

RESEARCH ARTICLE

Patterns of allele frequency differences among domestic cat breeds assessed by a 63K SNP array

Hasan Alhaddad^{1*}, Mona Abdi¹, Leslie A. Lyons²

1 Department of Biological Sciences, Kuwait University, Safat, Kuwait, **2** Department of Veterinary Medicine and Surgery, College of Veterinary Medicine, University of Missouri - Columbia, Columbia, Missouri, United States of America

* hassan.alhaddad@ku.edu.kw, hhalhaddad@gmail.com



Abstract

Cats are ubiquitous companion animals that have been keenly associated with humans for thousands of years and only recently have been intentionally bred for aesthetically appealing coat looks and body forms. The intense selection on single gene phenotypes and the various breeding histories of cat breeds have left different marks on the genomes. Using a previously published 63K Feline SNP array dataset of twenty-six cat breeds, this study utilized a genetic differentiation-based method (*d_i*) to empirically identify candidate regions under selection. Defined as three or more overlapping (500Kb) windows of high levels of population differentiation, we identified a total of 205 candidate regions under selection across cat breeds with an average of 6 candidate regions per breed and an average size of 1.5 Mb per candidate region. Using the combined size of candidate regions of each breed, we conservatively estimate that a minimum of ~ 0.1–0.7% of the autosomal genome is potentially under selection in cats. As positive controls and tests of our methodology, we explored the candidate regions of known breed-defining genes (e.g., *FGF5* for longhaired breeds) and we were able to detect the genes within candidate regions, each in its corresponding breed. For breed specific exploration of candidate regions under selection, eleven representative candidate regions were found to encompass potential candidate genes for several phenotypes such as brachycephaly of Persian (*DLX6*, *DLX5*, *DLX2*), curled ears of American Curl (*MCRIP2*, *PBX1*), and body-form of Siamese and Oriental (*ADGRD1*), which encourages further molecular investigations. The current assessment of the candidate regions under selection is empiric and detailed analyses are needed to rigorously disentangle effects of demography and population structure from artificial selection.

OPEN ACCESS

Citation: Alhaddad H, Abdi M, Lyons LA (2021) Patterns of allele frequency differences among domestic cat breeds assessed by a 63K SNP array. PLoS ONE 16(2): e0247092. <https://doi.org/10.1371/journal.pone.0247092>

Editor: Gregory S. Barsh, HudsonAlpha Institute for Biotechnology, UNITED STATES

Received: December 17, 2020

Accepted: February 1, 2021

Published: February 25, 2021

Copyright: © 2021 Alhaddad et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The cat, *Felis catus*, was likely domesticated around 10,000 years ago but only recently has imposed artificial selection created pedigreed cats that are generically referred to as cat breeds [1–3]. Since the first cat shows held in London's Crystal Palace in 1871 and later in New York's

Madison Square Garden in 1881, many new breeds have been developed [4–6]. Currently, 40–71 breeds are recognized by different cat fancy organizations (Cat Fanciers' Association (CFA)—42 [7], The International Cat Association (TICA)—71 [8], Governing Council of the Cat Fancy (GCCF)—40 [9], and Federation Internationale Feline (FIFE)—48 [10] breeds). The difference in the number of recognized breeds is related to the different criteria used to define a breed. Unlike other domesticated animals, which were selected for physical, behavioral, production, or functional traits and resulted in extreme breed variation [11–15], cat breeds were mostly selected for discrete, single gene traits that were aesthetically pleasing to breeders [16].

In general, cat selected phenotypes can be classified into two groups; one related to the coat (color, pattern, length and texture) and a second related to body morphology (face, ear, body size, legs and tail) [17]. Most of the external variation, especially those that define breeds, can be seen in the coat and less so in body morphology [16, 18]. Cat aesthetic traits differ in (1) type; coat, body variations, or combinations, (2) origin; recently propagated *de novo* mutation or selected from natural populations, (3) history; an old trait or recently discovered, (4) specificity; trait present in a single breed or shared amongst multiple breeds. These differences between the selected traits are likely to have different impacts on the genomes of various breeds.

Cat coat color traits belong to two groups; (1) patterned, which is displayed as a combination of more than one color in the form of spotted, striped, swirls and dorsal-ventral pigmentation, or (2) pigmentation, where colors range from white to heavily pigmented color. Cat coat color phenotypes are products of genetic variation within *KIT* (white and white spotting) [19–21], *TYR* (points pattern, mocha, albino) [22–24], *ASIP* (black, agouti) [25], *MC1R* (amber and russet) [26, 27], *TYRP1* (brown, cinnamon and chocolate) [28], and *MLPH* (dilute coloration) [29]. Similarly, coat length and texture arose from polymorphisms within *FGF5* (long-hair) [30, 31], *KRT71* (curly and naked) [32, 33], and *LPAR6* and *LIPH* (curly) [34, 35]. On the other hand, body morphology variations that were selected in various breeds include the face (Burmese—*ALX1*) [36, 37], ears (Scottish Fold folded ears—*TRPV4*) [38], size (dwarfism—*UGDH*) [39], and tail (Japanese Bobtail tail length and kink—*HES7*, Manx tail length and tailless—*T* gene) [40–42].

The origin of cat breed defining traits generally followed one of three paths [43]. One path depended on selecting individual cats from the standing variation within a feral population. Breeds developed by this route are referred to as natural breeds (i.e. Persian, Turkish Angora, Maine Coon and Siberian) [44]. A second relied on combining traits of two breeds or two cat species and develop a cross-hybrid. The Ocicat's defining look was developed using other established breeds (Abyssinian, American Shorthair and Siamese). The Bengal breed is a hybrid between domestic cats and the Asian leopard cat, *Prionailurus bengalensis* [6]. A third way was to identify a *de novo* mutation (i.e. an individual with unique trait) and carefully perform selective matings to fix and propagate the newly discovered phenotype (e.g., Selkirk Rex [45]).

Recent and rapid increase in cat breeds through strong selection directed for desired single traits and crossing of unrelated pre-existing breeds, created a bulk of interrelated and yet phenotypically divergent breeds. However, the majority of breeds were established by selecting from feral populations (new phenotype or pre-existing one) and can be traced to four regional populations: (1) Western derived (e.g. Persian, Main Coon, and Norwegian Forest Cat), (2) Eastern derived (e.g. Birman, Burmese, Bombay, Siamese and Oriental), (3) Mediterranean (Turkish Van and Turkish Angora), and (4) Arabian Sea (Sokoke) [16, 43, 44].

The development of high-throughput genotype arrays has permitted the possibility of identifying chromosomal regions and genes targeted by artificial selection in various domestic animals species (e.g. cats [46], dogs [47, 48], cattle, [11, 49, 50], and horse [15]). Correspondingly,

the current study aimed to investigate the signatures of selection of twenty-six recognized cat breeds using 63K SNP array data and provide an overall representation of the candidate genomic regions across all breeds, which we refer to as the “landscape of selection”. The specific objectives included: (1) identifying candidate regions under selection for each breed, (2) evaluating the biological significance of the detected candidate regions using known phenotype-causing variants, (3) exploring the gene content within a representative number of candidate regions for likely phenotype association, and (4) inspecting candidate regions that are present in several closely related breeds.

Results

Information on selected dataset

Seven hundred thirty-seven cats representing 26 recognized cat breeds were included in the study (S1 Table). Seven breeds (Persian, Scottish Fold, Siberian, Bengal, Ragdoll, Birman and Burmese) were represented by 50 samples. To conduct an intraspecific comparison within a single breed, the samples of these seven over represented breeds were randomly divided into two groups (e.g., Bengal 1 and Bengal 2). The breed sample sizes ranged from 10 to 26 cats and twenty of the breeds were represented by 20 or more cats.

As a preparation of the workable SNP dataset, autosomal markers ($n = 58,888$) were extracted from the genotype dataset by removing the X-chromosome SNPs ($n = 2,700$) and the unmapped SNPs ($n = 684$). Approximately 99.7% ($n = 58,768$) of the autosomal markers passed SNP genotyping call rate of 90% (each SNP genotyped in at least 90% of the samples), thus, only 120 SNPs were excluded across all samples. The average percentage of monomorphic SNPs within each breed was 31.7% ranging from (19.1%–54.8%) (S1 Table). Approximately 12% SNPs ($n = 6,955$) had a MAF between (0 and 0.05) across all populations and were excluded. Overall, 50709 autosomal SNPs that were accurately mapped to the latest cat genome assembly (felCat 9.0 [51]) were included in downstream analyses.

To understand the genetic relationships among the breeds under study, PCA was performed and the first two components PC1 and PC2 explained ~ 40% of the total genetic variation (Fig 1). A broad overview showed that the first component, PC1, which accounts for 28.9% of the total variation, structured the breeds according to their origin ancestry as a continuum from Western to Eastern (Fig 1a). Additionally, Western breeds and what we refer to as “Middle” breeds clustered more closely than to the Eastern breeds, which forms a distant distinct group. The population structure Western breeds can be summarized by: (1) the Persian family breeds (Persian, British Shorthair, Scottish Fold, and Selkirk Rex) formed a distinct group, (2) Maine Coon breed formed a single distinct group, and (3) individuals from the Munchkin and Siberian were less closely structured and extended to the “Middle” breeds (Fig 1b). The “Middle” breeds were less distinctly structured with the exception of Bengal, which formed a separate group (Fig 1c). Two general groups were detected in the Eastern breeds cluster where Birman separated from the rest (Fig 1d). Individuals of the Bombay breed did not form a unique group and some extended to the “Middle” and “Western” clusters (Fig 1a). The second component, PC2, that explains 9.5% of the total variation enabled separation of Bengal from other “Middle” breeds (Fig 1c) and Birman from Eastern breeds (Fig 1d).

Selection statistical test (*di*)

As a calculation of likely signature of selection, SNP-specific *di* values were calculated for each breed. averaged for SNPs within individual overlapping 500Kb windows (50% overlap between adjacent windows). The 99th percentile threshold (top 92 windows) were reported as candidate windows under selection (S1 Fig, S2 Table).

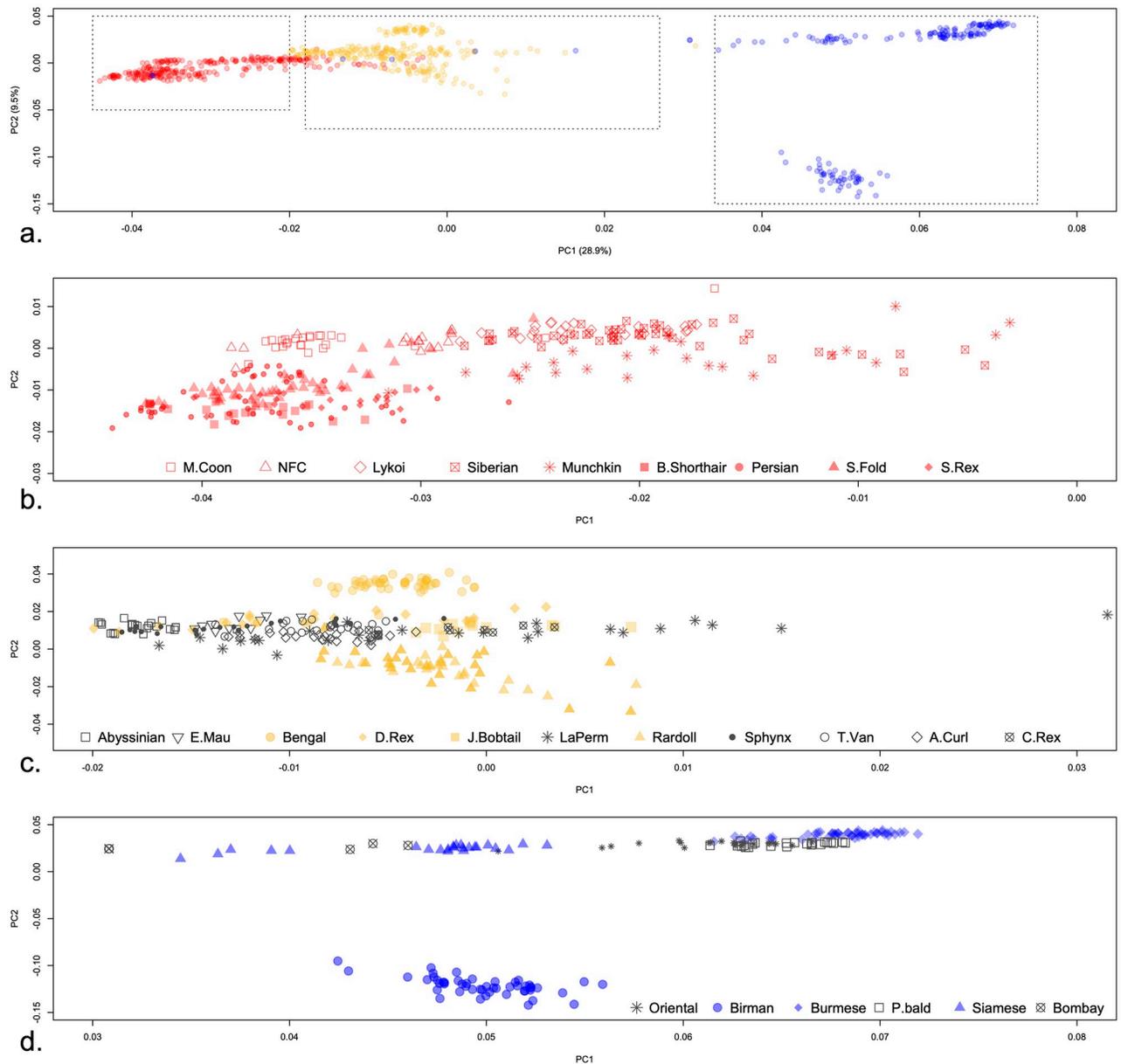


Fig 1. Population structure of cat breeds via principle component analysis (PCA). The first two principle components (PC1 and PC2) explain majority of the genetic variation among cat breeds. (a) A broad overview of the population structure of all breeds. Each individual cat is represented by a circle that is colored according to its ancestral group (Western-red, “Middle”-yellow, Eastern-blue) and dashed boxes are used to designate 9 Western breeds (left box), 11 intermediate breeds (middle box) and 6 Eastern breeds (left box). The blue dots extending to the “Middle” and Western groups belong to Bombay breed. Note that Bombay breed is formed by breeding of Burmese and American Shorthair. (b-d) Individual representations of Western, “Middle”, and Eastern breeds, respectively. Note the separation of the interspecies hybrid breed, Bengal, from other “Middle” breeds in (b) and separation of the phenotypically distinct Birman from other Eastern breeds in (c). Six Bombay individuals were outside of the ancestral grouping of Eastern breeds were removed in (d) for ease of presentation.

<https://doi.org/10.1371/journal.pone.0247092.g001>

Collectively across breeds, individual candidate windows under selection (top 92) from each breed were distributed across the autosomal chromosomes (S1 Fig, S2 Table). This posed challenges such as: (1) despite being on different locations, the number of candidate windows (92) is identical for each breed, which prevents objective comparisons between breeds, (2) the likelihood of false positive windows due to SNP ascertainment bias associated with the design

of the SNP array, (3) the uneven distribution of SNPs across windows, and (4) the lack of a systematic method that prioritizes certain candidate windows for gene content investigation. To overcome these challenges, we sought to identify candidate regions under selection rather than to explore individual scattered windows. We define a candidate region under selection as three or more overlapping candidate windows, which combined would be ≥ 1 Mb in size (S3 Table).

Using the aforementioned approach, nearly 50% of the candidate windows in each breed were single scattered windows or only two overlapping windows (S2 Fig, S2 Table). The categorization of candidate windows is only for prioritizing which candidate regions to be investigated for gene contents. These candidate windows were not totally dismissed from subsequent analyses but later used when discussing specific breeds (e.g., Persian) or inspecting specific chromosomal locations (e.g., *FGF5* on Chr. B1). Abyssinian breed showed the lowest percentage of nonoverlapping candidate windows (50%) whereas Bombay had over 90% of its candidate windows as single or two windows (S2 Fig).

Candidate regions of selection

Combining and overlaying the identified candidate regions under selection for all breeds provides an overall look at the genomic landscape of selection in the pedigreed domestic cats (Fig 2). Using the 205 candidate regions (S3 Table) of 26 breeds (not accounting for redundancy), the following regarding the genomic landscape of selection in cats was observed: (1) candidate regions were found variably in all chromosomes (S3a Fig), (2) the largest number of candidate region were observed in Chrs. B4, A1, and B1, respectively whereas the least number were in Chrs. E1, D4, and F1 (S3a Fig), (3) the number of candidate regions were in correlation with the size of the chromosome with the exception of Chr. B4 and E3 (S3b Fig), (4) the majority of candidate regions were composed of three overlapping windows and measured 1 Mb in size (S3c and S3d Fig), (5) the average size of candidate regions was 1.5 Mb, and the largest candidate region (10 Mb) was found on Chr. B4 (S1 Fig).

Whilst candidate regions expectedly varied in number and size between chromosomes, variation was also observed between breeds (S1 Table, S2 Fig). The number of candidate regions of selection averaged to ~ 6 regions per breed where the least was found in Bombay (two regions) and the most was in Birman (11 regions) (S3 Table). Similarly, the average combined size of candidate regions per breed was 9.3 Mb and ranged from 2.5 Mb in Bombay to a 16.25 Mb in Ragdoll (S1 Table, S4 Fig). When examining the number and the combined sizes of candidate regions per breed in relation to the ancestral groups (Western, “Middle”, Eastern), we observed that Eastern breeds (excluding Bombay) had larger number of candidate regions per breed (averaged ~ 9 regions) and covering larger regions (averaged ~ 12 Mb) compared to “Middle” (~ 6 regions and 10.4 Mb), and Western breeds (~ 5 regions and 7 Mb).

We adopted the exclusion of single or two windows regions as a strategy to prioritize the importance of candidate regions and to provide conservative estimates of candidate regions under selection for each breed. However, we find it more useful to include all candidate windows when comparing two populations of the same breed (e.g., Bengal 1 and Bengal 2) or between different breeds especially since partial shared windows are often observed (S1 Fig).

Intra-specific breed comparison was conducted on seven breeds representing the three ancestral groups (Western, “Middle”, and Eastern). The shared candidate windows within a breed ranged from 38% for Scottish Fold to over 84% for Burmese (S4 Table). The Western breeds (Persian, Scottish Fold, and Siberian) exhibited lower number of shared candidate windows among the subpopulations (38–53%) while Bengal and Ragdoll as representatives of “Middle” breeds shared (62–64%). Eastern breeds (Birman and Burmese) showed the highest number of shared windows (77–84%). The unique and unshared windows between the two

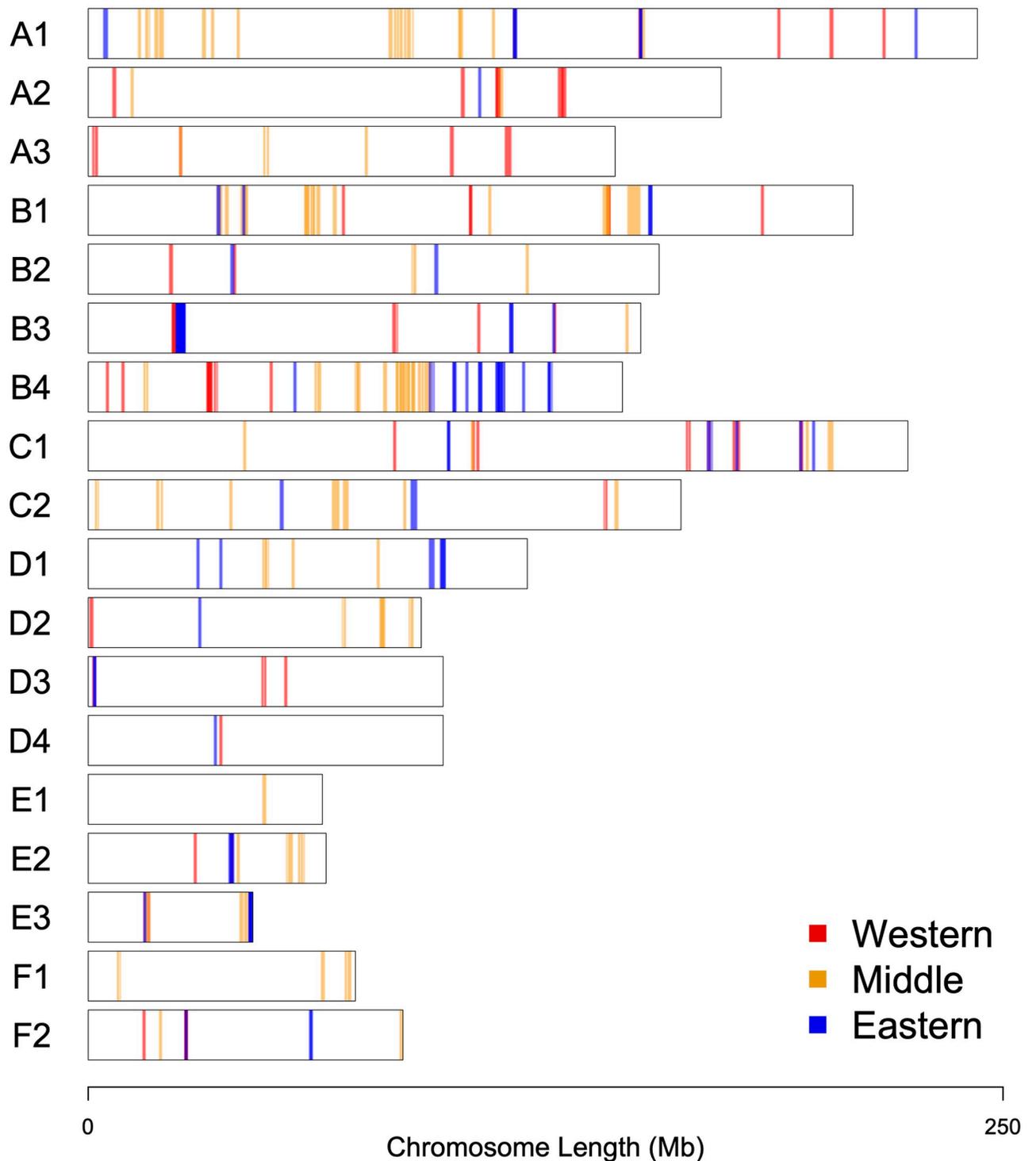


Fig 2. The genomic landscape of selection in domestic cat genome. Chromosomal representation of 205 candidate regions under selection across all cat breeds. Colored vertical bars represent candidate regions under selection and the intensity of color reflects the number of breeds having the particular candidate region. Red, yellow, and blue colored bars correspond to Western (9), “Middle” (11), and Eastern (6) breeds, respectively.

<https://doi.org/10.1371/journal.pone.0247092.g002>

populations within an individual breed were mostly of single scattered windows or two overlapping windows (S4 Table). This independently supports our strategy of defining candidate regions under selection by three or more overlapping windows.

Many individual windows were shared across several breeds (examples in S1 Fig Chrs. A1, B1, and B2). However, more significantly was the presence of shared candidate regions. One example of such regions was located on Chr.B3 and shared among Bombay, Burmese, Peterbald, Oriental, and Siamese and less uniformly with Scottish Fold and Siberian. A second region is located on Chr. E3 and is shared across seven breeds from the three ancestral groups (S1 Fig).

Candidate regions as positive controls

Genetic variants that are associated with phenotypes under selection in cats serve as useful positive controls for our conducted analyses and a display of the validity subsequent reported findings. Our five positive controls were: (1) Cornish Rex and multiple candidate regions surrounding its curly coat causative gene, *LPAR6* (S1 Fig-Chr.A1), (2) Abyssinian and a candidate region where *ASIP* resides, which explains the lack the non-agouti of its coat (S1 Fig-Chr.A3), (3) Persian, Norwegian Forest Cat, Siberian, and Turkish Van with candidate regions of different sizes around the longhair gene, *FGF5* (S1 Fig-Chr.B1), (4) Devon Rex's candidate region for the curly coat associated gene, *KRT71* (S1 Fig-Chr.B4), and (5) Siamese a *TYR* gene candidate region, which is responsible for the distinct dark coat coloration at the extremities (S1 Fig-Chr.D1).

Dominantly inherited genetic variants, previously identified for breed-defining phenotypes, were also observed. The candidate regions encompassing or adjacent to such variants were mostly represented by single or two windows. Examples of these regions included: (1) windows surrounding *UGDH* gene, which is related to dwarfism in Munchkins (S1 Fig-Chr.B1), (2) a single window adjacent to *KRT71*, which is responsible for the curly coat of Selkirk Rex (S1 Fig-Chr.B4), and (3) a window covering the tail-length causative gene, *HES*, in Japanese Bobtail (S1 Fig-Chr.E1).

Single breed candidate regions

We explored the candidate regions under selection for each breed and examined genes within for relationships to the breed defining phenotypes (S3 Table). However, we limited our discussion to a representative number of candidate regions selected based on their (1) sizes, (2) presence of phenotype related candidate genes, and when applicable (3) occurrence in replica populations (i.e., Persian1 and 2). Below, we provide detailed examples for three breeds while others are reported in Table 1.

Genome-wide analysis of the signatures of selection for the American Curl breed showed distinctly several overlapping windows with high levels of genetic population differentiation on Chr.E3 and F1 (Fig 3a). The size of the first candidate region was 3.75Mb and contained 123 genes among which *MCRIP2* representing a possible candidate gene for the backward curled ears. A second smaller candidate region was detected with *PBX1* as another candidate gene (Table 1). Similarly, Maine Coon's genome-wide survey showed two apparent regions on Chrs.A1 and A3 (Fig 3b). The candidate region on Chr.A1 was a 1Mb region and contained four genes (*GABRG2*, *GABRA1*, *GABRA6*, *GABRB2*) whereas the region on Chr.A3 was 2Mb with two genes, *FAM98A* and *RASGRP3*. Lastly, with a focus on the unique coloration of the Turkish Van, which is characterized by an overall white coat with different coloration of parts of the head and tail, we identified two candidate regions (Fig 3c). The regions were located on Chrs.A2 and E3 with the latter harboring *CLCN7* as candidate genes, (Table 1).

Table 1. Representative candidate regions under selection in cat breeds with likely phenotype related candidate genes.

Breed	Chr.	Start (Mb)	End (Mb)	Region size (Mb)	No. genes	Candidate gene(s)	Phenotype
Perisan ¹	A2	101.75	103	1.25	8	<i>DLX6, DLX5</i>	Brachycephaly
	C1	163.25	164.75	1.5	6	<i>DLX2</i>	
Maine Coon	A1	188.25	189.25	1	4	<i>GABRG2, GABRA1, GABRA6, GABRB2</i>	Eating behavior/size
Bengal ¹	D2	79.5	81.25	1.75	13	<i>FGFR2</i>	Coat type
Turkish Van	E3	43.75	44.75	1	48	<i>CLCN7</i> ²	Van color
American Curl	E3	41.25	44.75	3.75	123	<i>MCRIP2</i>	Curled ears
	F1	63.5	45	1.25	6	<i>PBX1</i> ³	
La Perm	B1	147.25	151	3.75	26	<i>EPGN, EREG</i>	Curly coat
Ragdoll ¹	A1	82	89	7	40	<i>LAMP1</i>	Multiple
Birman ¹	B4	11.25	114	2.75	8	<i>KITLG</i>	Gloves
Oriental	D3	1.25	2.25	1	5	<i>ADGRD1</i>	Body weight
Siamese	D3	1.25	2.25	1	5	<i>ADGRD1</i>	Body weight

¹ Candidate regions were found completely or partially shared among two populations (n = 25 cats each) of the same breed.

² Due to the low density of SNPs on Chr.E3, flanking regions to the candidate region were explored. Candidate gene (*CLCN7*) is located upstream of the region at ~ 43.3Mb.

³ Due to low density of SNPs on Chr.F1, flanking regions were explored. Candidate gene (*PBX1*) is located upstream of the region at ~ 63Mb.

<https://doi.org/10.1371/journal.pone.0247092.t001>

Common selected region across several breeds

A shared candidate region of relatively low SNP density was found in Turkish Van, American Curl, Japanese Bobtail, Peterbald, Oriental, and Siamese (S1 Fig, S3 Table). This shared candidate region among diverse breeds was located on Chr. E3 and contained 36 genes. A second shared candidate region was identified specifically in the breeds of slender body shape; Peterbald, Oriental, and Siamese (See Fig 4 for body shape). The Chr. D3 region was 1Mb in size and contained five genes among *ADGRD1* as potential candidate for the unique body form (Table 1).

Discussion

Recent and rapid selection in the domestic cat has resulted in the formation of 40–71 cat breeds that are recognized today [7–10]. Selection in the cat focused on single or few aesthetic external traits, which are considered breed defining [16]. The distinct curly coat of Selkirk Rex [52] and the tail morphology of Japanese Bobtail [53] are examples of cat breed development practices. A different scenario is the Somali breed, which is only a longhair variant of Abyssinian cats, sharing nearly identical breed standards [7]. This study (1) explored the fingerprints of the cat selection history and breeding practices using genome-wide SNPs genotyped in 26 breeds, (2) highlighted candidate regions under selection, and (3) reported likely phenotype-related genes.

Despite the intense selection imposed during the formation of the cat breeds, only ~ 2.5–16.25 Mb appeared under selection, which constituted 0.1–0.7% of the autosomal genome. Using similar analytical approaches, the number of candidate regions identified for the cat (205 in 26 breeds) is smaller than the ones identified for the dog (275 in 10 breeds) [48], the horse (695 in breeds) [15], and the cow (583 in 5 breeds) [54]. The difference in number of candidate regions under selection between cat breeds and other domesticated animals may largely be due to limited selection on a single or few genes in cats whereas in other animals the selection is on quantitative trait loci (QTL).

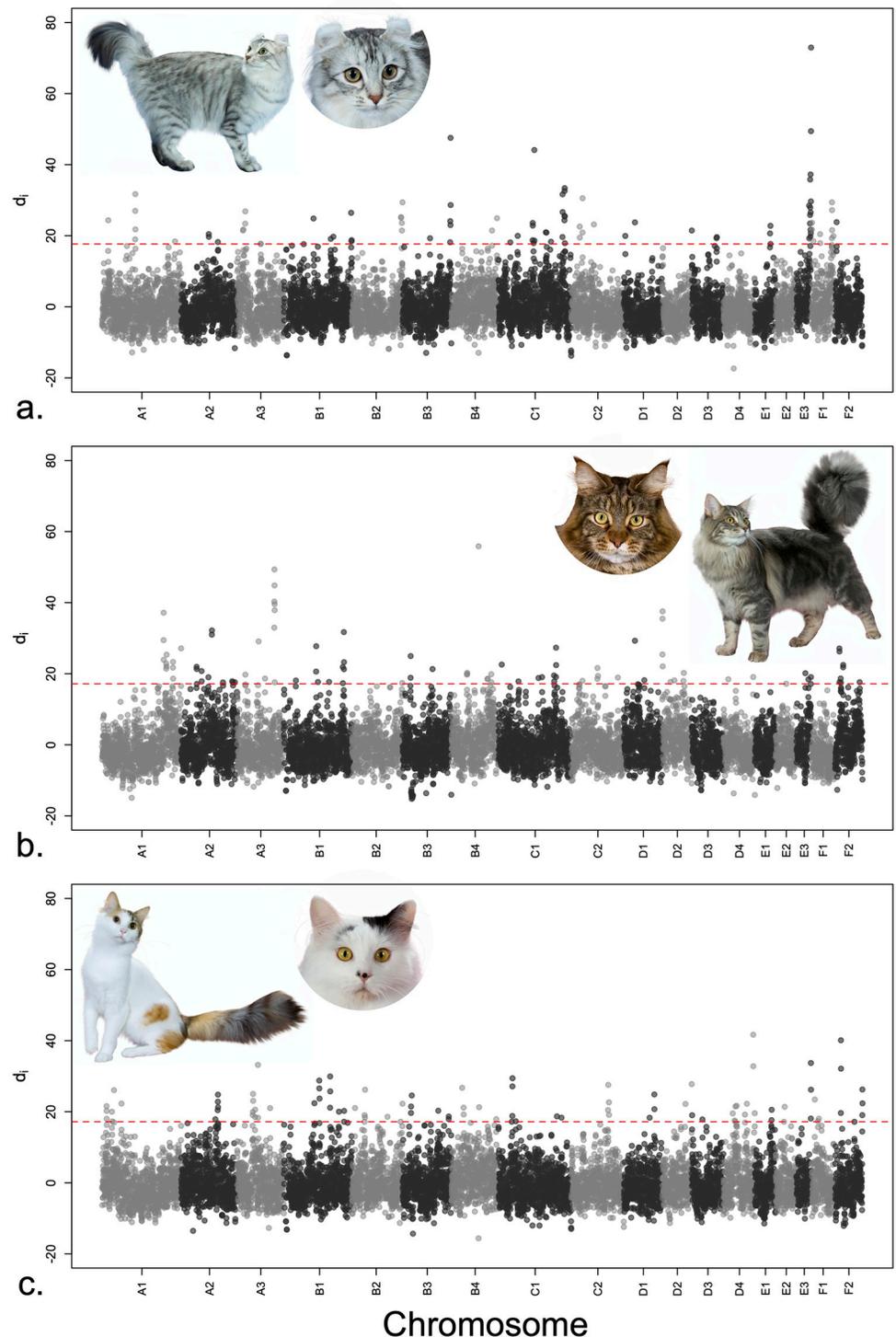


Fig 3. Manhattan plots of genome-wide (d_i) measures of overlapping 500Kb windows for three cat breeds. Each dot represents a window of 500 Kb (total 9108 windows) and the horizontal dashed red line indicates the 99th percentile threshold, which distinguishes the top 92 windows. (a) Genome-wide survey of selection for the American Curl, which is a breed defined by backward curled ears. A number of overlapping windows clearly show a signal on Chr. E3. (b) Genome-wide overview of selection for Maine Coon cats, which is a breed recognized by an overall large size and facial features such large oval eyes and a square chin. Visible signals of overlapping windows are shown on Chrs. A1 and A3. (c) Signal of selection for the Turkish Van breed, which is acknowledged for a coat color pattern named after the breed, “Van”. This pattern appears in an overall white coat with distinct coloration in parts of the head and tail. Few overlapping windows were detected on Chrs. A2 and E3. Cat photographs were provided with permission by Larry Johnson.

<https://doi.org/10.1371/journal.pone.0247092.g003>



Fig 4. Morphological variation in body-form between six Eastern breeds. Birman symbolizes one extreme end of the physical spectrum of Eastern breeds by displaying stocky body, broad and round head, rounded muzzle, and round eyes whereas Siamese, Oriental, and Peterbald display the other extreme end through a tubular body, wedged head, wedged muzzle, and almond-shaped eyes. Burmese, Bombay exhibit intermediate body-shapes and facial features between the two extremes. Cat photographs were provided with permission by Larry Johnson.

<https://doi.org/10.1371/journal.pone.0247092.g004>

Cat breeds and signatures of selection

While top 1% candidate windows ($n = 92$) under selection for each breed were reported (S2 Table), we prioritized the candidate windows into candidate regions based on the presence of

multiple overlapping windows in the same genomic location (S3 Table). This ensured reporting of conservative estimates of candidate regions under selection in each breed, allowed objective comparison between breeds, and avoided false positives due to ascertainment bias. On average, cat breeds had 2–12 regions under selection positioned on 2–8 chromosomes and with combined size of 2.5–16.25 Mb. The number and size of candidate regions varied among breeds based on: (1) ancestry (e.g., Western), (2) population size (i.e., popularity), (3) age of the breed-defining trait(s) (e.g., longhair), (4) mode of inheritance of the trait under selection, and (5) breeding practices (e.g., outcrossing). An example of the differences observed based on breed ancestry was the relative low number and small size of candidate regions under selection in Western breeds contrasted by the high number of large size candidate regions in Eastern breeds (S1 Table). This is consistent with previously reported differences in estimated population parameters such as linkage disequilibrium, genetic diversity, and levels of inbreeding [55, 56].

Persian, including Exotics as a short-haired variant, is the most popular and largest pure-bred cat population. The effect of population size of the Persian breed on the candidate region under selection was evident in the few small sized candidate regions (max = 1.5 Mb). This effect was especially apparent when comparing the two replica populations (Persian 1 and 2) where 50% of the candidate windows were shared (S4 Table). The variability in the detect candidate regions under selection mirrors the variability observed in its signature brachycephaly phenotype and the range of diseases [57]. Nonetheless, two shared candidate regions among the two Persian populations contained promising candidate genes for brachycephaly (Table 1). Both candidate regions contained distal-less homeobox genes (Chr.A2: *DLX6* and *DLX5*—Chr.C1: *DLX2*), which are involved in the craniofacial development [58, 59]. We suspect that several candidate genomic regions and several genes are responsible for the wide range of brachycephaly phenotypes in Persian cats.

A comparison between the Turkish Van and the Bengal breeds provides an illustration to the age the breed and its selected phenotypes on the numbers and sizes of candidate regions under selection. The Turkish Van is an old breed and belongs to a Mediterranean ancestral cluster [16, 44] with signature coat color pattern named after the breed, “Van”. This coat color appears as an overall white color with different colors randomly covering parts of the head and tail (Fig 3c). On the other hand, the Bengal breed is a comparatively younger breed that resulted from an interspecies hybridization [6] with a marking phenotype of patterned, soft, and glittered coat. Reminiscent of the old age of the Turkish Van breed and its phenotype, several small sized candidate regions were identified, most notably on Chr. E3 (Fig 3c) with *CLCN7* as a candidate gene for the hypopigmented appearance of the “Van” coat color [60]. By contrast, much fewer and considerably larger candidate regions were localized for the Bengal breed and two signifying a recent selection on Chrs.B4 (~ 5.5 Mb) and D2 (1.75 Mb). A focus on the candidate region on Chr. D2, which had a manageable number of genes, suggested *FGFR2* as a candidate gene for the softness of the Bengal’s coat [61].

The effect of the mode of inheritance of the breed-defining phenotype also has effects on the number and size of candidate regions. As an illustration, we refer to the regions of the Scottish Fold breed, which is defined by the dominantly-inherited folded ears [38] and compare it to American Curl, which characterized by the recessively-inherited backward curled ears (Fig 3a) [62]. Using two separate populations of the Scottish Fold breed (SFOLD1 and 2), we observed: (1) an averaged maximum candidate region size of 2 Mb, (2) significantly low number of shared windows between the two populations (S4 Table), and (3) lack of any candidate windows near the phenotype causing gene, *TRPV1* on Chr.D3. Conversely, the America Curl breed exhibited larger candidate regions two of which are indicative of selection of a recessive trait (Fig 3a). The largest candidate region (3.75 Mb) was located on Chr.E1 and among the

123 genes in the region, *MCRIP2* represents a likely candidate for curled ears. *MCRIP2* is a paralog of *FGFRL1*, which has high localized expression in cartilaginous tissue [63, 64]. A second candidate region also possessed a candidate gene for the curled ears, *PBX1*, and is implicated with ear malformations [65].

Lastly, breeding practices during the formation of a cat breed has apparent manifestations on the candidate regions under selection. We use the Abyssinian breed as an example of high inbreeding and the Bombay as an outcrossed breed. While Abyssinian is known to be an inbred breed [16, 44, 55], Bombay is documented to be a hybrid breed between the Eastern Burmese and the Western American Shorthair (Fig 1, S1 Table). The signs of inbreeding of Abyssinian contributed to eight candidate regions with a combined size of 15.5 Mb whereas the outcrossing of Bombay allowed detecting as two candidate regions with a sum of 2.5 Mb.

Other candidate regions and candidate genes

We examined the robustness of our analysis using known phenotype-related genes and inspected their colocalization in candidate regions under selection (e.g. Cornish Rex [46], Devon Rex [32], Abyssinian [25], Siamese [22], and in longhaired breeds [66, 67]). However, beyond the aforementioned known breed-defining variants, we suggested a number of candidate genes that resided in identified regions of different breeds and might be responsible for phenotypes with still unknown causative genes (Table 1).

The Maine Coon breed is known for its large size (Fig 3b) [68] and its candidate region on Chr.A1 might be related to the large body size. The four genes within the region, *GABRG2*, *GABRA1*, *GABRA6*, and *GABRB2*, are known to alter the tastebuds and enhance food craving [69, 70]. On the other extreme of body-form, we have detected a shared candidate region across Siamese, Oriental, and Peterbald on Chr.D3. The three Eastern breeds exhibit thin-tubular body (Fig 4). The shared candidate region highlighted *ADGRD1* as a potential candidate gene for its association with metabolism and body-weight [71, 72].

Although several genes responsible for the curly coat in cat breeds have been identified [32, 33, 35, 46], the genetic variant responsible for the curled coat of LaPerm is still unknown. We have identified 3.75 Mb candidate region in LaPerm with likely involvement of *EPM3* and *EGER* to the unique curly coat of the breed [73]. Examining the Birman via its two populations and the highly shared candidate windows showed several windows on Chr.B1 where *KIT* is located. The white coat that is located on Birman cats' feet and referred to as "gloves", has been linked to *KIT* (Fig 4) [18]. However, a large candidate region was observed on Chr.B4 with *KITLG* as a potential second player in the formation of the "gloves" phenotype.

The single largest candidate region across all breeds was observed in Ragdoll on Chr.A1 (7 Mb) and shared among its two populations (RAG1 and 2). Even though the region contained forty genes, the *Lysosomal Associated Membrane protein 1 (LAMP1)* represented a candidate gene that may explain multiple aspects related to Ragdoll cats. *LAMP1* is associated with *Lysosomal Trafficking Regulator (LYST)*, which causes the multi-system Chediak-Higashi Syndrome. The syndrome manifests as dilution of skin, hair, and iris colors, bleeding diathesis, and recurrent infections (reviewed in [74]). Although the *LYST* and its severe consequences have been identified and described in cats (cited in [75]), we suspect that the large candidate region on Chr.A1 and *LAMP1* more moderately contributes to the phenotypic characteristics of Ragdoll cats (e.g. light coat and eye colors) and explains its mucopolysaccharidoses [76].

We have empirically identified candidate regions under selection in twenty-six cat breeds and due to the nature of the data (i.e., Array SNPs ascertained from different breeds), it is not possible to rigorously account for and distinguish between the effects of demography and population structure from artificial selection. Using whole-genome sequence data can foreseeably

provide more definitive confirmations of the candidate regions presented. Furthermore, our suggested candidate genes also await an in-depth molecular investigation to reach satisfying conclusions.

Materials and methods

Ethics statement

This study does not require an ethics statement.

SNP genotype dataset acquisition

A previously published SNP dataset using the Illumina Feline-infinium 63K SNP array was used in the current study [55]. The dataset contains 34 domestic cat breeds ($n = 1,570$). A subset of the genotype dataset was selected by excluding cat breeds with sample size less than 10 individuals, and samples with genotyping rate less than 90%. Sample genotyping rate was determined by using the command (`--mind`) in plink v1.7 [77] (S1 and S2 Data).

Within each cat breed, only unrelated samples were retained to perform further analyses. The genetic relatedness between individuals within each breed was determined using the statistic (PI_HAT) that measures the proportion of Identity by descent (IBD) of alleles for pairs of individuals and was calculated by the command (`--genome`) in plink v1.7 [77]. A (PI_HAT) of 0.25 or lower was used to select samples that were unrelated at least to the grandparent level. The final list of samples included in the study and their corresponding breed assignment are shown in S5 Table.

SNP genotyping and quality control

The original datasets consists of ~63K SNP markers [55] and the following markers were excluded from downstream analyses: (1) X-chromosome SNPs, (2) unmapped SNPs, (3) SNPs with genotyping rate less than 90%, and (4) SNPs unmapped to felCat 9.0 [51]. The final list of SNPs included in the study and their updated positions to felCat 9.0 are shown in S6 Table. SNP genotyping rate was determined using (`--geno`) while percentage of monomorphic SNPs was determined with (`--freq`) both implemented in plink v1.7 [77].

For each population independently, the following summary statistics were calculated using PLINK v1.7 [77]: (1) the function (`--freq`) was used to calculate the mean and standard deviation of minor allele frequency (MAF), (2) the mean and standard deviation of observed and expected heterozygosity were obtained using the function (`--hardy`), and (3) the mean individual inbreeding coefficient (F) in each breed was determined with the command (`--het`).

Population structure

Population structure and relationships between different cat breeds ($n = 26$) were analyzed using Principal Component Analysis (PCA) that is based on the variance-standardized relationship matrix implemented in PLINK v1.9 [78]. With the default setting, (`--pca`) function was used to generate the first 20 principle components. The first two components of the PCA (PC1, PC2) were used to illustrate the overall population structure of the breeds.

F_{st} and d_i calculations

F_{st} was calculated following Weir and Cockerham method [79], which provides accurate estimates of F_{st} especially when sample sizes are uneven across breeds or are small [80] and using vcftools v 0.1.13 [81]. Original files (.ped and .map) were first converted into Variant Call

Format (vcf) file using (--recode-vcf) command in plink v1.9 [78]. Using vcf files and (--weir-fst-pop) command of vcftools, F_{st} was calculated pairwise between all breeds and for each SNP.

A simple summary statistic (d_i) that measure locus specific divergence in allele frequencies was previously developed and implemented [15, 46, 48]. A script written in R was used to calculate d_i values according to the equation for each breed independently against all breeds. At each SNP, the d_i represents the sum of the values calculated above.

Candidate regions under selection

Based on SNP distribution and ensure sufficient SNP density in our scans, each chromosome was divided into 500Kb overlapping windows (e.g., window1: 0-500Kb, window2: 250-750Kb). The d_i estimates were averaged for all SNPs within each window ($n = 9108$ windows). The d_i averaged values of each window within each chromosome were visualized via an R Manhattan plot script. The 99th percentile was used as a threshold for window prioritization and the top 92 windows were considered candidate windows under selection. To avoid false positives and to allow comparison across breed beyond the 99th percentile threshold, the candidate windows were further categorized into candidate regions if three or more windows overlapped. The resulting candidate regions were visualized across breeds in chromosomal plots based on the latest cat genome assembly (felCat 9.0 [51]).

Supporting information

S1 Fig. Chromosomal representation of candidate windows under selection across cat breeds. Each autosomal chromosome is plotted separately. Chromosomes are positioned on the x-axis with gray thin lines representing SNP position. Each breed is represented on the y-axis and traced by a dotted horizontal line for ease of reference. Western, “Middle”, and Eastern, are colored red, yellow, and blue, respectively. Individual colored-bars corresponds to a 500 Kb candidate window (92 for each breed). Overlapping windows are displayed as darker colored bars and combined are considered candidate region under selection whereas individual scattered are of lightly colored and dismissed from gene content survey. Vertical dashed lines are placed according to the genomic positions (felCat9) of known genetic disease/phenotype implicated genes in cats. Boxed candidate regions were chosen as representative candidate regions in individual breeds or in multiple breeds and further investigated for candidate genes. (PDF)

S2 Fig. Proportions of candidate windows under selection categorized into four classes based on number of overlapping windows.

(TIFF)

S3 Fig. Summary of candidate regions under selection for all cat breeds combined.

(a) Overview of the conservative number of candidate regions per chromosome. (b) Number of candidate regions in relation to the chromosome size. Chr. B4 and Chr. E3 deviates from linear relationship between chromosome size and number of candidate regions. (c) Number of overlapping windows and (d) size (Mb) per candidate region. The majority of candidate regions are composed of three overlapping widows and 1Mb size.

(TIFF)

S4 Fig. A comparison of the combined size (Mb) of candidate regions under selection across cat breeds.

(TIFF)

S1 Table. Summary statistics of investigated cat breeds.
(XLSX)

S2 Table. A List of candidate 500Kb windows under selection of 26 cat breeds.
(XLSX)

S3 Table. A List of candidate regions under selection including lists of genes within each candidate region for 26 cat breeds.
(XLSX)

S4 Table. An intra-specific comparison of candidate windows between two populations within a breed.
(XLSX)

S5 Table. A list of the samples used in the study.
(XLSX)

S6 Table. A list of the SNPs used in the study with updated positions to felCat 9.0.
(XLSX)

S1 Data. SNP genotype data used in the study (ped file).
(ZIP)

S2 Data. SNP positions of the genotype data used in the study (map file).
(ZIP)

Acknowledgments

We are grateful to the four anonymous reviewers who provided constructive criticisms, detailed comments, and valuable suggestions. We are also thankful to the two editors-in-chief, Gregory Barsh and Gregory Copenhaver, for their comments and guidance towards the improvement of the manuscript and enhancing the clarity of its findings. We lastly thank Larry Johnson for supplying high quality photographs of cat breeds and Tasneem Maraqa for editing the photos.

Author Contributions

Conceptualization: Hasan Alhaddad, Leslie A. Lyons.

Data curation: Hasan Alhaddad, Mona Abdi.

Formal analysis: Hasan Alhaddad, Mona Abdi.

Investigation: Hasan Alhaddad, Leslie A. Lyons.

Methodology: Mona Abdi.

Supervision: Hasan Alhaddad, Leslie A. Lyons.

Validation: Hasan Alhaddad, Mona Abdi.

Visualization: Hasan Alhaddad, Mona Abdi.

Writing – original draft: Hasan Alhaddad, Mona Abdi, Leslie A. Lyons.

Writing – review & editing: Hasan Alhaddad.

References

1. Vigne JD, Guilaine J, Debue K, Haye L, Gerard P. Early taming of the cat in Cyprus. *Science*. 2004; 304(5668):259. <https://doi.org/10.1126/science.1095335> PMID: 15073370.
2. Driscoll CA, Menotti-Raymond M, Roca AL, Hupe K, Johnson WE, Geffen E, et al. The Near Eastern origin of cat domestication. *Science*. 2007; 317(5837):519–23. <https://doi.org/10.1126/science.1139518> PMID: 17600185.
3. Gentry A, Clutton-Brock J, Groves CP. The naming of wild animal species and their domestic derivatives. *Journal of Archaeological Science*. 2004; 31(5):645–51. <https://doi.org/10.1016/j.jas.2003.10.006>
4. Crystal Palace—Summer concert today Cat Show on July 13. Penny Illustrated Paper, Amusement:. 1871;510, July 08:11. Epub 510.
5. The First Cat Show in America. *New York Times*. 1881 March 06.
6. Morris D. *Cat breeds of the world*. New York: Penguin Books; 1999.
7. CFA. Cat Fanciers' Association www.cfa.org/2019 [cited 2017 May 01]. www.cfa.org/.
8. TICA. The International Cat Association www.tica.org/2019 [cited 2017 May 01]. www.tica.org/.
9. GCCF. The Governing Council of the Cat Fancy. <http://www.gccfcats.org> 2019 [cited 2017 May 01]. <http://www.gccfcats.org>.
10. FIFe. Federation Internationale Feline fifeweb.org/2019 [cited 2017 May 01]. fifeweb.org/.
11. Ramey HR, Decker JE, McKay SD, Rolf MM, Schnabel RD, Taylor JF. Detection of selective sweeps in cattle using genome-wide SNP data. *BMC Genomics*. 2013; 14:382. <https://doi.org/10.1186/1471-2164-14-382> PMID: 23758707.
12. de Simoni Gouveia JJ, da Silva MV, Paiva SR, de Oliveira SM. Identification of selection signatures in livestock species. *Genet Mol Biol*. 2014; 37(2):330–42. <https://doi.org/10.1590/s1415-47572014000300004> PMID: 25071397.
13. Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, et al. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature*. 2013; 495(7441):360–4. <https://doi.org/10.1038/nature11837> PMID: 23354050.
14. Schlamp F, van der Made J, Stambler R, Chesebrough L, Boyko AR, Messer PW. Evaluating the performance of selection scans to detect selective sweeps in domestic dogs. *Mol Ecol*. 2016; 25(1):342–56. <https://doi.org/10.1111/mec.13485> PMID: 26589239.
15. Petersen JL, Mickelson JR, Rendahl AK, Valberg SJ, Andersson LS, Axelsson J, et al. Genome-wide analysis reveals selection for important traits in domestic horse breeds. *PLoS genetics*. 2013; 9(1): e1003211. <https://doi.org/10.1371/journal.pgen.1003211> PMID: 23349635.
16. Kurushima JD, Lipinski MJ, Gandolfi B, Froenicke L, Grahn JC, Grahn RA, et al. Variation of cats under domestication: genetic assignment of domestic cats to breeds and worldwide random-bred populations. *Animal genetics*. 2013; 44(3):311–24. <https://doi.org/10.1111/age.12008> PMID: 23171373.
17. Lyons LA. DNA mutations of the cat: The good, the bad and the ugly. *Journal of Feline Medicine and Surgery*. 2015; 17(3):203–19. <https://doi.org/10.1177/1098612X15571878> PMID: 25701860.
18. Montague MJ, Li G, Gandolfi B, Khan R, Aken BL, Searle SM, et al. Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111(48):17230–5. <https://doi.org/10.1073/pnas.1410083111> PMID: 25385592.
19. Cieslak M, Reissmann M, Hofreiter M, Ludwig A. Colours of domestication. *Biol Rev Camb Philos Soc*. 2011; 86(4):885–99. <https://doi.org/10.1111/j.1469-185X.2011.00177.x> PMID: 21443614.
20. Cooper MP, Fretwell N, Bailey SJ, Lyons LA. White spotting in the domestic cat (*Felis catus*) maps near KIT on feline chromosome B1. *Animal genetics*. 2006; 37(2):163–5. <https://doi.org/10.1111/j.1365-2052.2005.01389.x> PMID: 16573531.
21. David VA, Menotti-Raymond M, Wallace AC, Roelke M, Kehler J, Leighty R, et al. Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. *G3 (Bethesda)*. 2014; 4(10):1881–91. <https://doi.org/10.1534/g3.114.013425> PMID: 25085922.
22. Lyons LA, Imes DL, Rah HC, Grahn RA. Tyrosinase mutations associated with Siamese and Burmese patterns in the domestic cat (*Felis catus*). *Anim Genet*. 2005; 36(2):119–26. <https://doi.org/10.1111/j.1365-2052.2005.01253.x> PMID: 15771720.
23. Imes DL, Geary LA, Grahn RA, Lyons LA. Albinism in the domestic cat (*Felis catus*) is associated with a tyrosinase (TYR) mutation. *Animal genetics*. 2006; 37(2):175–8. Epub 2006/04/01. <https://doi.org/10.1111/j.1365-2052.2005.01409.x> PMID: 16573534.
24. Yu Y, Grahn RA, Lyons LA. Mocha tyrosinase variant: a new flavour of cat coat coloration. *Animal genetics*. 2019; 50(2):182–6. Epub 2019/02/04. <https://doi.org/10.1111/age.12765> PMID: 30716167.

25. Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, O'Brien SJ. Molecular genetics and evolution of melanism in the cat family. *Curr Biol*. 2003; 13(5):448–53. [https://doi.org/10.1016/s0960-9822\(03\)00128-3](https://doi.org/10.1016/s0960-9822(03)00128-3) PMID: 12620197.
26. Gustafson NA, Gandolfi B, Lyons LA. Not another type of potato: MC1R and the russet coloration of Burmese cats. *Animal genetics*. 2017; 48(1):116–20. <https://doi.org/10.1111/age.12505> PMID: 27671997
27. Peterschmitt M, Grain F, Arnaud B, Deleage G, Lambert V. Mutation in the melanocortin 1 receptor is associated with amber colour in the Norwegian Forest Cat. *Animal genetics*. 2009; 40(4):547–52. <https://doi.org/10.1111/j.1365-2052.2009.01864.x> PMID: 19422360.
28. Lyons LA, Foe IT, Rah HC, Grahn RA. Chocolate coated cats: TYRP1 mutations for brown color in domestic cats. *Mammalian Genome*. 2005; 16(5):356–66. <https://doi.org/10.1007/s00335-004-2455-4> PMID: 16104383
29. Ishida Y, David VA, Eizirik E, Schaffer AA, Neelam BA, Roelke ME, et al. A homozygous single-base deletion in MLPH causes the dilute coat color phenotype in the domestic cat. *Genomics*. 2006; 88(6):698–705. <https://doi.org/10.1016/j.ygeno.2006.06.006> PMID: 16860533.
30. Kehler JS, David VA, Schaffer AA, Bajema K, Eizirik E, Ryugo DK, et al. Four independent mutations in the feline Fibroblast Growth Factor 5 gene determine the long-haired phenotype in domestic cats. *Journal of Heredity*. 2007; 98(6):555–66. <https://doi.org/10.1093/jhered/esm072> PMID: 17767004
31. Drögemüller C, Rüfenacht S, Wichert B, Leeb T. Mutations within the FGF5 gene are associated with hair length in cats. *Animal Genetics*. 2007; 38(3):218–21. <https://doi.org/10.1111/j.1365-2052.2007.01590.x> PMID: 17433015
32. Gandolfi B, Outerbridge CA, Beresford LG, Myers JA, Pimentel M, Alhaddad H, et al. The naked truth: Sphynx and Devon Rex cat breed mutations in KRT71. *Mamm Genome*. 2010; 21(9–10):509–15. <https://doi.org/10.1007/s00335-010-9290-6> PMID: 20953787.
33. Gandolfi B, Alhaddad H, Joslin SEK, Khan R, Filler S, Brem G, et al. A splice variant in KRT71 is associated with curly coat phenotype of Selkirk Rex cats. *Scientific Reports*. 2013; 3. <https://doi.org/10.1038/srep02000> PMID: 23770706
34. Gandolfi B, Alhaddad H, Affolter VK, Brockman J, Haggstrom J, Joslin SEK, et al. To the Root of the Curl: A Signature of a Recent Selective Sweep Identifies a Mutation That Defines the Cornish Rex Cat Breed. *Plos One*. 2013; 8(6). <https://doi.org/10.1371/journal.pone.0067105> PMID: 23826204
35. Manakhov AD, Andreeva TV, Rogaev EI. The curly coat phenotype of the Ural Rex feline breed is associated with a mutation in the lipase H gene. *Animal Genetics*. 2020;n/a(n/a). <https://doi.org/10.1111/age.12958> PMID: 32463158
36. Bertolini F, Gandolfi B, Kim ES, Haase B, Lyons LA, Rothschild MF. Evidence of selection signatures that shape the Persian cat breed. *Mamm Genome*. 2016. <https://doi.org/10.1007/s00335-016-9623-1> PMID: 26956354.
37. Lyons LA, Erdman CA, Grahn RA, Hamilton MJ, Carter MJ, Helps CR, et al. Aristaless-Like Homeobox protein 1 (ALX1) variant associated with craniofacial structure and frontonasal dysplasia in Burmese cats. *Developmental Biology*. 2016; 409(2):451–8. <https://doi.org/10.1016/j.ydbio.2015.11.015> PMID: 26610632
38. Gandolfi B, Alamri S, Darby WG, Adhikari B, Lattimer JC, Malik R, et al. A dominant TRPV4 variant underlies osteochondrodysplasia in Scottish fold cats. *Osteoarthritis Cartilage*. 2016; 24(8):1441–50. <https://doi.org/10.1016/j.joca.2016.03.019> PMID: 27063440.
39. Lyons LA, Fox DB, Chesney KL, Britt LG, Buckley RM, Coates JR, et al. Localization of a feline autosomal dominant dwarfism locus: a novel model of chondrodysplasia. *bioRxiv*. 2019:687210. <https://doi.org/10.1101/687210>
40. Xu X, Sun X, Hu X-S, Zhuang Y, Liu Y-C, Meng H, et al. Whole Genome Sequencing Identifies a Missense Mutation in HES7 Associated with Short Tails in Asian Domestic Cats. *Scientific Reports*. 2016; 6:31583. <https://www.nature.com/articles/srep31583#supplementary-information>. PMID: 27560986
41. Lyons LA, Creighton EK, Alhaddad H, Beale HC, Grahn RA, Rah H, et al. Whole genome sequencing in cats, identifies new models for blindness in AIPL1 and somite segmentation in HES7. *BMC genomics*. 2016; 17:265-. <https://doi.org/10.1186/s12864-016-2595-4> PMID: 27030474.
42. Buckingham KJ, McMillin MJ, Brassil MM, Shively KM, Magnaye KM, Cortes A, et al. Multiple mutant T alleles cause haploinsufficiency of Brachyury and short tails in Manx cats. *Mammalian genome: official journal of the International Mammalian Genome Society*. 2013; 24(9–10):400–8. Epub 2013/08/15. <https://doi.org/10.1007/s00335-013-9471-1> PMID: 23949773.
43. Gandolfi B, Alhaddad H. Investigation of inherited diseases in cats: genetic and genomic strategies over three decades. *J Feline Med Surg*. 2015; 17(5):405–15. <https://doi.org/10.1177/1098612X15581133> PMID: 25896240.

44. Lipinski MJ, Froenicke L, Baysac KC, Billings NC, Leutenegger CM, Levy AM, et al. The ascent of cat breeds: genetic evaluations of breeds and worldwide random-bred populations. *Genomics*. 2008; 91(1):12–21. <https://doi.org/10.1016/j.ygeno.2007.10.009> PMID: 18060738.
45. Filler S, Alhaddad H, Gandolfi B, Kurushima JD, Cortes A, Veit C, et al. Selkirk Rex: morphological and genetic characterization of a new cat breed. *J Hered*. 2012; 103(5):727–33. <https://doi.org/10.1093/jhered/ess039> PMID: 22837475.
46. Gandolfi B, Alhaddad H, Affolter VK, Brockman J, Haggstrom J, Joslin SE, et al. To the Root of the Curl: A Signature of a Recent Selective Sweep Identifies a Mutation That Defines the Cornish Rex Cat Breed. *PLoS one*. 2013; 8(6):e67105. <https://doi.org/10.1371/journal.pone.0067105> PMID: 23826204.
47. Vaysse A, Ratnakumar A, Derrien T, Axelsson E, Rosengren Pielberg G, Sigurdsson S, et al. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS genetics*. 2011; 7(10):e1002316. <https://doi.org/10.1371/journal.pgen.1002316> PMID: 22022279.
48. Akey JM, Ruhe AL, Akey DT, Wong AK, Connelly CF, Madeoy J, et al. Tracking footprints of artificial selection in the dog genome. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(3):1160–5. <https://doi.org/10.1073/pnas.0909918107> PMID: 20080661.
49. Barendse W, Harrison BE, Bunch RJ, Thomas MB, Turner LB. Genome wide signatures of positive selection: the comparison of independent samples and the identification of regions associated to traits. *BMC Genomics*. 2009; 10:178. <https://doi.org/10.1186/1471-2164-10-178> PMID: 19393047.
50. Hayes BJ, Chamberlain AJ, Maceachern S, Savin K, McPartlan H, MacLeod I, et al. A genome map of divergent artificial selection between *Bos taurus* dairy cattle and *Bos taurus* beef cattle. *Animal genetics*. 2009; 40(2):176–84. <https://doi.org/10.1111/j.1365-2052.2008.01815.x> PMID: 19067671.
51. Buckley RM, Davis BW, Brashear WA, Farias FHG, Kuroki K, Graves T, et al. A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. *bioRxiv*. 2020:2020.01.06.896258. <https://doi.org/10.1371/journal.pgen.1008926> PMID: 33090996
52. Filler S, Alhaddad H, Gandolfi B, Kurushima JD, Cortes A, Veit C, et al. Selkirk Rex: Morphological and Genetic Characterization of a New Cat Breed. *Journal of Heredity*. 2012; 103(5):727–33. <https://doi.org/10.1093/jhered/ess039> PMID: 22837475
53. Pollard RE, Koehne AL, Peterson CB, Lyons LA. Japanese Bobtail: vertebral morphology and genetic characterization of an established cat breed. *Journal of Feline Medicine and Surgery*. 2014; 17(8):719–26. <https://doi.org/10.1177/1098612X14558147> PMID: 25488973
54. Xu L, Bickhart DM, Cole JB, Schroeder SG, Song J, Tassell CP, et al. Genomic signatures reveal new evidences for selection of important traits in domestic cattle. *Mol Biol Evol*. 2015; 32(3):711–25. <https://doi.org/10.1093/molbev/msu333> PMID: 25431480.
55. Gandolfi B, Alhaddad H, Abdi M, Bach LH, Creighton EK, Davis BW, et al. Applications and efficiencies of the first cat 63K DNA array. *Scientific reports*. 2018; 8(7014).
56. Alhaddad H, Khan R, Grahn RA, Gandolfi B, Mullikin JC, Cole SA, et al. Extent of linkage disequilibrium in the domestic cat, *Felis silvestris catus*, and its breeds. *PLoS One*. 2013; 8(1):e53537. Epub 2013/01/12. <https://doi.org/10.1371/journal.pone.0053537> PMID: 23308248.
57. O'Neill DG, Romans C, Brodbelt DC, Church DB, Černá P, Gunn-Moore DA. Persian cats under first opinion veterinary care in the UK: demography, mortality and disorders. *Scientific Reports*. 2019; 9(1):12952. <https://doi.org/10.1038/s41598-019-49317-4> PMID: 31530836
58. Qiu M, Bulfone A, Martinez S, Meneses JJ, Shimamura K, Pedersen RA, et al. Null mutation of *Dlx-2* results in abnormal morphogenesis of proximal first and second branchial arch derivatives and abnormal differentiation in the forebrain. *Genes Dev*. 1995; 9(20):2523–38. Epub 1995/10/15. <https://doi.org/10.1101/gad.9.20.2523> PMID: 7590232.
59. Robledo RF, Rajan L, Li X, Lufkin T. The *Dlx5* and *Dlx6* homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes Dev*. 2002; 16(9):1089–101. Epub 2002/05/10. <https://doi.org/10.1101/gad.988402> PMID: 12000792.
60. Nicoli E-R, Weston MR, Hackbarth M, Becerril A, Larson A, Zein WM, et al. Lysosomal Storage and Albinism Due to Effects of a *De Novo CLCN7* Variant on Lysosomal Acidification. *The American Journal of Human Genetics*. 2019; 104(6):1127–38. <https://doi.org/10.1016/j.ajhg.2019.04.008> PMID: 31155284
61. Fujimoto A, Nishida N, Kimura R, Miyagawa T, Yuliwulandari R, Batubara L, et al. *FGFR2* is associated with hair thickness in Asian populations. *Journal of Human Genetics*. 2009; 54(8):461–5. <https://doi.org/10.1038/jhg.2009.61> PMID: 19590514
62. Robinson R. The American Curl Cat. *Journal of Heredity*. 1989; 80(6):474–5. <https://doi.org/10.1093/oxfordjournals.jhered.a110902> PMID: 2614061

63. Trueb B, Taeschler S. Expression of FGFR1, a novel fibroblast growth factor receptor, during embryonic development. *Int J Mol Med*. 2006; 17(4):617–20. <https://doi.org/10.3892/ijmm.17.4.617> PMID: 16525717
64. Wiedemann M, Trueb B. Characterization of a Novel Protein (FGFR1) from Human Cartilage Related to FGF Receptors. *Genomics*. 2000; 69(2):275–9. <https://doi.org/10.1006/geno.2000.6332> PMID: 11031111
65. Slavotinek A, Risolino M, Losa M, Cho MT, Monaghan KG, Schneidman-Duhovny D, et al. De novo, deleterious sequence variants that alter the transcriptional activity of the homeoprotein PBX1 are associated with intellectual disability and pleiotropic developmental defects. *Hum Mol Genet*. 2017; 26(24):4849–60. <https://doi.org/10.1093/hmg/ddx363> PMID: 29036646.
66. Kehler JS, David VA, Schaffer AA, Bajema K, Eizirik E, Ryugo DK, et al. Four independent mutations in the feline fibroblast growth factor 5 gene determine the long-haired phenotype in domestic cats. *J Hered*. 2007; 98(6):555–66. <https://doi.org/10.1093/jhered/esm072> PMID: 17767004.
67. Drogemuller C, Rufenacht S, Wichert B, Leeb T. Mutations within the FGF5 gene are associated with hair length in cats. *Animal genetics*. 2007; 38(3):218–21. <https://doi.org/10.1111/j.1365-2052.2007.01590.x> PMID: 17433015.
68. Kienzle E, Moik K. A pilot study of the body weight of pure-bred client-owned adult cats. *British Journal of Nutrition*. 2011; 106(S1):S113–S5. Epub 2011/10/12. <https://doi.org/10.1017/S0007114511001802> PMID: 22005404
69. Cao Y, Zhao FL, Kolli T, Hivley R, Herness S. GABA expression in the mammalian taste bud functions as a route of inhibitory cell-to-cell communication. *Proc Natl Acad Sci U S A*. 2009; 106(10):4006–11. Epub 2009/02/19. <https://doi.org/10.1073/pnas.0808672106> PMID: 19223578.
70. Han DH, Bolo N, Daniels MA, Lyoo IK, Min KJ, Kim CH, et al. Craving for alcohol and food during treatment for alcohol dependence: modulation by T allele of 1519T>C GABAAalpha6. *Alcohol Clin Exp Res*. 2008; 32(9):1593–9. Epub 2008/07/12. <https://doi.org/10.1111/j.1530-0277.2008.00734.x> PMID: 18616664.
71. Chan Yingguang F, Jones Felicity C, McConnell E, Bryk J, Bünger L, Tautz D. Parallel Selection Mapping Using Artificially Selected Mice Reveals Body Weight Control Loci. *Current Biology*. 2012; 22(9):794–800. <https://doi.org/10.1016/j.cub.2012.03.011> PMID: 22445301
72. Fischer L, Wilde C, Schöneberg T, Liebscher I. Functional relevance of naturally occurring mutations in adhesion G protein-coupled receptor ADGRD1 (GPR133). *BMC Genomics*. 2016; 17(1):609. <https://doi.org/10.1186/s12864-016-2937-2> PMID: 27516204
73. Schneider MR, Werner S, Paus R, Wolf E. Beyond Wavy Hairs: The Epidermal Growth Factor Receptor and Its Ligands in Skin Biology and Pathology. *The American Journal of Pathology*. 2008; 173(1):14–24. <https://doi.org/10.2353/ajpath.2008.070942> PMID: 18556782
74. Introne W, Boissy RE, Gahl WA. Clinical, Molecular, and Cell Biological Aspects of Chediak–Higashi Syndrome. *Molecular Genetics and Metabolism*. 1999; 68(2):283–303. <https://doi.org/10.1006/mgme.1999.2927> PMID: 10527680
75. Buckley RM, Grahn RA, Gandolfi B, Herrick JR, Kittleson MD, Bateman HL, et al. Assisted reproduction mediated resurrection of a feline model for Chediak–Higashi syndrome caused by a large duplication in LYST. *Scientific Reports*. 2020; 10(1):64. <https://doi.org/10.1038/s41598-019-56896-9> PMID: 31919397
76. Lyons LA, Grahn RA, Genova F, Beccaglia M, Hopwood JJ, Longeri M. Mucopolysaccharidosis VI in cats—clarification regarding genetic testing. *BMC Vet Res*. 2016; 12(1):136-. <https://doi.org/10.1186/s12917-016-0764-y> PMID: 27370326.
77. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*. 2007; 81(3):559–75. <https://doi.org/10.1086/519795> PMID: 17701901.
78. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015; 4:7. <https://doi.org/10.1186/s13742-015-0047-8> PMID: 25722852.
79. Weir BS, Cockerham CC. Estimating F-Statistics for the Analysis of Population Structure. *Evolution*. 1984; 38(6):1358–70. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x> PMID: 28563791.
80. Willing EM, Dreyer C, van Oosterhout C. Estimates of genetic differentiation measured by F(ST) do not necessarily require large sample sizes when using many SNP markers. *PloS one*. 2012; 7(8):e42649. <https://doi.org/10.1371/journal.pone.0042649> PMID: 22905157.
81. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011; 27(15):2156–8. <https://doi.org/10.1093/bioinformatics/btr330> PMID: 21653522.