high-discrimination STR profiling [3]. However, even hair shafts typically contain enough mtDNA for sequencing, which is still capable of excluding many potential sources despite lower discriminating power than nuclear DNA [4].

For both marker types it is necessary to create databases to assess the frequency of alleles or haplotypes within a relevant population [5]. For haploid markers such as mtDNA, where profiles are expected to be shared by many matrilineally-related individuals, the strength of the evidence is determined not only by the variability of the sequence but also the size of the geographically relevant genetic database [6], which should be large enough to accurately reflect the local diversity.

In the United Kingdom (UK) an estimated 10.2 million cats live in 26% of households [7], and regularly 'tag' the furnishings and clothes of their owners with their shed hairs. Consequently it is inevitable that cases will occur in which cat hairs could provide critical evidence. In July 2012 a dismembered human torso was washed up on a beach in Southsea, a seaside resort in Portsmouth, England (Figure 1a). The torso was wrapped in a shower curtain and sealed within a black plastic bin liner. The victim was identified as a local resident, and one of his associates became the prime suspect after voluntarily presenting himself at a police station in a confused and intoxicated condition stating that he

had done something “really bad” and thought that he may have killed someone, but was unable to remember any details. Whilst no human traces other than those of the victim were discovered associated with the body, eight cat hairs were found amongst a population of fine animal hairs recovered on tapes from the outside of the shower curtain. Both suspect and victim were cat owners, and the hairs matched the colour and morphology of those of the suspect’s cat, but were dissimilar to those of the victim’s cat (D. Hopwood pers. comm.).

The reporting officer of the forensic service provider (LGC) carrying out the examination invited the Veterinary Genetics Laboratory (VGL) at UC Davis to compare DNA recovered from the hairs on the shower curtain with the DNA of the cats belonging to suspect and victim. The hairs from the curtain yielded insufficient DNA for STR profiling [8], but a 402-bp fragment of the mtDNA control region flanked by repeat sequences RS2 and RS3 was successfully amplified and Sanger sequenced. The resultant sequence did not match any of the 149 mtDNA haplotypes (mitotypes) seen within a database of 1394 cats from 25 distinct populations in the USA, Asia and Europe [9]. The victim’s cat possessed a different mitotype and was excluded as a potential source, but the suspect’s cat was a match for this globally rare sequence.

Hampshire Constabulary, the investigating police force, required an opinion on the relevance of the database used by VGL for determining the evidential significance. Given that there was significant variation in the frequencies of the 12 common mitotypes between countries it was clear that a database relevant to the UK and possibly the immediate area of the crime scene was needed, in order for the evidence to be evaluated appropriately. We describe here the creation of a UK domestic cat mtDNA database spanning 402-bp of the control region, its application in the investigation of this case, and its implications for future uses of domestic cat hair profiling within the UK and beyond.

## Materials and Methods

Surplus blood collected during routine examinations of 152 cats at 105 veterinary surgeries throughout England was provided by IDEXX, the largest national veterinary diagnostic company. With only six weeks to build and analyse the database before the court submission date, 32 samples were accumulated over 22 days from seven surgeries in the Portsmouth postal code area, including 23 samples from two surgeries within a mile of the homes of both suspect and victim and the location where the body was recovered (Figure 1a). These were supplemented with 120 samples collected from male cats over two days from elsewhere in England with a mean of 1.25 samples per surgery.

The ages of 93% of donor cats were known and there was only one case where two samples from the same location had the same age and could thus represent duplicate samples from the same individual. With an average age at sampling of 12.2 years, close to the average life expectancy of cats, and a mean number of cats per cat-owning household of 1.66 [7] there is a very low probability that maternally related individuals were sampled. Cat geographic origins matched closely with human population density (Figure 1b) and also the density of cats (Figure 1c), which in the UK are predominantly “owned” rather than feral. Breed cats accounted for 12.5% of the total (N=19), while 81.5% were non-breed and 9% were of undefined breed status; this is in close accordance with the 89% of non-breed cats recorded by microchipping companies [10]. Therefore we believe this is a representative sample of cats within England.

DNA was extracted from 200µl EDTA-treated whole blood using a QIAcube robotic platform and the QIAamp Mini (QIAGEN) blood kit according to the manufacturer’s protocol. After NanoDrop quantification (Thermo Scientific) all samples were diluted to 5ng/µl before PCR amplification with primers JHmtF3—gatagtgccttaatcgtgc and JHmtR3—gtcctgtggaacaatagg following a published procedure [9]. Subsequent Exonuclease 1 and shrimp alkaline phosphatase treatments degraded the primers and removed excess dNTPs prior to Big Dye v3.1 terminator sequencing of the 492-bp products on an ABI3730 Genetic Analyzer. Forward and reverse sequences generated with the above primers were aligned using CodonCode Aligner (CodonCode Corp.) and the consensus sequences of the 402-bp region previously reported [9] were

compared across all individuals in the dataset. The relationship between mitotypes was investigated with median-joining networks constructed with Network 5 software (Fluxus Technology Ltd). Population frequencies of mitotypes were extracted from several published reports and geographic variation visualised using pie charts created in Excel. Populations were also compared using MDS plots created with the R package isoMDS. Upper-bound frequency estimates for court reporting were calculated with commonly used methods [5]. Random match probability (RMP) was calculated as  $(1 - \sum p_i^2)$  where  $p$  is the frequency of the observed haplotype.

## Results

Sixteen distinct mitotypes, defined by 25 variable sites, were observed among the 152 cats (Table 1; Figure 2a); 137 (90%) possessed the widespread mitotypes A-J previously reported [9], whilst another carried the previously observed outlier mitotype OL1. Mitotype B1, seen seven times in the published database of 1394 cats [9], was also observed seven times among the 152 English cats, suggesting that the frequency might be locally increased. Four previously unreported mitotypes were also found, namely B-UK1 (N=3), B-UK2 (N=1), C-UK1 (N=2) and C-UK2 (N=1).

B-UK1 matched the sequence of the evidential hairs and the suspect's cat and represents a variant of mitotype B that is globally rare, but within the UK is merely uncommon. B-UK2 is a singleton mitotype, differing by one base from B1. C-UK1 (N=2) & C-UK2 (N=1) are both distinguished by single SNP variants upon the globally abundant mitotype C, which has its highest reported frequency in our UK dataset. Since B-UK2 and C-UK2 have not been observed elsewhere, and are confined to single individuals in our dataset, they have been recorded as globally unique observations.

In the English dataset there are 23 variable positions within the 402-bp and two sites (134 and 181) which, in common with all other sequenced cats, differ from the cat mtDNA reference genome (U20753, [11]). Both of the English unique mitotypes are generated by recurrent transitions - B-UK2 shares 173G/A with mitotypes A and I, and C-UK2 169T/C separates it from C, just as it does B from D.

Representation of the mitotypes in the form of a median-joining network (Figure 2a) illustrates the similarity of frequencies in the local (N=32) and rest-of-England (N=120) subsets; they do not differ significantly ( $p=0.92$ , Exact Test of Sample Differentiation based on haplotype frequencies). However, the English dataset differs significantly ( $p<1\times10^{-5}$ ) from the recently published Dutch population (Figure 2b) [12]. Despite the geographic proximity of England and the Netherlands (Southsea is as close to Amsterdam as it is to the furthest sampling location in England) the English Channel has allowed the two populations to diverge considerably. In addition, the English data set is

significantly different from those populations grouped in the published Grahn et al. [9] dataset (Figure 3; Table 2).

A comparison of the geographic distribution of the twelve main haplogroups between the grouped populations (Figure 3a) and a multidimensional scaling (MDS) plot of  $F_{st}$  values based on major haplogroup frequencies (Figure 3b), shows that geographic proximity is a poor predictor of the similarity of populations when they are separated by a physical barrier. The neighbouring English and Dutch populations differ greatly ( $F_{st}=0.1861$ ,  $p<1\times10^{-5}$ ) whereas the English and US populations are much more similar to each other ( $F_{st}=0.0087$ ,  $p\sim0.04$ ), raising interesting questions about the source of founder populations and the role of genetic drift.

The random match probability (RMP) was 0.24 for Portsmouth, 0.18 for the rest of England dataset and 0.19 for the combined national population. As there was no significant differentiation between the 32 Portsmouth area cats and the 120 from the rest of England it is legitimate to pool the samples as a single English population. The evidential profile shared with the suspect's cat was seen three times in our dataset - once in the Portsmouth area 3.1% (1/32) and twice elsewhere - giving a national frequency of 2.0% (3/152). Based upon these observations we calculated upper-bound frequency estimates as shown in Table 3 for court reporting [5].

## Discussion

In this study we have generated a database of English domestic cat mitotypes, used this to interpret the strength of evidence in a particular criminal case, and undertaken a comparison of regional cat mtDNA databases. This analysis reveals high degrees of population structure, emphasising the importance of careful choice of relevant databases for casework.

Of the twelve "universal" mitotypes (A-L) which had a frequency of  $>0.7\%$  among the 1394 cats in a large published database (Table 2; [9]), ten (A-J) were found in the English sample at frequencies from 30.3% (C) to 0.7% (G & J), whilst mitotypes L and K, which are strongly associated with the Eastern Mediterranean/ Middle East and US populations respectively, and may represent



local expansions, were absent from the UK database. Most geographically separated populations show evidence of locally abundant variants that are significantly less common elsewhere. In the Netherlands mitotype A (represented by extended 641-bp mitotypes NL-A1 to NL-A11 inclusive [12]) was present in 66.4% of cats, compared with just 24.3% in England. Minor mitotypes, <0.7% in the large dataset [9], can also reach much higher frequencies in local populations: mitotype B6, seen just twice in the continental US random-bred population, once in the Netherlands (equivalent to NL-B7 in a Maine Coon breed cat) and not at all in the UK, varies in frequency from 0% in Vancouver (0/15) to 28.6% in Ottawa (8/28) in a survey of Canadian cats [13].

In most countries studied so far, the universal mitotypes account for the majority of individuals (e.g. 83.6% of cats (Table 2; [9])). However, a further 30 subtypes of A-L, each observed more than once but at frequencies <0.7%, account for an additional 8.0%, while singleton haplotypes together account for 7.5% of cats. In England, 90.7% of cats carried major mitotypes (A-L), 7.9% had subtypes (B1=7, B-UK1=3, C-UK1=2, OL1=1), and just 1.3% had unique mitotypes (B-UK2 & C-UK2). The low frequency of unique mitotypes is in striking contrast to the published European dataset (Table 2; [9]) which showed the highest proportion of globally unique mitotypes, with more than half being represented by novel singletons (although each of the national datasets comprised fewer than 25 individuals - Germany (12/21) & Italy (13/23)). This result had suggested that the English population might be very different to that of the US, which proved not to be the case.

Interestingly there is also a low proportion of unique mitotypes in the Dutch population when considering just the 402-bp rather than the extended 641-bp region ( $5/113 = 4.4\%$ ), and one of these (NL-B6) actually matches B-UK1 and thus represents an example of a rare trans-national mitotype. These unique-mitotype frequencies are much more in keeping with those of other regions in the large published dataset (Table 2; [9]), where the proportions of unique mitotypes were < 1% in the Indian Ocean trade route countries, ~4% in Iran/Iraq and the Eastern Mediterranean, 5% in the USA, and 14% in East Asia. Grahn et al. [3] described computer simulations based on random resampling of previously published cat populations which demonstrated that a database size of

150 with a small number of observed mitotypes (16) and low frequency of unique mitotypes (<2%), as in our study, should achieve 99% saturation of mitotype representation (defined as the point where a sample size increase of 100 individuals corresponds to an increment in the diversity measure of <1% [14]). In their study a random-bred Eastern US cat population sampled in New York (N=101) and Florida (n=50) displaying similar allele frequencies to the UK but slightly higher diversity with 21 different mitotypes, five more than the UK, attained 99% saturation at a simulated sample size of 100. The same sample size was also considered appropriate based upon simulations using European random-bred cats which included many more novel mitotypes. Therefore we believe that our sample size provides adequate representation of the mitotype diversity within the UK and that any unsampled variants will be rare.

Our samples included 19 cats from 12 breeds, of which the Burmese was most common with three representatives. Fifteen pedigree cats belonged to the three main mitotypes (A-C) but two of the less common mitotypes in the English population were also represented - D by two Burmese and a Tonkinese, and E by the sole Bengal. As the Tonkinese is a derivative of the Burmese, which in turn has been shown by both STR and SNP typing to have almost the lowest levels of variability among cat breeds [15, 16], it is unsurprising that they share mitotype D with five of 18 Burmese in the previously published dataset (Table 2; [9]). Whilst our sample of pedigree animals is too small to draw conclusions regarding differences between breeds, this result hints at increased frequencies associated with certain breeds which should be taken into account if a pedigree animal is thought to be the source of an evidential hair. In England, and probably many other countries, the majority of dogs can be visually assigned to particular breeds, the popularity of which can fluctuate, potentially leading to significant shifts in breed and mitotype frequencies in those instances where there is an association with breed [17]. By contrast, for cats random-bred animals are clearly in the majority and fashionable breeds only exert a very slight influence on national mitotype frequencies. In both species, individuals assigned to a breed usually display a range of mitotypes representative of the regional population within which the breed was developed. For example, the Burmese and derivative breeds whilst originating from imports from Burma (Myanmar)

where the aforementioned mitotype D is likely to be abundant (see Figure 3a) have been rapidly expanded from the founding stock by introducing other matriline in both the US and the UK.

The gene pool of random-bred cat populations has also been affected by founder effects and admixture. The establishment of a large pool of predominantly feral cats living in close association with human habitation historically helped to keep rodent populations under control, and only recently have the majority become owned as pets [16]. Cats were frequently taken on board ships to control pests, and so early European colonisation of North America by emigrants from the British Isles probably brought over many cats originating from the ports of embarkation, contributing to close similarity of the English and North American populations. Subsequent introductions may have had a much reduced impact due to the relative size of the established population. One example of possible back migration is the B6 variant which has attained its highest reported frequency in Ottawa, Canada, and may be represented in moderate frequency in nearby Maine, USA, from whence the Maine Coon breed originated. Importation of this breed has contributed the sole B6 variant so far reported in Europe [12].

## **Conclusion**

Our results provided no significant evidence of a difference in mitotypes between cats sampled in the vicinity of a crime scene and the wider English population, which allowed us to use a national estimate of the frequency to demonstrate the rarity of the evidential sequence. Combined with the finding of fibres matching the shower curtain on the suspect's clothing, the demonstration of likely two-way transfer played a key role in securing a conviction.

Although Sanger sequencing of cat hairs in this case proved useful in demonstrating a link with the disposal of the victim's body, the relative abundance of just a few control region mitotypes highlights that current approaches to the analysis of cat hairs will usually lack evidential weight: in our English population only 13% of cats had a sequence as uncommon as the one

described here. For cat hairs to be used more widely in forensic investigations a more discriminating test is clearly needed. Data from other hypervariable regions [12] and cytochrome b [18] show that discrimination power can be substantially enhanced by sequencing additional parts of the mtDNA molecule. We are exploring this through next-generation sequencing of whole mitochondrial genomes from a representative sample of cats from this study, and combining this with the development of short-amplicon multiplexes that will allow amplification of the entire mitochondrial genome from single shed hairs – work that will be described in future publications.

**Acknowledgments**

BO was supported by Hampshire Police, and JHW by the Leverhulme Trust, grant number F/00 212/AM, and by a University of Leicester Impact Development Fund award. We thank Beth Wictum and the Forensic Unit of the Veterinary Genetics Laboratory at the University of California Davis for sharing both evidential and primer sequences. We thank Deb Hopwood (formerly of LGC Forensics) for helpful suggestions and details of the UK casework findings.

**Conflicts of interest**

None

## References

- [1] Menotti-Raymond MA, David VA, O'Brien SJ 1997. Pet cat hair implicates murder suspect. *Nature*; 386(6627):774.
- [2] Lyons, L.A., Grahn, R.A., Kun, T.J., Netzel, L.R., Wictum, E.E. and Halverson, J.L., 2014. Acceptance of domestic cat mitochondrial DNA in a criminal proceeding. *Forensic Science International: Genetics*, 13, pp.61-67.
- [3] Tarditi CR, Grahn RA, Evans JJ, Kurushima JD, Lyons LA. 2011. Mitochondrial DNA sequencing of cat hair: an informative forensic tool. *Journal of forensic sciences* 56 Suppl 1: S36-46.
- [4] Menotti-Raymond, M., David, V.A., Stephens, J.C., Lyons, L.A., and O'Brien, S.J. 1997. Genetic individualization of domestic cats using feline STR loci for forensic applications. *J. Forensic Sci.* 42: 1039–1051.
- [5] Tully, G. and Wetton, J. H. 2014. Interpretation of Mitochondrial DNA Evidence. Wiley Encyclopedia of Forensic Science. DOI: 10.1002/9780470061589.
- [6] Grahn, Robert A., et al. 2015 "Feline mitochondrial DNA sampling for forensic analysis: When enough is enough!" *Forensic Science International: Genetics* 16: 52-57.
- [7] Murray JK, Browne WJ, Roberts MA, Whitmarsh A, Gruffydd-Jones TJ. 2010. Number and ownership profiles of cats and dogs in the UK. *Veterinary Record* 166: 163-168.
- [8] Muller K, Brugger C, Klein R, Miltner E, Reuther F, Wiegand P, . 2008. STR typing of hairs from domestic cats. *Forensic Science International: Genetics* Suppl 1: 607-609.
- [9] Grahn RA, Kurushima JD, Billings NC, Grahn JC, Halverson JL, Hammer E, Ho CK, Kun TJ, Levy JK, Lipinski MJ et al. 2011. Feline non-repetitive mitochondrial DNA control region database for forensic evidence. *Forensic Science International: Genetics* 5: 33-42.
- [10] Mirror PetLog data (<http://www.mirror.co.uk/news/ampp3d/cat-map-britain-wheres-most-5184092>)
- [11] Lopez JV, Cevario S, O'Brien SJ. 1996. Complete nucleotide sequences of the domestic cat (*Felis catus*) mitochondrial genome and a transposed mtDNA tandem repeat (Numt) in the nuclear genome. *Genomics* 33: 229-246.
- [12] Wesselink M, Bergwerff L, Hoogmoed D, Kloosterman AD, Kuiper I. 2015. Forensic utility of the feline mitochondrial control region - A Dutch perspective. *Forensic Science International: Genetics* 17: 25-32.
- [13] Arcieri M, Agostinelli G, Gray Z, Spadaro A, Lyons LA, Webb KM. 2016. Establishing a database of Canadian feline mitotypes for forensic use. *Forensic Science International: Genetics* 22: 169-174.
- [14] Pereira, L., Cunha, C. and Amorim, A., 2004. Predicting sampling saturation of mtDNA haplotypes: an application to an enlarged Portuguese database. *International journal of legal medicine*, 118(3), pp.132-136.
- [15] Lipinski MJ, Froenicke L, Baysac KC, Billings NC, Leutenegger CM, Levy AM, Longeri M, Niini T, Ozpinar H, Slater MR, Pedersen NC, Lyons LA 2008. The ascent of cat breeds: genetic evaluations of breeds and worldwide random-bred populations. *Genomics*. Jan;91(1):12-21.

- [16] Alhaddad, H., Khan, R., Grahn, R.A., Gandolfi, B., Mullikin, J.C., Cole, S.A., Gruffydd-Jones, T.J., Häggström, J., Lohi, H., Longeri, M. and Lyons, L.A., 2013. Extent of linkage disequilibrium in the domestic cat, *Felis silvestris catus*, and its breeds. *PLoS One*, 8(1), p.e53537.
- [17] Wetton, J.H., Higgs, J.E., Spriggs, A.C., Roney, C.A., Tsang, C.S. and Foster, A.P., 2003. Mitochondrial profiling of dog hairs. *Forensic Science International*, 133(3), pp.235-241.
- [18] Tamada, T., Kurose, N. and Masuda, R., 2005. Genetic diversity in domestic cats *Felis catus* of the Tsushima Islands, based on mitochondrial DNA cytochrome b and control region nucleotide sequences. *Zoological Science*, 22(6), pp.627-633.

## Figure legends

### **Figure 1: Geographical distribution of cat samples, crime-scene location, and human and feline population densities.**

a) Map of British Isles showing the distribution of English veterinary surgeries contributing samples. Inset below is an enlarged map of Southsea showing local vets: suspect's and victim's addresses, and the location of the victim's remains, all lie within the black circle. Map data © copyright Google. b) Human population density in the British Isles (contains Ordnance Survey data © Crown copyright and database right; sourced from [en.wikipedia.org/wiki/British\\_Isles#/media/File:British\\_Isles\\_population\\_density\\_2011\\_NUTS3.svg](https://en.wikipedia.org/wiki/British_Isles#/media/File:British_Isles_population_density_2011_NUTS3.svg)). c) UK distribution of microchipped cats based on data from Petlog ([www.petlog.org.uk](http://www.petlog.org.uk)) as a proxy for feline population density. Note that microchipping of cats is not obligatory and that the practice may vary by socio-economic status and geography.

### **Figure 2: Median-joining networks of mtDNA haplotypes showing frequencies of mitotypes in Portsmouth and the rest of England.**

a) Comparison of Portsmouth and the rest of England; b) Comparison of the Dutch [9] and the total English datasets. Circles represent haplotypes with area proportional to frequency, and lines between them represent SNP mutational steps between haplotypes. Populations are indicated by shading as shown in the keys. The networks each include two mitotypes taken from the feline mtDNA reference sequences, labelled LOPEZ [11] and SRS (Sylvester reference sequence) [3].

### **Figure 3: Comparison of mitotype frequencies in English and other populations.**

a) Distribution of mitotypes in different cat population groups. Populations are shown as pie-charts with sectors proportional to mitotype frequencies haplogroups as indicated in the key bottom right. Countries corresponding to particular population groups are shaded on the map. b) MDS of samples, treated as populations.



## Tables

**Table 1: Variable mtDNA nucleotide positions and mitotypes found among English cats.**

site	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
SRS <sup>a</sup> position	7	11	46	54	144	148	149	160	172	173	175	176	184	203	257	261	262	327	331	356	357	358	366	370	378	
Lopez <sup>b</sup> position	16820	16824	16859	16867	16957	16961	16962	16973	16985	16986	16988	16989	16997	7	58.1	62	63	130	134	159	160	161	169	173	181	
Lopez	T	A	T	C	C	A	A	A	G	T	C	T	G	A	D	T	T	T	T	T	C	T	C	G	G	
SRS	.	.	C	.	.	.	.	T	A	.	.	.	.	.	C	.	D	.	A	.	.	.	T	.	A	
mitotype																										number
A	.	.	C	.	.	.	.	T	A	C	.	.	.	.	C	.	D	C	A	.	.	.	T	A	A	37
B	.	G	C	.	.	.	.	T	A	.	.	.	.	.	C	D	A	.	A	C	.	.	T	.	A	24
B1	.	G	C	.	.	.	.	T	A	.	.	C	.	.	C	D	A	.	A	C	.	.	T	.	A	7
B-UK1	.	G	C	T	.	.	.	T	A	.	.	.	.	.	C	D	A	.	A	C	.	.	T	.	A	3
B-UK2	.	G	C	.	.	.	.	T	A	.	.	C	.	.	C	D	A	.	A	C	.	.	T	A	A	1
C	.	G	.	.	.	.	.	T	.	.	.	.	.	.	.	.	A	.	A	C	.	.	T	.	A	46
C-UK1	.	G	.	.	.	.	.	T	.	.	.	.	.	.	C	.	A	.	A	C	.	.	T	.	A	2
C-UK2	.	G	.	.	.	.	.	T	.	.	.	.	.	.	.	.	A	.	A	C	.	.	.	.	A	1
D	.	G	C	.	.	.	.	T	A	.	.	.	.	.	C	D	A	.	A	C	.	.	.	.	A	7
E	C	.	.	.	T	.	.	T	A	.	T	.	A	.	C	.	A	.	A	.	.	.	.	.	A	3
F	.	.	C	.	.	.	.	T	A	.	.	.	.	.	C	D	A	.	A	.	.	.	.	.	A	10
G	.	.	C	.	.	.	.	T	A	.	.	.	.	.	C	D	A	.	A	C	.	.	T	.	A	1
H	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	A	4
I	C	.	.	.	T	.	.	T	A	C	.	.	.	G	C	.	A	C	A	.	.	.	T	A	A	4
J	.	G	C	.	.	.	.	T	A	.	.	.	.	.	.	.	A	.	A	C	.	.	.	.	A	1
OL1	.	.	.	.	.	G	G	T	.	.	.	.	.	.	C	.	D	.	A	C	T	C	.	.	A	1

<sup>a</sup> Sylvester reference sequence [3]; <sup>b</sup> Lopez reference sequence [11].

**Table 2: Frequencies of mitotypes in different cat populations.**

Mitotype	England	United States	Europe	Eastern Mediterranean	Iran/Iraq	Indian Ocean trade route	East Asia	Fancy breeds	Netherlands
<b>A</b>	37	123	10	73	10	24	21	102	80
<b>B</b>	34	143	9	7	0	25	5	77	12
<b>C</b>	47	98	7	10	2	18	2	67	3
<b>D</b>	7	24	1	73	36	27	0	33	3
<b>E</b>	3	10	1	9	14	3	22	9	0
<b>F</b>	10	15	0	17	1	13	0	8	8
<b>G</b>	2	11	0	13	0	2	0	11	0
<b>H</b>	4	14	4	4	0	0	0	1	1
<b>I</b>	4	6	1	1	0	0	0	11	1
<b>J</b>	1	0	0	8	6	0	0	1	0
<b>K</b>	0	18	0	0	0	0	0	2	0
<b>L</b>	0	4	0	9	0	0	0	2	0
<b>OL1</b>	1	0	0	0	0	0	0	7	0
<b>OL2</b>	0	3	0	0	0	0	0	0	0
<b>OL3</b>	0	0	0	0	0	2	0	0	0
<b>Unique mitotype</b>	2	24	25	9	3	1	6	36	5
<b>Total</b>	152	493	58	233	72	115	56	367	113

English data are from the current study, Netherlands data from [12], and remaining data from [9].

**Table 3: Strength of evidence based on mitotype frequencies.**

a) Upper-bound frequency estimates	Number of observations (x)	Number of samples (n)	Frequency of B-UK1 (pobs)	$p \pm \sqrt{p(1-p)/n}$ upper bound	$p = (x+1)/(n+2)$	$p = (x+2)/(n+2)$
Southsea	1	23	0.043	0.086	0.080	0.120
Portsmouth	1	32	0.031	0.062	0.059	0.088
Rest of England	2	120	0.017	0.028	0.025	0.033
England combined	3	152	0.020	0.031	0.026	0.032
b) Estimates of number of B-UK1 carriers in the population (based on combined English upper-bound frequencies)	Estimated number of cats	Expected number of carriers (observed freq.)	Number of carriers (upper-bound estimate)			
Southsea *	1,322	26	41	34	43	
Portsmouth *	14,671	290	455	381	476	
England	8,620,000	170,132	267,383	223,896	279,870	

\* Assumes one cat/14 people in Portsmouth area (PetLog data <http://www.mirror.co.uk/news/ampp3d/cat-map-britain-wheres-most-5184092>)

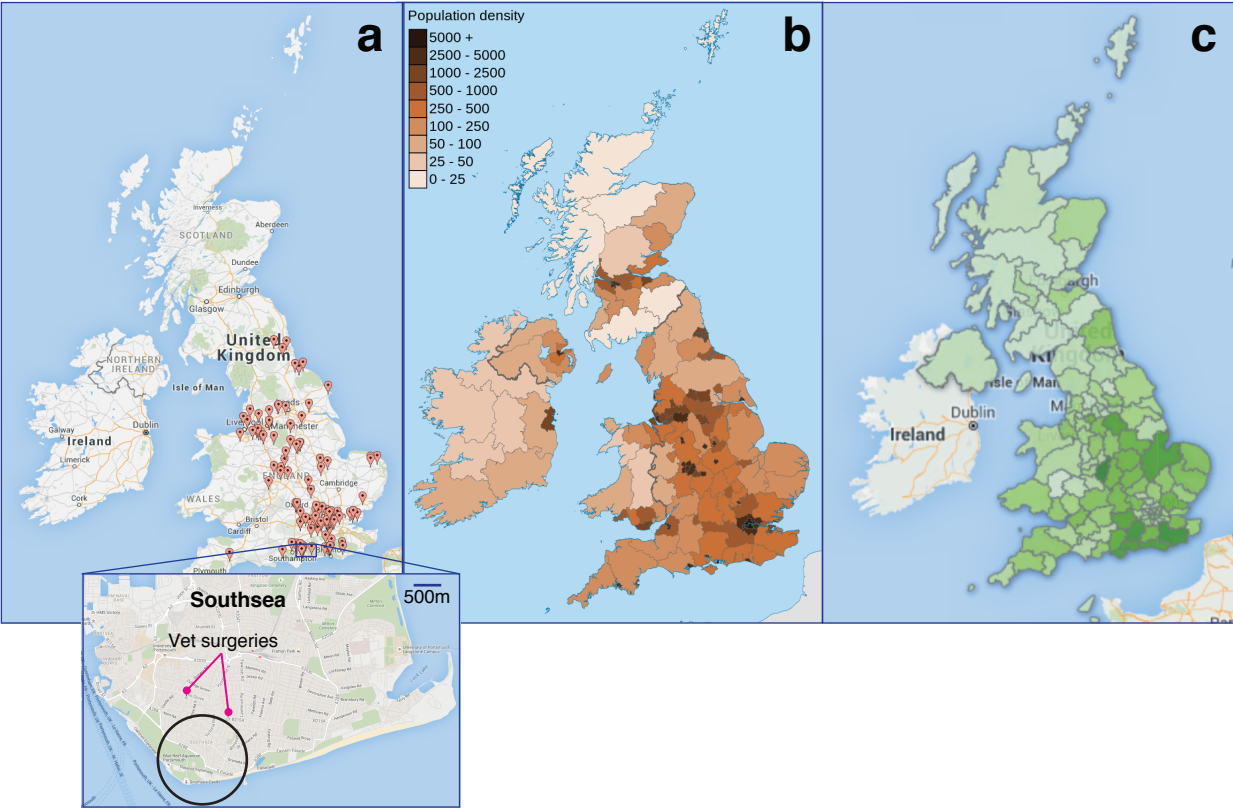
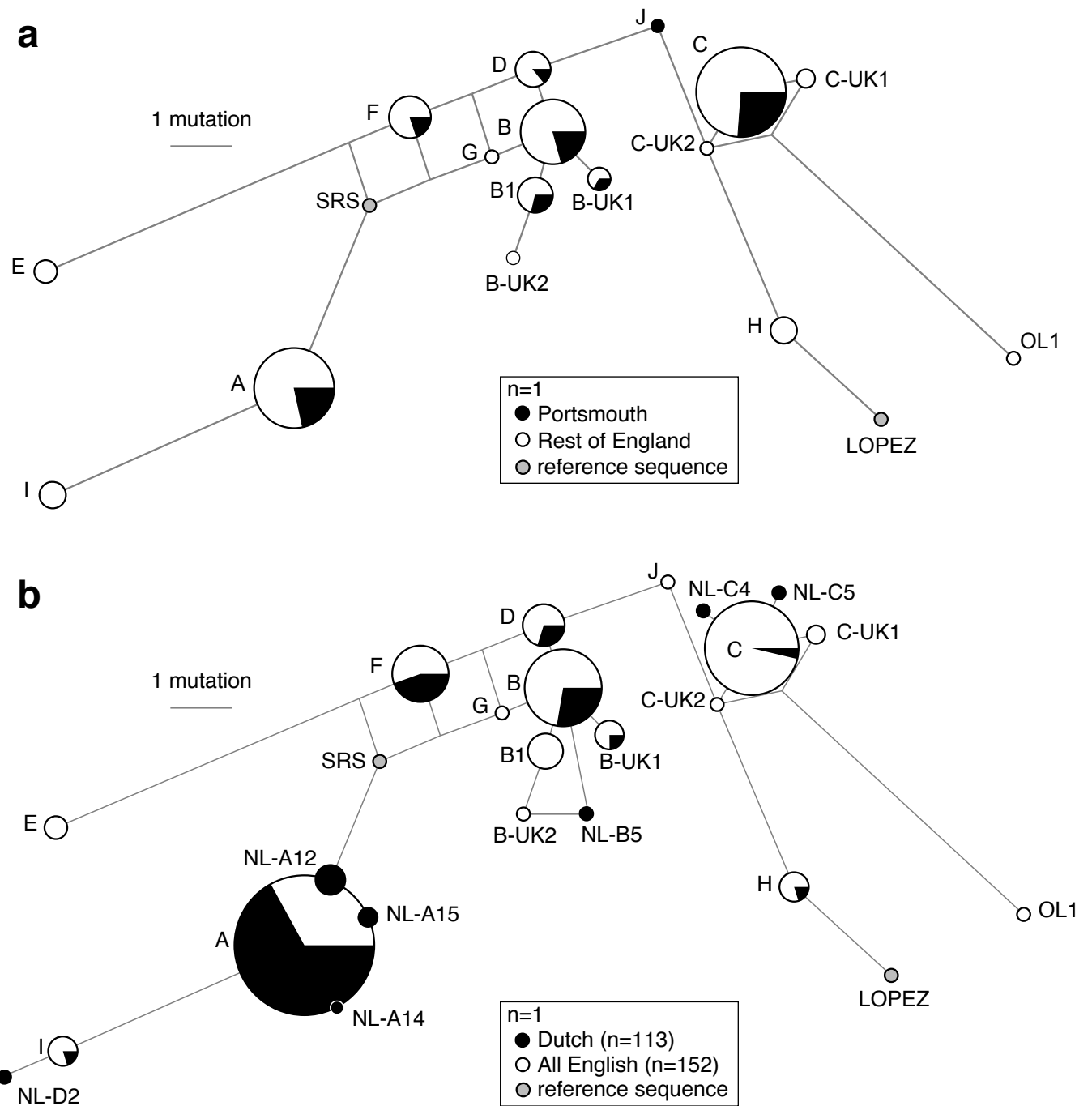
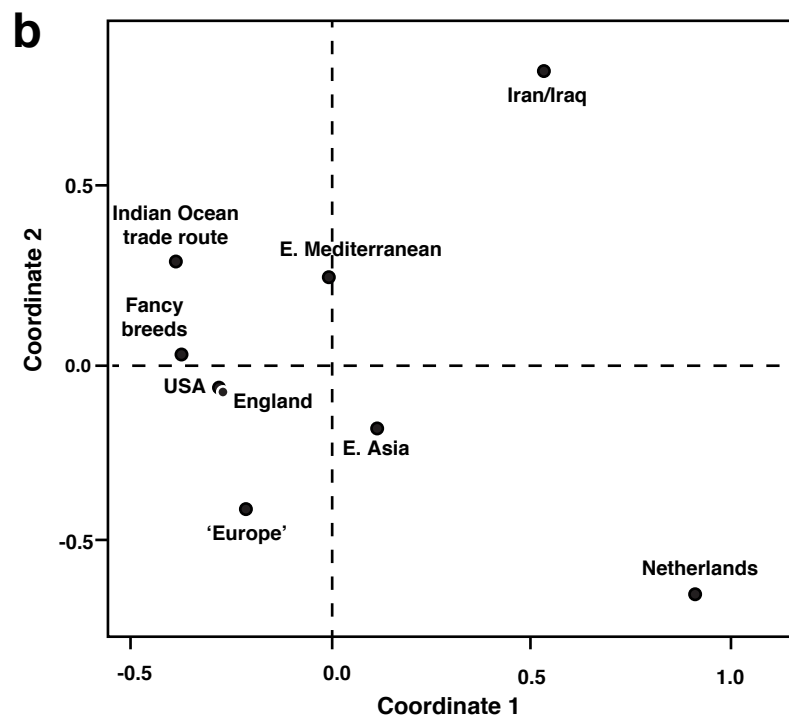
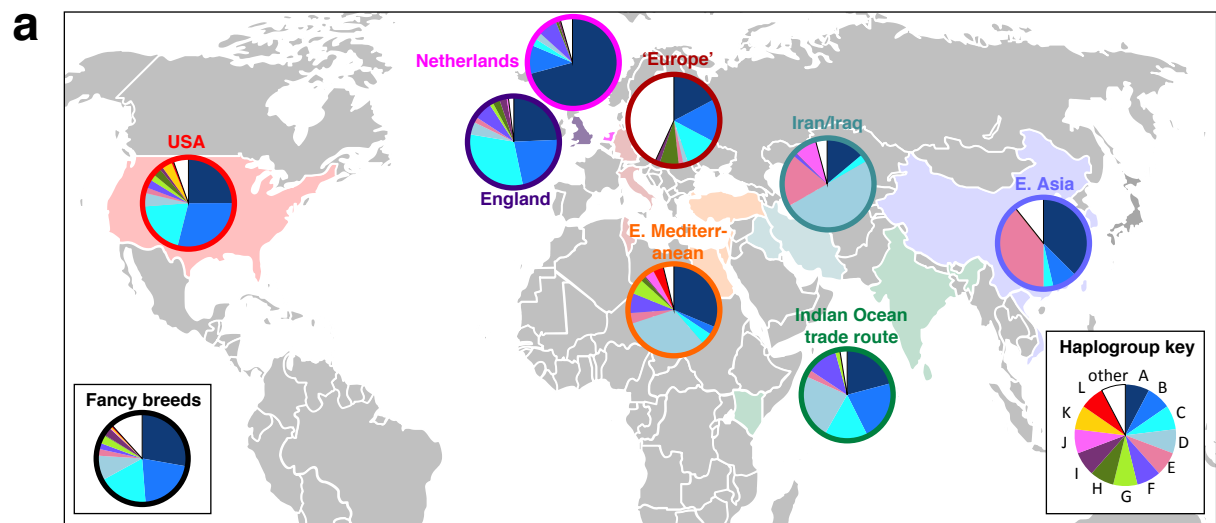


Figure 1



**Figure 2**



**Figure 3**