

RESEARCH

Open Access



Whole-genome sequencing of CRFK and PG-4 cells to infer the phenotype of the original donor cats

Ganma Tanaka^{1†}, Rikuto Goto^{1†}, Tetsushi Komoto^{1,2}, Akiko Kubota¹, Reo Hayashi¹, Takeshi Igawa^{1,2}, Naoaki Sakamoto^{1,3} and Akinori Awazu^{1,3*}

Abstract

Background Crandell-Rees Feline Kidney (CRFK; kidney-derived) cells and PG-4 cells (astrocyte-derived) have been in use and have been passaged for decades in laboratories worldwide; however, no detailed information on the genetic background of the donor individuals is available, particularly regarding phenotype characteristics such as coat and iris color.

Results We performed whole-genome sequencing of CRFK and PG-4 cells. We analyzed the resulting data to infer the phenotype of the individual from which the cells were derived, specifically for the coat color, coat length, coat pattern, and iris color. Our data suggested that CRFK cells originated from a cat with long black fur lacking stripes and with non-blue irises; PG-4 cells originated from a cat with long bicolored white and black fur without stripes, and with non-blue irises. Analysis of publicly available RNA-seq data confirmed that genes associated with coat phenotype and iris color are expressed in the skin and eyes, as well as in various other organs.

Conclusions Variants of the genes affecting coat phenotype and iris color may influence physiological functions throughout the body. These results shed light on the previously unknown genetic background of commonly used feline cultured cells and the phenotype of the donor individuals. These findings may facilitate more accurate interpretation of data obtained from cultured feline cells and provide guidelines for developing cell lines with domestic cat genotypes exhibiting diverse phenotypes through genome editing. This will help to provide clarity regarding deafness in cats with white fur and blue irises and elucidate the effects of melanocyte destruction caused by KIT gene mutations on the nervous system and the influence of other coat-related factors in organs other than the skin or functions independent of coat formation.

Keywords Genomics, Coat pattern, Iris color, RNA sequencing

[†]Ganma Tanaka and Rikuto Goto contributed equally to this work and share first authorship.

*Correspondence:
Akinori Awazu
awa@hiroshima-u.ac.jp

¹Graduate School of Integrated Sciences for Life, Hiroshima University, 1-3-1, Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan

²Amphibian Research Center, Hiroshima University, Higashihiroshima, Hiroshima, Japan

³Research Center for the Mathematics on Chromatin Live Dynamics, Hiroshima University, Higashihiroshima, Hiroshima, Japan



Background

Domestic cats (*Felis catus*) are widely kept as companion animals around the world, and the phenotypic diversity of their coat color, coat pattern, coat length, and iris color has been a topic of interest to breeders and scientists alike. Recent advances in genomics have markedly improved our understanding of the genetic basis controlling these diverse phenotypes, and the respective genotype-phenotype relationships are being increasingly elucidated [1]. Examples include the *TYR* gene, which is the color locus associated with albinism and a color point phenotype [2–4]; the *TYRP1* gene, conveying brown coat color [2, 5]; the *MC1R* genes, associated with in red coat color [6]; the *KIT* gene, determining white coat and white spotting [7–11]; the *ARHGAP36* gene, involved in orange coat color [12, 13]; the *MLPH* gene, causing dilute pigmentation [14]; the *ASIP* gene, controlling the expression of the agouti pattern [15]; the *LVRN* gene, determining the type of tabby pattern [16]; the *DKK4* gene, associated with the ticked tabby pattern [17]; the *FGF5* gene, involved in long-fur variants [18, 19]; the *KRT71* and *LPAR6* genes, involved in hairless and rexing (curly hair) phenotypes [20–22]; and the *PAX3* gene, which is associated with blue irises [23, 24]. Furthermore, technological advancements such as long-read sequencing facilitated the construction of a more continuous and accurate draft genome of the domestic cat [25–27], and a plethora of genetic polymorphisms associated with disease have been identified [17, 25, 28–32].

Immortalized cell lines are an indispensable tool in current biomedical and veterinary science. Crandell-Rees Feline Kidney (CRFK) cells, derived from the kidneys of a domestic cat [33, 34], have been used for numerous virological studies, including feline calicivirus [35–37], feline immunodeficiency virus [38–41], feline coronavirus [42–45], and various feline endogenous retroviruses such as RD114 [46–48], as well as for vaccine development [49–51]. PG-4 cells, derived from feline fetal astrocytes, have also played an important role in retrovirus research [52, 53].

Additionally, Crandell et al. [33] reported that CRFK cells exhibited aneuploidy ($2n=37$) compared with the normal feline autosome count of $2n=38$, and demonstrated a characteristic karyotype with three marker chromosomes. Lasfargues et al. [54] also observed that the chromosome count fluctuated between 36 and 38 in each cell. By contrast, no such findings have been reported for PG-4 cells.

These cell lines have been passaged for decades in laboratories worldwide and have facilitated numerous scientific breakthroughs. However, despite their widespread use, no detailed information on the genetic background of the donor individuals, particularly with regard to their phenotype, such as coat and iris color, is available.

Genomic information on such cell lines is crucial for a deeper understanding of their biological properties, and elucidating the phenotypes of their donors may provide new insights into the influence of specific genotypes on cell properties.

In this study, we performed whole-genome sequencing of CRFK and PG-4 cell lines. The sequencing data were used for in-depth analysis of the nucleotide sequences of genes known to be associated with coat and iris phenotypes (i.e., *TYR*, *TYRP1*, *MC1R*, *KIT*, *ARHGAP36*, *MLPH*, *ASIP*, *LVRN*, *DKK4*, *FGF5*, *KRT71*, *LPAR6*, and *PAX3*; Table 1). Based on the genotyping results for each gene, we inferred the previously unknown phenotypes of the donor individuals. Furthermore, analysis of publicly available RNA-seq data confirmed that these genes were expressed in skin and eyes as well as in various other tissues throughout the body. The findings and genome sequence data are expected to provide fundamental insights into the genetic background of extensively used feline cell lines. This information will support the interpretation of future research outcomes involving these cell lines and contribute to the development of guidelines for establishing cell lines with various genotypes through genome editing.

Methods

Cell lines and culture conditions

CRFK cells (JCRB9035, RRID: CVCL_2426) and a feline sarcoma-positive leukemia-negative (S+L-) astrocyte cells termed PG-4 (S+L-) (JCRB9125, RRID: CVCL_3322) were obtained from the JCRB Cell Bank (National Institutes of Biomedical Innovation, Health and Nutrition, Osaka, Japan). CRFK cells were cultured in Dulbecco's Modified Eagle Medium (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) supplemented with 10% fetal bovine serum (FBS) (Gibco), penicillin (100 IU/mL), streptomycin (100 ng/mL) (Sigma-Aldrich, St. Louis, MO, USA), and 1% non-essential amino acids (FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan) at 37 °C and 5% CO₂. PG-4 cells were cultured in McCoy's 5A Modified Medium (Gibco) supplemented with 10% heat-inactivated FBS (Gibco), penicillin (100 IU/mL), and streptomycin (100 ng/mL) (Sigma-Aldrich) at 37 °C and 5% CO₂. For cell passaging, cells were washed using BASIC DPBS (no calcium or magnesium; Gibco) and then treated with TrypLE™ Express Enzyme (1X; Gibco) for 5 min at 37 °C and 5% CO₂ conditions.

Paired-end short-read whole-genome sequencing

Genomic DNA was extracted from logarithmically growing cells (approximately 1×10^6 cells) using the DNeasy Blood & Tissue Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. The concentration and purity (A260/A280 ratio) of the extracted DNA were

measured using a NanoDrop 2000c Spectrophotometer (Thermo Fisher Scientific). Genomic DNA extracted from CRFK cells and PG-4 cells was used for library preparation and paired-end sequencing (150 bp read length) at Novogene (Beijing, China) using a Rapid Plus DNA Lib Prep Kit for Illumina V2 (Illumina) and an Illumina NovaSeq X Plus platform (Illumina). Library quality control was performed by Novogene using qPCR and fragment size analysis according to Novogene's standard protocols. The raw sequence data were obtained as .fastq files.

Mapping of paired-end reads and quantification

Adapter sequences were trimmed, and low-quality reads were removed from the original .fastq files using fastp software (ver. 0.21.0) with default parameters. Reads in processed .fastq files were mapped to the reference genome sequence of *Felis catus*, Fcatus_Fca126_mat1.0 (RefSeq assembly GCF_018350175.1) or AnAms1.0 (GenBank assembly GCA_013340865.2) [27] using bwa mem software (ver. 0.7.19) with the setting “-R”@RG\tID:sample1\tSM:sample1\tLB:lib1\tPL:ILLUMINA” and default parameters otherwise. Mapping results were obtained as .sam files. Text data in .sam files were transformed to binary data in sorted .bam files using Samtools (ver. 1.7) with default parameters. Index files were created for the obtained .bam files in the form of .bam.bai files.

Small variations in coat morphology-related genes

The variations of coat morphology-related phenotypes of the donor cats from which CRFK and PG-4 cell lines originated were inferred based on the comparisons between the wild-type sequences and mapped sequences on variant positions of each coat morphology-related gene (Table 1). Coat phenotype-related polymorphisms in *TYR*, *TYRP1*, *MC1R*, *KIT*, *MLPH*, *ASIP*, *LVRN*, *DDK4*, *FGF5*, *KRT71*, and *LPAR6* genes, which relate to phenotypes such as the color, brown, red, white glove, dilution, agouti, tabby, ticked, long-fur, hairless, and rexing (Table 1), were examined using alignments of paired-end sequence reads with wild-type sequences.

Large deletions in intron 1 of *ARHGAP36* and the intergenic region between *KIT* and *KDR* genes

The deletion of a part of intron 1 of the *ARHGAP36* gene (X: 109,186,183_109,191,258 in AnAms1.0; Table 1) was reported to produce orange coat color [12, 13], and that of the intergenic region between the *KIT* and *KDR* genes (B1: 161142880_161237957 in Fcatus_Fca126_mat1.0, Table 1) produces salmiak coat color [11]. The occurrence of these deletions in CRFK and PG-4 was estimated by comparing the mapped read number distributions

between these regions and their neighboring regions with the same region length. It should be noted that Fcatus_Fca126_mat1.0 has a deletion of the genomic region corresponding to X: 109186183_109191258 in AnAms1.0 (Table 1). Thus, AnAms1.0 (or felCat9 [25]) was used as the reference genome for analysis of the orange locus.

Insertions of FERV1-LTR and RD114-LTR into *PAX3*

The insertions of homologous sequences of the long terminal repeat (LTR) of the feline endogenous retrovirus 1 (FERV1), termed FERV1-LTR, into a specific region in intron 4 of *PAX3* (C1: g.205833101_205833102 in Fcatus_Fca126_mat1.0, Table 1) or the LTR of RD114 retrovirus, termed RD114-LTR, into another specific region in the same intron (C1: g.205834854_205834855 in Fcatus_Fca126_mat1.0, Table 1) are associated with blue iris color in cats [23, 24]. The occurrence of such insertions was estimated by extracting the genome reads that were mapped in both FERV1-LTR and intron 4 of *PAX3* or both RD114-LTR sequences and intron 4 of *PAX3*. The genome sequence of RD114-LTR was obtained from the NCBI database [55].

Long-read whole-genome sequencing

Genomic DNA of CRFK cells and PG-4 cells was extracted for long-read genome sequencing using the NucleoBond HMW DNA kit (TaKaRa, Tokyo, Japan), and libraries for the Oxford Nanopore Technologies (ONT, UK) sequencer were constructed using a ligation library preparation kit (SQK-LSK114, ONT). All procedures were performed according to the manufacturer's instructions, and the quality and molecular weight of the genomic DNA were measured using the Qubit fluorometric quantification system (Thermo Fisher Scientific). Sequencing was conducted using a PromethION sequencer (PromethION 2 Solo, ONT) and R10.4.1 flow cells. POD5 files were basecalled to .fastq files using MinKNOW software v25.05.14.

Insertions of FERV1 homolog into *KIT* examined by long-read sequencing

The insertion of FERV1 into a specific region of intron 1 of the *KIT* gene (B1: 161388687_161388688 in Fcatus_Fca126_mat1.0, Table 1) and that of FERV1-LTR make the coat color of cats white-spotted and white throughout, respectively [8]. The occurrence of such insertions in CRFK and PG-4 cells was estimated by mapping the reads of long-read sequencing to the genome sequence named KIT-FERV1 sequence using minimap2 (ver. 2.30), where the KIT-FERV1 sequence is a part of the *KIT* intron 1 sequence with the insertion of the FERV1 sequence, as shown in Supplementary Fig. 1 of the study by David et al. [8].

Table 1 Phenotypic traits determined by DNA variants, and genotypes of CRFK and PG-4 cells

Locus Gene	Phenotype (Alleles)	Mutation in CDS (or related noncoding region): F.catus_Fca126_mat1.0 Genome: F.catus_Fca126_mat1.0 Genome: AnAms1.0	Depth of alleles, and Genotype in CRFK	Depth of alleles, and Genotype in PG-4
Color TYR [2–4]	Burmese (c ^b)	c0.679G>T	+ : 18 c ^b : 0	+ : 16 c ^b : 0
		D1: g0.44022819C>A D1: g0.46081417C>A	+/+	+/+
	Siamese (c ^s)	c0.904G>A	+ : 25 c ^s : 0	+ : 17 c ^s : 20
		D1: g0.44013031C>T D1: g0.46071661C>T	+/+	+/c ^s
Brown TYRP1 [2, 5]	Albino (c ^a)	c0.939del	+ : 25 c ^a : 0	+ : 34 c ^a : 0
		D1: g0.44012996del D1: g0.46071626del	+/+	+/+
	Mocha (c ^m)	c0.820_936delinsAATCTC	+ : 40 c ^m : 0	+ : 54 c ^m : 0
		D1: g0.44012836_44012998delinsGAGATT D1: g0.46071466_46071628delinsGAGATT	+/+	+/+
Amber MC1R ⁶	Chocolate (b)	c0.1261+5G>A	+ : 22 b : 0	+ : 26 b : 0
		D4: g0.38142265 G>A D4: g0.39391308 G>A	+/+	+/+
	Cinnamon (b ^l)	c0.298C>T	+ : 20 b ^l : 0	+ : 19 b ^l : 0
		D4: g0.38130163C>T D4: g0.39379209C>T	+/+	+/+
White KIT [7–11]	Red (e)	c0.250G>A	+ : 35 e : 0	+ : 34 e : 0
		E2: g0.61570294G>A E2: g0.62623146G>A	+/+	+/+
	White (w ^w)	(Intron 1: FERV1-LTR insertion) B1: g0.161388687_161388688insFERV1-LTR B1: g0.163130991_163130992insFERV1-LTR	F10.3 +/+	Fg. 3 +/+
		(Intron 1: FERV1 insertion) B1: g0.161388687_161388688insFERV1 B1: g0.163130991_163130992insFERV1	Fg. 3 +/+	Fg. 3 +/w ^s
Orange ARHGAP36 [12, 13]	White spotted (w ^s)	c0.1035_1036delinsCA	+ : 27 w ^s : 0	+ : 28 w ^s : 0
		B1: g0.161337385_161337386delinsTG B1: g0.163079561_163079562delinsTG	+/+	+/+
	Salmiak (w ^{sal})	(KIT-KDR intergenic region: 95 kb deletion) B1: g0.161142880_161237957del B1: g0.162885414_162980052del	Fg. 2a +/+	Fg. 2a +/+
		(Intron 1: 5 kb deletion) (Lacking) X: g0.109186183_109191258del c0.83del C1: g0.218197448del C1: g0.219535339del	Fg. 2b +/+	Fg. 2b +/+
Dilution MLPH [14]	Dilute (d)	+ : 15 d : 4	+ : 15 d : 4	+ : 38 d : 0
		+/d	+/d	+/+

Table 1 (continued)

Locus Gene	Phenotype (Alleles)	Mutation in CDS (or related noncoding region): Fcatus_Fca126_mat1.0 Genome: Fcatus_Fca126_mat1.0 Genome: AnAms1.0	Depth of alleles, and Genotype in CRFK	Depth of alleles, and Genotype in PG-4
Agouti ASIP [15]	Nonagouti (a)	c0.123_124del	+ : 0	+ : 0
		A3: g0.24831083-24831084del	a : 48	a : 25
Tabby LVRN [16]	Blotched tabby (t ^b)	A3: g0.25283750_25283751del	a/a	a/a
		c0.176C>A	+ : 16	+ : 17
	Blotched tabby (t ^b)	A1: g0.94477275C>A	t ^b : 0	t ^b : 0
		A1: g0.97133509C>A	+/+	+/+
	Blotched tabby (t ^b)	c0.416C>A	+ : 34	+ : 10
		A1: g0.94477515C>A	t ^b : 0	t ^b : 12
	Blotched tabby (t ^b)	A1: g0.97133749C>A	+/+	+/t ^b
		c0.682G>A	+ : 33	+ : 22
	Blotched tabby (t ^b)	A1: g0.94477781G>A	t ^b : 0	t ^b : 0
		A1: g0.97134015G>A	+/+	+/+
	Blotched tabby (t ^b)	c0.2522G>A	+ : 0	+ : 5
		A1: g0.94536796G>A	t ^b : 10	t ^b : 3
	Blotched tabby (t ^b)	A1: g0.97192983G>A	t ^b /t ^b	+/t ^b
		c0.188G>A	+ : 36	+ : 20
Ticked DKK4 ¹⁷	Ticked (ti)	B1: g0.40429492G>A	ti : 0	ti : 0
		B1: g0.41805975G>A	+/+	+/+
	Ticked (ti)	c0.53C>T	+ : 33	+ : 30
		B1: g0.40428846C>T	ti : 0	ti : 0
	Long fur (l)	B1: g0.41805329C>T	+/+	+/+
		c0.356_357insT	+ : 24	+ : 22
Longhair FGF5 [18, 19]	Long fur (l)	B1: g0.139634675_139634676insA	l : 0	l : 0
		B1: g0.141368532_141368533insA	+/+	+/+
	Long fur (l)	c0.406C>T	+ : 24	+ : 25
		B1: g0.139645946G>A	l : 0	l : 0
	Long fur (l)	B1: g0.141379803G>A	+/+	+/+
		c0.474del	+ : 26	+ : 0
	Long fur (l)	B1: g0.139634261del	l : 0	l : 7
		B1: g0.141368118del	+/+	+/l
	Long fur (l)	c0.475A>C	+ : 0	+ : 0
		B1: g0.139634260T>G	l : 26	l : 13
	Long fur (l)	B1: g0.141368117T>G	l/l	l/+
		c0.577G>A	+ : 37	+ : 27
	Long fur (l)	B1: g0.139634158C>T	l : 0	l : 0
		B1: g0.141368015C>T	+/+	+/+

Table 1 (continued)

Locus Gene	Phenotype (Alleles)	Mutation in CDS (or related noncoding region): Fcatus_Fca126_mat1.0 Genome: Fcatus_Fca126_mat1.0 Genome: AnAms1.0	Depth of alleles, and Genotype in CRFK	Depth of alleles, and Genotype in PG-4
Hairless Rexing KRT71 [20, 21]	Hairless (hr) Rexing (re)	c0.816+1 G>A B4: g0.7894169C>T B4: g0.80471148C>T	+ : 40 hr : 0 +/+	+ : 25 hr : 0 +/+
		c0.1108-4_1184del, c0.1184_1185insAGTTGGAG, and c0.1196_1197insT B4: g0.78939390_78939470del, B4: g0.78939389_78939390insCTCCAACT, and B4: g0.78939377_78939378insA B4: g0.80468919_80468919del, B4: g0.80468918_80468919insCTCCAACT, and B4: g0.80468826_80468827insA	+ : 51 re : 0 +/+	+ : 32 re : 0 +/+
Rexing LPAR6 [22]	Rexing (re)	c0.445-1 G>C B4: g0.78943283C>G B4: g0.80472733C>G	+ : 37 re : 0 +/+	+ : 29 re : 0 +/+
		c0.250_253del A1: g0.22865220_22865223del A1: g0.23372953_23372956del	+ : 16 re : 0 +/+	+ : 19 re : 0 +/+
Blue iris PAX3 [23, 24]	Blue iris (DBE ^{Cel}) Blue iris (DBE ^{ALT})	(Intron4: FERV1-LTR homolog insertion) C1: g0.205833101_205833102insN[395] C1: g0.207180017_207180018insN[395] (Intron4: RD114-LTR homolog insertion) C1: g0.205834854_205834855insN[433] C1: g0.207181773_207181774insN[433]	+ : DBE ^{Cel} : 0 +/+ + : DBE ^{ALT} : 0 +/+	+ : DBE ^{Cel} : 0 +/+ + : DBE ^{ALT} : 0 +/+

List of genes and phenotypes we focused on in this study, mutation position on CDS and genomic coordinates provided by Fca126_mat1.0 and AnAms1.0, coverage depth of alleles, and a summary of genotypes of CRFK and PG-4 cells as determined in the present study

Transcriptome data analysis of coat phenotype- and iris color-related genes in various tissues of domestic cats

Mapped RNA-seq data on the F.catus_Fca126_mat1.0 were obtained in .bam and .bam.bai formats for various tissues from the Ensembl public database [56], specifically, the file transfer protocol (ftp) site [57]. The annotation file of the F.catus_Fca126_mat1.0 in gene transfer format (GTF) was obtained from the NCBI database (Bethesda, MD, USA). In the file, the chromosome numbers were converted from NC_058368.1, NC_058369.1, etc. to A1, A2, etc.

Using FeatureCounts with the options -M-O --fraction [58], read counts data for each gene described in the GTF file were obtained from the .bam and .bam.bai formatted files. This count data could also be obtained from a file “F.catus_Fca126_mat1.0.ENA_gene_exp_mat_by_FeatureCounts.csv” deposited in Fcatlas [59]. From these counts, the transcript level of each gene was calculated by dividing the read count by the total exon length. The total transcript level was defined as the sum of all the transcript levels of the genes. Based on this result, the transcripts per million (TPM) of each gene were calculated as $1,000,000 * (\text{transcript level}) / (\text{total transcript level})$.

Results

Short- and long-read whole-genome sequencing

For CRFK and PG-4 cells, the total output of raw sequence data of 150 bp paired-end sequencing was 60.6 Gb and 60.1 Gb, and the [average]±[standard deviation] of mapping depth of the reads for F.catus_Fca126_mat1.0 were 23.59 ± 9.79 and 23.45 ± 8.17 , respectively; the total output of raw sequence data of long-read sequencing was 9.6 Gb and 11.8 Gb, N50 values were 23.5 kb and 38.2 kb, and the [average]±[standard deviation] of mapping depth of reads for F.catus_Fca126_mat1.0 were 3.91 ± 3.11 and 4.79 ± 3.34 , respectively. The average and standard deviations specified above were estimated through the entire genome region without the regions with singular sequences where the read coverage exhibited > 100.

Predicted phenotype of the CRFK cell donor

The mapping of paired-end reads from the whole-genome sequences of the CRFK cell revealed that most coat morphology-related genes were conserved (Table 1). However, the following variations were identified in mapped genome reads from CRFK cells (Table 1): a homozygous mutation of c.123_124del of the *ASIP* gene (*nonagouti* allele), a homozygous mutation of c.475A>C of the *FGF5* gene (*long coat* allele; Fig. 1), a homozygous mutation of c.2522G>A of the *LVRN* gene (*blotched-tabby* allele), and a heterozygous mutation of c0.83delT of the *MLPH* gene (*dilute* allele). No 95 kbp deletion was detected in the intergenic region between the *KIT* and *KDR* genes,

nor a 5 kbp deletion was observed in intron 1 of the *ARHGAP36* gene (Fig. 2). Additionally, no genome reads were identified that could be simultaneously mapped to both the *PAX3* intron 4 and FERV1-LTR or to both the *PAX3* intron 4 gene and RD114-LTR sequences. Furthermore, four reads from the long-read sequencing mapped to the KIT-FERV1 sequence; however, these reads contained no insertions of FERV1 and FERV1-LTR (Fig. 3).

The abovementioned homozygous mutation of the *FGF5* gene was known to exhibit long fur [18]. However, the mutation of the *LVRN* did not influence coat morphology in this case because of the mutation of the *ASIP* gene that suppresses the stripe pattern on the cat's skin coat, and the mutant genotype of the *MLPH* gene was recessive [1]. Therefore, the cat from which the CRFK cells originated is expected to have had long, black, unstriped fur and non-blue eyes (Fig. 4).

Predicted phenotype of the PG-4 cell donor

Mapped paired-end reads from the whole-genome sequences of the PG-4 cell indicated that most coat morphology-related genes were conserved but revealed the following variations (Table 1): a homozygous mutation of c.123_124del of the *ASIP* gene, a compound heterozygous mutation of c.474delT and c0.475A>C of the *FGF5* gene (Fig. 1) where one allele exhibits c.474delT and the other allele exhibits c.475A>C as mentioned below, the heterozygous mutations of c.416C>A and c.2522G>A of the *LVRN* gene, and a heterozygous mutation of c.904G>A of the *TYR* gene (*Siamese-color* allele). No 95 kbp deletion was identified in the intergenic region between the *KIT* and *KDR* genes, nor a 5 kbp deletion was found in intron 1 of the *ARHGAP36* gene (Fig. 2), as was similarly observed with CRFK. Additionally, we did not identify any genome reads that mapped concurrently to both the *PAX3* intron 4 and FERV1-LTR or to both the *PAX3* intron 4 gene and RD114-LTR sequences.

Notably, the genome sequence reads obtained from PG-4 cells containing bases 474 and 475 of the CDS of the *FGF5* gene revealed that seven reads exhibited only a deletion at base 474, while the remaining 13 reads showed only a substitution at base 475 (Fig. 1). Accordingly, we expected a compound heterozygous mutation in the *FGF5* gene of PG-4 cells. Similarly, PG-4 cells exhibited heterozygous substitutions at bases 416 and 2522 of the CDS of the *LVRN* gene. However, owing to the considerable distance between the genomic locations of these two bases, none of our sequencing data, including the long-read sequencing data, produced reads that contained both bases together. Therefore, we were unable to conclude whether the *LVRN* gene in PG-4 cells harbors a compound heterozygous mutation.

Six long-read sequences mapped to the KIT-FERV1 sequence, with four of them containing the insertions of

(a) CRFK

(a) CRFK		FGF5: c. 474, c. 475	
		↓	
WT		tccccccaccaggccaaatt	ta ccgatgactgcaagttcaggggagcgattccaagaaaaac
FGF5: c. 474del		tccccccaccaggccaaatt	-a ccgatgactgcaagttcaggggagcgattccaagaaaaac
FGF5: c. 475A>C		tccccccaccaggccaaatt	tc ccgatgactgcaagttcaggggagcgattccaagaaaaac
Read name and sequence			
_R_LH00504:186:22YKYJLT3:1:1136:23098:8018/2	-----	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
_R_LH00504:186:22YKYJLT3:1:1192:25992:16846/1	-----aatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
_R_LH00504:186:22YKYJLT3:1:1187:1481:21444/1	-----caggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
_R_LH00504:186:22YKYJLT3:1:1220:37134:16734/1	----ccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:2219:7629:10021/2	--ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
_R_LH00504:186:22YKYJLT3:1:2212:14989:24120/1	tccccccaccaggccaaatt	tc	ccgatgactacaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:2139:51206:15580/1	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1247:37291:5630/1	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
_R_LH00504:186:22YKYJLT3:1:2246:43264:25113/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
_R_LH00504:186:22YKYJLT3:1:2246:43255:25129/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:2185:31669:12744/2	tccccccaccaggccaaatt	t-	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1132:13233:8867/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1106:21286:25834/1	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1172:1749:11559/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1172:1740:11575/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1181:35396:19906/1	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:2274:13473:12616/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:2274:18586:13209/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1267:31882:27341/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
_R_LH00504:186:22YKYJLT3:1:1204:11014:26171/1	tccccccaccgggccaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1192:25992:16846/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1161:30662:6960/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:1161:30671:6976/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:2212:14989:24120/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:2157:26510:25049/2	tccccccaccaggccaaatt	tc	ccgatgactgcaagttcaggggagcgattccaagaaaaac
LH00504:186:22YKYJLT3:1:2144:20306:10389/2	tccccccaccgggccaatt	tc	ccgatgac-----

(b) PG-4

WT		FGF5: c. 474del		FGF5: c. 475A>C	
Read name and sequence					
_R_LH00504:188:22Y5KLLT3:4:2128:16034:3147/2	-----aatt	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag
_R_LH00504:186:22YKYJLT3:1:1209:31290:3628/1	-----aggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2102:7361:11959/1	-----caggccaaatt	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag
_R_LH00504:186:22YKYJLT3:1:1153:39436:4957/2	--ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2154:16238:26411/2	ccccaccaggccaaatt	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2217:22645:11783/2	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
_R_LH00504:186:22YKYJLT3:1:2203:41359:11014/2	ccccaccaggccaaatt	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:188:22Y5KLLT3:4:1147:27379:6303/2	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2249:9710:19490/2	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:1153:39436:4957/1	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2149:37162:21492/2	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:1270:2452:27453/1	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2276:50356:6095/2	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2154:27490:12200/2	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2136:34998:5022/1	ccccaccaggccaaatt	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag
_R_LH00504:186:22YKYJLT3:1:1233:2174:8963/2	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
_R_LH00504:186:22YKYJLT3:1:1233:2165:8979/2	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:2287:47203:29279/2	ccccaccaggccaaatt	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag
_R_LH00504:186:22YKYJLT3:1:1202:22719:14218/1	ccccaccaggccaaatt	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag	tc	ccgatgactgcaagttcaggagcgattccaagaaaaacag
LH00504:186:22YKYJLT3:1:1222:6446:7233/1	ccccaccaggccaaatt	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag	-a	ccgatgactgcaagttcaggagcgattccaagaaaaacag

Fig. 1 Alignment results of whole genome sequencing reads of CRFK and PG-4 cells with WT and mutant sequences of FGF5 genes. Alignment results of genome sequencing reads around FGF5: c.0.474 and c.0.475; CRFK (a) and PG-4 (b) cells. Left string indicates the name of sequenced read, and red colored symbols indicate coat phenotype-related mutations

the FERV1 homolog sequence exceeding 7 kbp (Fig. 3). This indicates that the PG-4 genome contains the heterozygous insertion of the entire FERV1 homolog sequence.

The mutant genotype of the *TYR* gene was recessive, and in this case, the two heterozygous mutations of the

LVRN did not influence coat morphology because of the *ASIP* gene mutation, the same as in CRFK [1]. However, the effect of *FERV1* is known to be dominant [8]. Additionally, the compound heterozygous mutation of the *FGF5* gene produces long fur [18]. Therefore, the PG-4

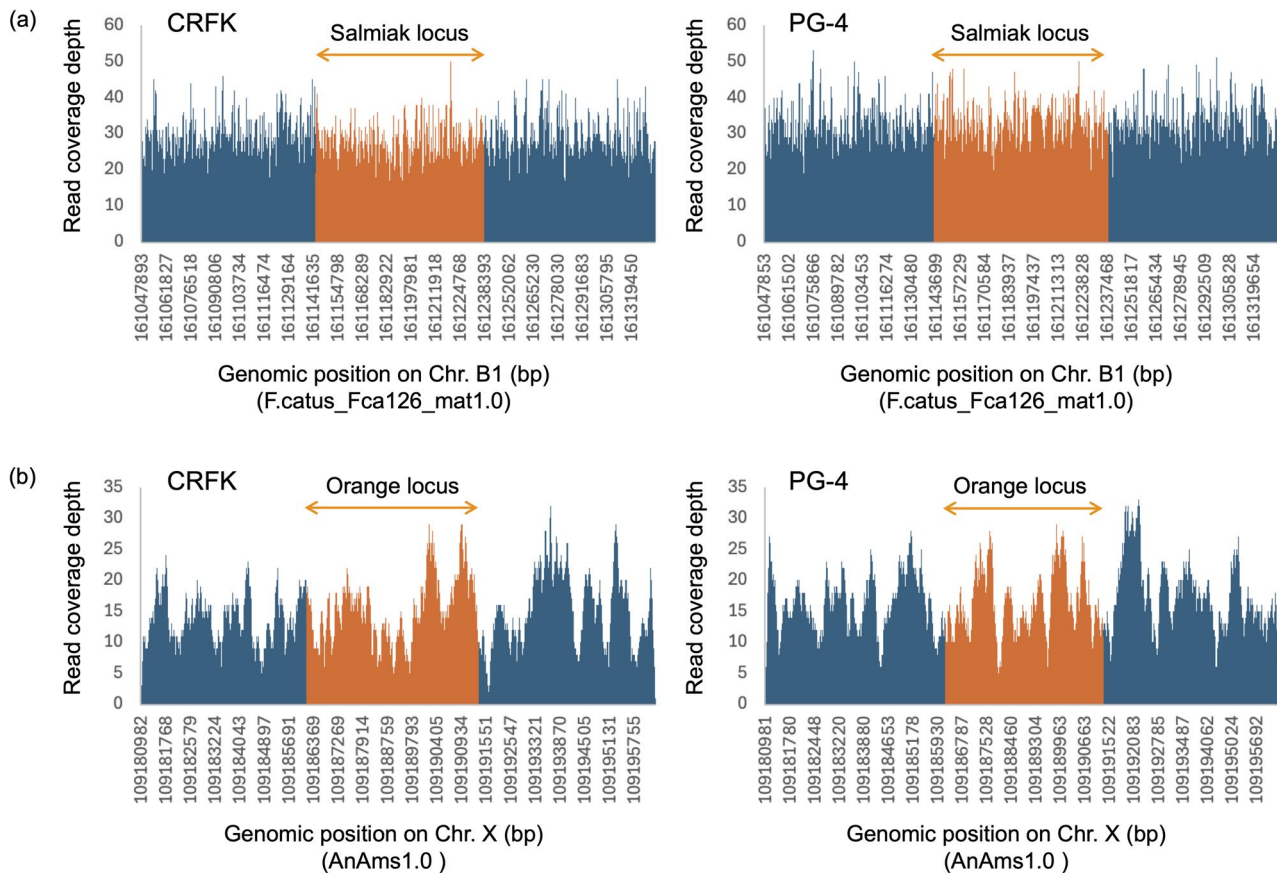


Fig. 2 Read coverage depth distributions of CRFK and PG-4 cells around the salmiak and orange loci. Read coverage depth distributions from 95 kbp upper stream to 95 kbp lower stream of the salmiak locus (a), and those from 5.1 kbp upper stream to 5.1 kbp lower stream of the orange locus (b) of CRFK and PG-4 cells. Bar graph colored orange indicates coverage depth of reads mapped to the salmiak (a) and orange loci (b), respectively. In all distributions, no notable declines in read coverage depth were observed in both salmiak and orange loci, indicating that both cats from which the CRFK and PG-4 cells express no salmiak and orange coat

cells likely originated from a cat with white and black (white spotting), unstriped, long fur (Fig. 4).

Notably, certain cats with white spotting may have blue eyes or atypical eye coloration; however, this occurrence is less common than that in cats exhibiting the Dominant White allele in the *KIT* gene [8]. Additionally, no insertions of FERV1-LTR and RD114-LTR in intron 4 of the *PAX3* gene were observed. Therefore, the PG-4 cells likely originated from a cat with non-blue irises (Fig. 4).

Expression features of coat phenotype- and iris color-related genes in various tissues

The analysis of publicly available RNA-seq data from various organs of domestic cats in the Ensembl database (https://ftp.ensembl.org/pub/data_files/felis_catus/F.catus_Fca126_mat1.0/rnaseq/) revealed that coat phenotype-related genes are expressed in various organs at levels (TPM) comparable to or greater than those in the skin (Fig. 5, Table S2). For example, *TYR*, *TYRP1*, and *MLPH* were highly expressed in the ear tip, retina, and optic nerve, and *MC1R*, *KIT*, *DDK4*, and *FGF5* were highly expressed

in the brain. *KIT*, *DDK4*, and *LPAR6* tend to be expressed throughout the body, with prominent expression in the reproductive organs. Additionally, *ARHGAP36* was highly expressed in the spinal cord, and *ASIP*, along with *MLPH*, was highly expressed in the lungs. The iris color-related gene *PAX3* was also highly expressed in the cerebellum, in addition to visual organs, and, along with *DDK4*, in the kidney and embryonal tissue.

Discussion

We performed whole-genome sequencing of CRFK, a cultured feline kidney cell line, and PG-4, a feline fetal astrocyte-derived cell line, and analyzed the genome data to determine the coat color, hair length, coat pattern, and iris color of the donor individuals. The results suggested that the CRFK cell line originated from a cat with long black hair without stripes and non-blue irises, and the PG-4 cell line came from a cat with long fur, black and white bicolor coat without stripes, and non-blue irises.

Coat morphology and iris color are phenotypes that can be readily identified and used to predict genotypes.

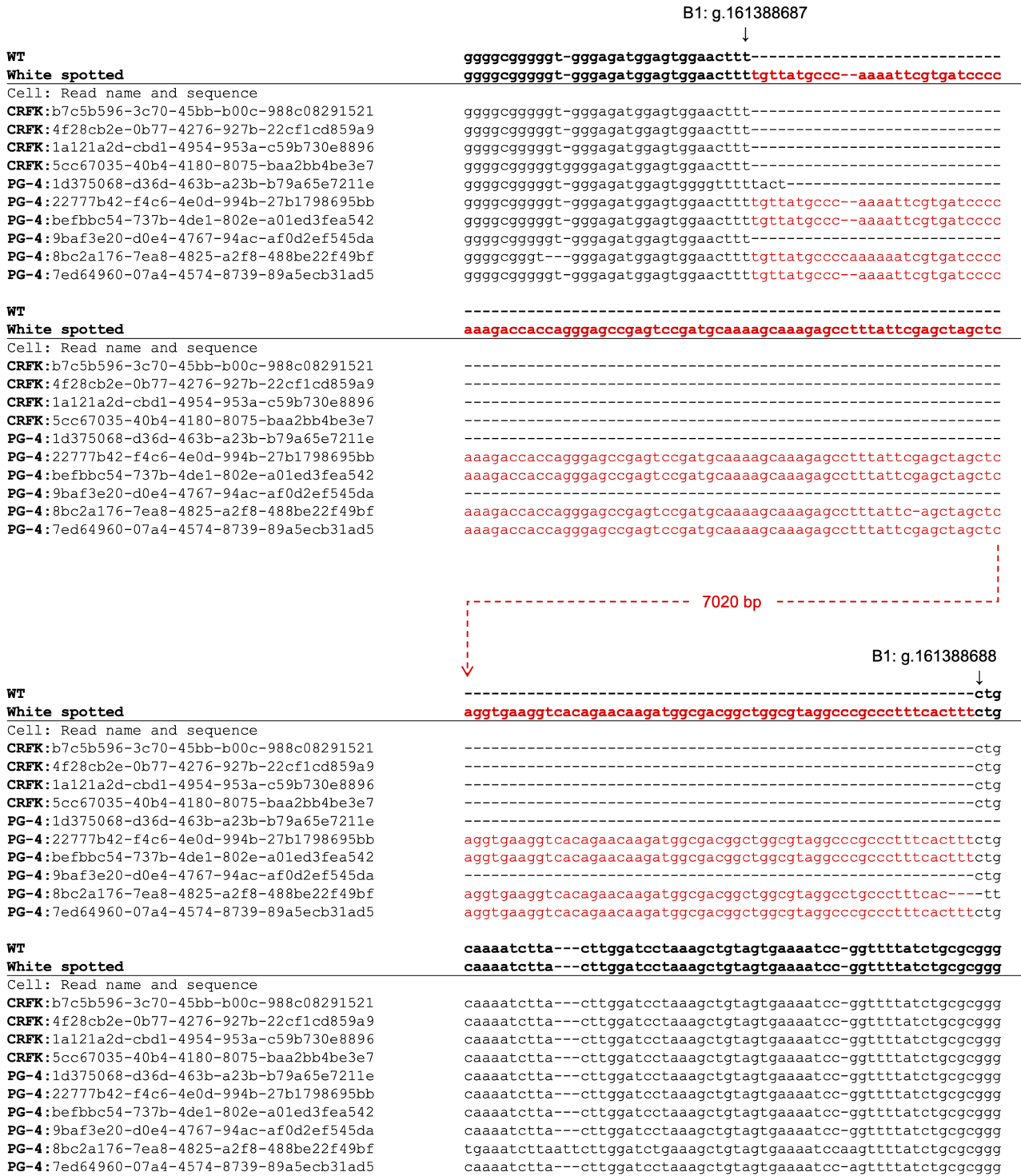


Fig. 3 Parts of upper and lower streams of alignment results of long-read sequencing reads of CRFK and PG-4 cells with WT and KIT-FERV1 sequences of KIT intron 1. Alignment results of long-read genome sequencing reads from CRFK and PG-4 cells to the KIT-FERV1 that contains sequence around the region B1: g.161388687_161388688 (F.Catus_Fca126_mat1.0). Left string indicates [the name of cell (CRFK or PG-4): name of sequenced read], and red colored sequences indicate FERV1 and aligned sequences with it. Four long reads from CRFK could align, but no reads contained the FERV1 homologue, while 6 reads from PG-4 could align, and four of the six reads contained the FERV1 homologue. This figure shows only a part of the results at the upper and lower streams of the KIT-FERV1 sequence. See table S1 for the alignment result with the full-length KIT-FERV1 sequence

(a) Original cat of CRFK cell

Coat: Long
Color: black
Stripes: No
Irises: Not blue



(b) Original cat of PG-4 cell

Coat: Long
Color: White and black
Stripes: No
Irises: Not blue



Fig. 4 Example illustrations of inferred original cats of CRFK and PG-4 cells. An example illustration of the inferred phenotypes of coat morphologies and iris colors of the original cats of the CRFK cell (a) and the PG-4 cell (b). CRFK cells are expected to originate from a cat with long, black fur, lacking stripes, and with non-blue irises. PG-4 cells were expected to be derived from a cat with long, white and black bicolored fur without stripes and non-blue irises

However, the expressions of these genes are not limited to the skin and eyes; rather, they are expressed in various tissues throughout the body (Fig. 5). Therefore, differences in these genotypes may contribute to disease susceptibility in specific organs. The effects of these genetic variations are likely weaker than those of the rearing environment and may not be significant, especially in young, healthy individuals. However, as domestic cats have become longer-lived in recent years, individual-specific risks may become more apparent, and the underlying genotypes may exert increasing effects as the cats age. Therefore, elucidating the relationships between genotypes underlying readily recognizable phenotypes and disease risk is crucial from the perspectives of both medical care and welfare, such as prevention and symptom suppression/alleviation.

Results from research on cultured cells to elucidate disease mechanisms and develop therapeutic drugs must be interpreted with caution, i.e., not only on the experimental results obtained using established, immortalized cells but also with regard to their genotype and expression profiles of the genes harboring polymorphisms in various tissues and organs. However, so far, such information has not been available, even for certain frequently used cell lines such as CRFK. The present study and genome sequence data provide such information and facilitate a more comprehensive interpretation of research on feline cultured cell lines.

Furthermore, with the availability of detailed genomic information, genome editing can be applied to these cell lines to establish various derivative lines that differ only in specific aspects, e.g., regarding coat morphology or other traits, and compare their properties. This is expected to advance research on the regulatory mechanisms of coat color, for example, the validation of the role of the novel *KIT* isoform caused by FERV1-LTR insertion

in the formation of white coat [60], as well as the relationships between the visually identifiable phenotypes examined in the present study and disease risk and treatment methods. Simultaneously, this comprehensive genomic information can help to develop guidelines for establishing cell lines with various disease-associated genotypes using genome editing. The development of these derivative cell lines may provide valuable insights into the mechanism underlying deafness in cats with white fur and blue irises, as well as elucidate the influence of other coat-related factors in organs beyond the skin or functions unrelated to coat formation.

The limitations and future directions of this study are outlined below.

First, notably, the findings of this study are specific to the cell types used. The cells used in this study were CRFK cells (JCRB9035, RRID: CVCL_2426), which were prepared from CRFK (ATCC: CCL 94) and PG-4 (S+L-) cells (JCRB9125, RRID: CVCL_3322) obtained from PG-4 (S+L-) cells (ATCC: CRL-2032). Consequently, these cells were expected to have similar characteristics to those of the CRFK and PG-4 cells used in numerous previous studies. For example, however, the number of chromosomes has been reported to vary from 36 to 38 in CRFK [33, 54]. Additionally, the karyotype of PG-4 cells, which has not yet been reported, likely differs from the normal domestic cat genome. Therefore, using the same cell line may still produce results that differ slightly from those used by other researchers. Investigating additional variation among cells in these lines remains an important future consideration.

As mentioned previously, the karyotype of CRFK differs from that of the normal domestic cat genome [33, 54]. Additionally, the specific chromosomes that cause aneuploidy remain unidentified. However, identifying large structural variations using only whole-genome



Fig. 5 Expression levels of coat phenotype-related and iris color-related genes in various tissues. The expression level of each gene in each tissue is quantified as $\log_2(\text{transcripts per million}+1)$. The length of the red bar for gene a for each tissue is proportional to the [expression levels of gene a in each tissue]/[the maximum expression level of gene a among all focused tissues]. Most coat phenotype-related genes were found to exhibit similar or higher expressions in various tissues other than skin, and the iris color-related gene (PAX3) was also found to exhibit similar or higher expressions in a variety of tissues other than the opt-functional organs

analysis of the present sequence read data, approximately 60 Gb paired-end sequencing data and 10 Gb long-read data, was challenging. Similarly, accurately identifying sequences and locations of various long repeat sequences, endogenous retroviral sequences, and their homologs was challenging owing to limited read coverage, particularly for long-read coverage. However, these sequences may also be associated with various diseases, as demonstrated by the involvement of FERV1-LTR in deafness in cats [8, 23, 24] and the positive and negative roles of enFeLV-LTR and enFeLV-env in resistance to exogenous feline leukemia virus [61–63]. The increased

volume of ongoing long-read sequencing, along with the draft genome assembly using these data, is expected to address these challenges in the future.

This study focused on coat and iris phenotypes in domestic cats. However, major histocompatibility complex (MHC) polymorphisms are strongly correlated with breed in domestic dogs [64–68]. Although research on MHC polymorphisms in domestic cats is ongoing [69–71], the relationship with breed remains unclear. In this study, we showed the results of the primary analysis of the small indels and SNPs in MHC class I and II regions of CRFK and PG-4 cells described by Variant Call format

(VCF) [72] (Table S3). However, because the MHC gene cluster comprises a complex and large genomic region, it is difficult to identify the donor breeds of CRFK and PG-4 using only the sequence data and analytical methods presented in this study. However, as mentioned above, further long-read sequencing and its data analysis may enable breed estimation of these cell lines.

In addition, the present genome sequence data also included mitochondrial DNA data. In this study, we presented the results of the primary analysis of the small indels and SNPs in the mitochondrial DNA of CRFK and PG-4 cells described by VCF [72] (Table S4 [72]). Although the correlations are not substantially strong, an association may exist between the polymorphisms of mitochondrial DNA and breeds of domestic cats [73–76]. Therefore, a comprehensive analysis of this data may also contribute to breed estimation.

These analyses are also critical future challenges.

Conclusions

We performed whole-genome sequencing of CRFK and PG-4 cells and inferred the phenotype of the donors of these cells. We suggested that CRFK cells originated from a cat with long, black fur lacking stripes and non-blue irises; PG-4 cells originated from a cat with long, bicolored white and black fur without stripes, and with non-blue irises. Additionally, analysis of publicly available RNA-seq data confirmed that genes associated with coat phenotype and iris color are expressed in the skin and eyes, as well as in various other organs, indicating that variants of these genes, which affect coat phenotype and iris color, may influence physiological functions throughout the body. These insights may facilitate a more accurate interpretation of data derived from feline cultured cells and inform guidelines for developing cell lines with domestic cat genotypes that exhibit various phenotypes using genome editing techniques. This will help to elucidate the role of various coat phenotype-related and iris color-related factors in organs other than the skin and functions independent of coat formation, including the effects of melanocyte destruction caused by KIT gene mutations on the nervous system.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40575-026-00150-9>.

Supplementary Material 1: Legends of Table S1 - S4.

Supplementary Material 2: Table S1

Supplementary Material 3: Table S2

Supplementary Material 4: Table S3

Supplementary Material 5: Table S4

Acknowledgements

Not applicable.

Author contributions

Conceptualization: A.A., Data curation: T.K., A.K., R.H., and T.I., Formal analysis: G.T. and R.G., Funding acquisition: A.A., Investigation: G.T., R.G., T.K., R.H., and A.A., Methodology: G.T., R.G., T.K., T.I., and A.A., Project administration: N.S. and A.A., Supervision: N.S. and A.A., Validation: R.G. and A.A., Visualization: T.G., N.S., and A.A., Writing - original draft: A.A., Writing - review & editing: T.G., T.K., T.I., and N.S.

Funding

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI grants (24K01783 (T. I.) and 24K09248 (A. A.)).

Data availability

These raw read data of the short and long read sequencings were deposited in the DDBJ Sequence Read Archive (DRA) with BioProject accession PRJDB37530 (DRR748391-DRR748394, DRR786948, and DRR786949). All other data generated or analyzed during this study were included in this published article (and Supplementary Information files).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 10 November 2025 / Accepted: 26 February 2026

Published online: 06 March 2026

References

1. Lyons LA. DNA mutations of the cat: the good, the bad and the ugly: the good, the bad and the ugly. *J Feline Med Surg*. 2015;17:203–19. <https://doi.org/10.1177/1098612X15571878>.
2. Schmidt-Kuntzel A, Eizirik E, O'Brien SJ, et al. Tyrosinase and tyrosinase related protein 1 alleles specify domestic cat coat color phenotypes of the Albino and brown loci. *J Hered*. 2005;96:289–301.
3. Lyons LA, Imes DL, Rah HC, et al. Tyrosinase mutations associated with Siamese and Burmese patterns in the domestic cat (*Felis catus*). *Anim Genet*. 2005;36:119–26.
4. Imes DL, Geary LA, Grahm RA, et al. Albinism in the domestic cat (*Felis catus*) is associated with a tyrosinase (TYR) mutation. *Anim Genet*. 2006;37:175–78.
5. Lyons LA, Foe IT, Rah HC, et al. Chocolate coated cats: TYRP1 mutations for brown color in domestic cats. *Mamm Genome*. 2005;16:356–66.
6. Peterschmitt M, Grain F, Arnaud B, et al. Mutation in the melanocortin 1 receptor is associated with amber colour in the Norwegian forest cat. *Anim Genet*. 2009;40:547–52.
7. Montague MJ, Li G, Gandolfi B, et al. Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proc Natl Acad Sci USA*. 2014;111:17230–35.
8. David VA, Menotti-Raymond M, Wallace AC, et al. Endogenous retrovirus insertion in the KIT oncogene determines white and white spotting in domestic cats. *G3 (Bethesda)*. 2014;4:1881–91.
9. Frischknecht M, Jagannathan V, Leeb T. Whole genome sequencing confirms KIT insertions in a white cat. *Anim Genet*. 2015;46:98.
10. Górka A, Drobik-Czwaro W, Bryś J. Genetic determination of the amount of white spotting: a case study in Siberian cats. *Genes (Basel)*. 2022;13(6).
11. Anderson H, Salonen M, Toivola S, Blades M, Lyons L A, Forman OP, Hytönen MK, Lohi H A new Finnish flavor of feline coat coloration, “salmiak,” is associated with a 95-kb deletion downstream of the KIT gene. *Anim Genet*. 2024;55:676–80.
12. Kaelin CB, McGowan KA, Trotman JC, Koroma DC, David VA, Menotti-Raymond M, Graff EC, Schmidt-Küntzel A, Oancea E, Barsh GS Molecular and

- genetic characterization of sex-linked orange coat color in the domestic cat. *Curr Biol.* 2025;35:2826–36.e5. Pubmed reference: 40378841. doi: <https://doi.org/10.1016/j.cub.2025.04.055>.
13. Toh H, Au Yeung WK, Unoki M, Matsumoto Y, Miki Y, Matsumura Y, Baba Y, Sado T, Nakamura Y, Matsuda M, Sasaki H A deletion at the X-linked ARH-GAP36 gene locus is associated with the orange coloration of tortoiseshell and calico cats. *Curr Biol.* 2025;35:2816–25.e3, 2025. Pubmed reference: 40378840. doi: <https://doi.org/10.1016/j.cub.2025.03.075>.
 14. Ishida Y, David VA, Eizirik E, et al. A homozygous single-base deletion in MLPH causes the dilute coat color phenotype in the domestic cat. *Genomics.* 2006;88:698–705.
 15. Eizirik E, Yuhki N, Johnson WE, et al. Molecular genetics and evolution of melanism in the cat family. *Curr Biol.* 2003;13:448–53.
 16. Kaelin CB, Xu X, Hong LZ, et al. Specifying and sustaining pigmentation patterns in domestic and wild cats. *Science.* 2012;337:1536–41.
 17. Lyons LA, Buckley RM, Harvey RJ. 99 Lives cat genome Consortium. Mining the 99 Lives cat genome sequencing Consortium database implicates genes and variants for the Ticked locus in domestic cats (*Felis catus*). *Anim Genet.* 2021;52:321–32. <https://doi.org/10.1111/age.13059>.
 18. Kehrer JS, David VA, Schaffer AA, et al. Four independent mutations in the feline fibroblast growth factor 5 gene determine the long-haired phenotype in domestic cats. *J Hered.* 2007;98:555–66.
 19. Drogemüller C, Rufenacht S, Wicher B, et al. Mutations within the FGF5 gene are associated with hair length in cats. *Anim Genet.* 2007;38:218–21.
 20. Gandolfi B, Outerbridge C, Beresford L, et al. The naked truth: sphynx and Devon Rex cat breed mutations in KRT71. *Mamm Genome.* 2010;21:509–15.
 21. Gandolfi B, Alhaddad H, Joslin SE, et al. A splice variant in KRT71 is associated with curly coat phenotype of selkirk Rex cats. *Sci Rep.* 2013;3:2000.
 22. Gandolfi B, Alhaddad H, Affolter VK, 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:e67105.
 23. Abitbol M, Couronné A, Dufaure de Citres C, Gache V. A PAX3 insertion in the celestial breed and certain feline breeding lines with dominant blue eyes. *Anim Genet.* 2024;55:670–75.
 24. Abitbol M, Dufaure de Citres C, Rudd Garces G, Lühken G, Lyons LA, Gache V. Different founding effects underlie dominant blue eyes (DBE) in the domestic cat. *Anim (Basel).* 2024;14:1845. Pubmed reference: 38997957. doi: <https://doi.org/10.3390/ani14131845>.
 25. 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. *PLoS Genet.* 2020;16(10):e1008926. <https://doi.org/10.1371/journal.pgen.1008926>.
 26. Bredemeyer KR, Harris AJ, Li G, Zhao L, Foley NM, Roelke-Parker M, O'Brien SJ, Lyons LA, Warren WC, Murphy WJ Ultracontinuous single haplotype genome assemblies for the domestic cat (*Felis catus*) and Asian leopard cat (*Prionailurus bengalensis*). *J Hered.* 2021;112(2):165–73. <https://doi.org/10.1093/jhered/esaa057>.
 27. Matsumoto Y, Chung CYL, Isobe S, Sakamoto M, Lin X, et al. Chromosome-scale assembly with improved annotation provides insights into breed-wide genomic structure and diversity in domestic cats. *J Adv Res.* 2025;75:863–74.
 28. Lyons LA, Creighton EK, Alhaddad 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>.
 29. Rodney AR, Buckley RM, Fulton RS, et al. A domestic cat whole exome sequencing resource for trait discovery. *Sci Rep.* 2021;11:7159. <https://doi.org/10.1038/s41598-021-86200-7>.
 30. Anderson H, Davison S, Lytle KM, Honkanen L, Freyer J, Mathlin J, et al. Genetic epidemiology of blood type, disease and trait variants, and genome-wide genetic diversity in over 11,000 domestic cats. *PLoS Genet.* 2022;18(6):e1009804. <https://doi.org/10.1371/journal.pgen.1009804>.
 31. Hernandez I, Hayward JJ, Brockman JA, White ME, et al. Complex feline disease mapping using a dense genotyping array. *Front Vet Sci.* 2022;9:862414.
 32. Raffle J, Novo Matos J, Wallace M, et al. Identification of novel genetic variants associated with feline cardiomyopathy using targeted next-generation sequencing. *Sci Rep.* 2025;15:3871.
 33. Crandell RA, Fabricant CG, Nelson-Rees WA. Development, characterization, and viral susceptibility of a feline (*Felis catus*) renal cell line (CRFK). *Vitro.* 1973;9:176–85.
 34. Lawson JS, Syme HM, Wheeler-Jones CPD, Elliott J. Characterisation of Crandell-Rees feline kidney (CRFK) cells as mesenchymal in phenotype. *Res Vet Sci.* 2019;127:99–102.
 35. Sosnovtsev SV, Prikhod'ko EA, Belliot G, Cohen JI, Green KY. Feline calicivirus replication induces apoptosis in cultured cells. *Virus Res.* 2003;94:1–10.
 36. Bidawid S, Malik N, Adegbeunrin O, Sattar SA, Farber JM. A feline kidney cell line-based plaque assay for feline calicivirus, a surrogate for Norwalk virus. *J Virol Methods.* 2003;107:163–67.
 37. Tian J, Zhang X, Wu H, Liu C, Liu J, Hu X, Qu L Assessment of the IFN- β response to four feline caliciviruses: infection in CRFK cells. *Infect Genet Evol.* 2015;34:352–60.
 38. Siebelink KH, Karlas JA, Rimmelzwaan GF, Osterhaus AD, Bosch ML. A determinant of feline immunodeficiency virus involved in Crandell feline kidney cell tropism. *Vet Immunol Immunopathol.* 1995;46:61–69.
 39. Verschoor EJ, Boven LA, Blaak H, Van Vliet AL, Horzinek MC, De Ronde A. A single mutation within the V3 envelope neutralization domain of feline immunodeficiency virus determines its tropism for CRFK cells. *J Virol.* 1995;69:4752–57.
 40. Baldinotti F, Matteucci D, Mazzetti P, Giannelli C, Bandecchi P, Tozzini F, Bendinelli M Serum neutralization of feline immunodeficiency virus is markedly dependent on passage history of the virus and host system. *J Virol.* 1994;68:4752–79.
 41. Mizuno T, Goto Y, Baba K, Masuda K, Ohno K, Tsujimoto H. TNF- α -induced cell death in feline immunodeficiency virus-infected cells is mediated by the caspase cascade. *Virology.* 2001;287:446–55.
 42. Drechsler Y, Vasconcelos EJ, Griggs LM, Diniz PP. CRFK and primary macrophages transcriptomes in response to feline coronavirus infection differ significantly. *Front Genet.* 2020;11:584744.
 43. Drechsler Y, Vasconcelos EJ, Diniz PP. Host transcriptome studies in response to feline coronavirus reveal differences in macrophages vs CRFK. *J Immunol.* 2020;204(1_Supplement):92–27.
 44. Drechsler Y, Vasconcelos EJ, Griggs LM, Diniz PP. Host responses to feline coronavirus are significantly different in primary macrophages compared to CRFK cells. *J Immunol.* 2021;206(1_Supplement):19–13.
 45. Camero M, Lanave G, Catella C, Lucente MS, et al. ERDRP-0519 inhibits feline coronavirus in vitro. *BMC Vet Res.* 2022;18:55.
 46. Baumann JG, Günzburg WH, Salmons B. Crfk feline kidney cells produce an RD114-like endogenous virus that can package murine leukemia virus-based vectors. *J Virol.* 1998;72. <https://doi.org/10.1128/jvi.72.9.7685-7687.1998>.
 47. Shimode S, Nakagawa S, Miyazawa T. Multiple invasions of an infectious retrovirus in cat genomes. *Sci Rep.* 2015;5:8164.
 48. Shimode S, Sakuma T, Yamamoto T, Miyazawa T. Establishment of CRFK cells for vaccine production by inactivating endogenous retrovirus with TALEN technology. *Sci Rep.* 2022;12:6641.
 49. Lappin MR, Jensen WA, Jensen TD, Basaraba RJ, Brown CA, Radecki SV, Bendinelli M Investigation of the induction of antibodies against Crandell-Rees feline kidney cell lysates and feline renal cell lysates after parenteral administration of vaccines against feline viral rhinotracheitis, calicivirus, and panleukopenia in cats. *Am J Vet Res.* 2005;66:506–11.
 50. Whittemore JC, Hawley JR, Jensen WA, Lappin MR. Antibodies against Crandell-Rees feline kidney (CRFK) cell line antigens, α -enolase, and annexin A2 in vaccinated and CRFK hyperinoculated cats. *J Vet Intern Med.* 2010;24:306–13.
 51. Yoshikawa R, Sato E, Igarashi T, Miyazawa T. Characterization of RD-114 virus isolated from a commercial canine vaccine manufactured using CRFK cells. *J Clin Microbiol.* 2010;48:3366–69.
 52. Haapala DK, Robey WG, Oroszlan SD, Tsai WP. Isolation from cats of an endogenous type C virus with a novel envelope glycoprotein. *J Virol.* 1985;53:827–33.
 53. Bassin RH, Ruscetti S, Ali I, Haapala DK, Rein A. Normal DBA/2 mouse cells synthesize a glycoprotein which interferes with MCF virus infection. *Viol.* 1982;123:139–51.
 54. Lasfargues EY, Lasfargues JC, Dion AS, Greene AE, Moore DH. Experimental infection of a cat kidney cell line with the mouse mammary tumor virus. *Cancer Res.* 1976;36:67–72.
 55. RD114 retrovirus, complete genome. https://www.ncbi.nlm.nih.gov/nucleotide/NC_009889.1/. Accessed 4 August 2025.
 56. Cunningham F, Allen JE, Allen J, Amode MR, Armean IM, Azov AG, Barnes I, Bennett R, Berry A, Bhai J, Bignell A, Billis K, Boddus S et al. Ensembl 2022. *Nucl Acids Res.* 2022;50(D1):D988–95.
 57. Ensembl complete datasets and databases. https://ftp.ensembl.org/pub/data_files/felis_catus/F.catus_Fca126_mat1.0/maseq/. Accessed 26 January 2023.
 58. Liao Y, Smyth GK, Shi W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30:923–30.
 59. F catlas: <https://sites.google.com/view/fcatlas/>. Accessed 1 November 2025.

60. Awazu A, Takemoto D, Watanabe K, Sakamoto N. Possibilities of skin coat color-dependent risks and risk factors of squamous cell carcinoma and deafness of domestic cats inferred via RNA-seq data. *Genes To Cells: Devoted To Mol Cellular Mechanisms*. 2023;28:893–905.
61. Chiu ES, VandeWoude S. Presence of endogenous viral elements negatively correlates with feline leukemia virus susceptibility in puma and domestic cat cells. *J Virol*. 2020;94. <https://doi.org/10.1128/jvi.01274-20>.
62. Erbeck K, Gagne RB, Kraberger S, Chiu ES, Roelke-Parker M, VandeWoude S. Feline leukemia virus (FeLV) endogenous and exogenous recombination events result in multiple FeLV-B subtypes during natural infection. *J Virol*. 2021;95. <https://doi.org/10.1128/jvi.00353-21>.
63. Pramono D, Takeuchi D, Katsuki M, AbuEed L, Abdillah D, Kimura T, Kawasaki J, Miyake A, Nishigaki K FeLIX is a restriction factor for mammalian retrovirus infection. *J Virol*. 2024;98:e01771–23.
64. Kennedy LJ, Barnes A, Happ GM, Quinnell RJ, Bennett D, Angles JM, Day MJ, Carmichael N, Innes JF, Isherwood D, Carter SD, Thomson W, Ollier WE. Extensive interbreed, but minimal intrabreed, variation of DLA class II alleles and haplotypes in dogs. *Tissue Antigens*. 2002;59:194–204.
65. Angles JM, Kennedy LJ, Pedersen NC. Frequency and distribution of alleles of canine MHC-II DLA-DQB1, DLA-DQA1 and DLA-DRB1 in 25 representative American Kennel Club breeds. *Tissue Antigens*. 2005;66:173–84.
66. Miyamae J, Okano M, Katakura F, Kulski JK, Moritomo T, Shiina T. Large-scale Polymorphism analysis of dog leukocyte antigen class I and Class II genes (*DLA-88*, *DLA-12/88L* and *DLA-DRB1*) and comparison of the haplotype diversity between breeds in Japan. *Cells*. 2023;12:809.
67. Graumann MB, DeRose SA, Ostrander EA, Storb R. Polymorphism analysis of four canine MHC class I genes. *Tissue Antigens*. 1998;51(4 Pt 1):374–81.
68. Wagner JL, Burnett RC, Storb R. Organization of the canine major histocompatibility complex: current perspectives. *J Hered*. 1999;90:35–38.
69. Morris KM, Kirby K, Beatty JA, Barrs VR, Cattley S, David V, O'Brien SJ, Menotti-Raymond M, Belov K Development of MHC-Linked microsatellite markers in the domestic cat and their use to evaluate MHC diversity in domestic cats, cheetahs, and Gir Lions. *J Hered*. 2014;105:493–505.
70. Okano M, Miyamae J, Suzuki S, Nishiya K, Katakura F, Kulski JK, Moritomo T, Shiina T Identification of novel alleles and structural haplotypes of major histocompatibility complex class I and DRB genes in domestic cat (*Felis catus*) by a newly developed NGS-Based genotyping method. *Front Genet*. 2020;11:750.
71. Holmes JC, Holmer SG, Ross P, Buntzman AS, Frelinger JA, Hess PR. Polymorphisms and tissue expression of the feline leukocyte antigen class I loci FLAI-E, FLAI-H, and FLAI-K. *Immunogenetics*. 2013;65:675–89.
72. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R Genomes Project analysis group. The variant call format and VCFtools. *Bioinformatics*. 2010;2011(27):2156–58.
73. Grahn RA, Kurushima JD, Billings NC, Grahn JC, Halverson JL, Hammer E, Ho CK, Kun TJ, Levy JK, Lipinski MJ, Mwenda JM, Ozpinar H, Schuster RK, Shoorijeh SJ, Tarditi CR, Waly NE, Wictum EJ, Lyons LA Feline non-repetitive mitochondrial DNA control region database for forensic evidence. *Forensic Sci Int Genet*. 2011;5:33–42.
74. Wesselink M, Bergwerff L, Hoogmoed D, Kloosterman AD, Kuiper I. Forensic utility of the feline mitochondrial control region - a Dutch perspective. *Forensic Sci Int Genet*. 2015;17:25–32.
75. Bilgen N, Güvendi M, Sezgin E, Can H, Köseoğlu AE, Erkunt Alak S, Döşkaya M, Ün C Mitotyping of random bred cats and pure breed cats (Turkish Angora and Turkish Van) using non-repetitive mitochondrial DNA control region. *Gene*. 2024;892:147849.
76. Jaito W, Singchat W, Patta C, Thatukan C, Kumnan N, Chalermwong P, Budi T, Panthum T, Wongloet W, Wattanadilokchatkun P, Thong T, Muangmai N, Han K, Duengkae P, Phatcharakullawarawat R, Srikulnath K Shared alleles and genetic structures in different Thai domestic cat breeds: the possible influence of common racial origins. *Genomics Inf*. 2024;22:12.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.