basepaws

Feline Health and Trait Markers

Basepaws launched a DNA amplicon sequencing–based health and trait marker panel for pedigree and mixed–breed cats, testing for 115 manually–curated mutations associated with different diseases and physical traits. Our technology is agile and allows testing for diverse types of mutations, such as Single Nucleotide Polymorphisms (SNPs), inversions, duplications, deletions, and insertions with variable sequence length. As research progresses, we will be updating our health and trait marker panel to include the latest discoveries in feline genetic research.

Table of Contents:

Genetic Health and Trait Markers: Key Concepts

An SNP represents a single nucleotide difference in the same genomic region between organisms of the same species. SNPs are the most common type of genetic variation within species. SNPs occur normally in the DNA, and a randomly selected human genome is estimated to have four to five million SNPs, with over 100 million SNPs reported across human populations worldwide.[1] Research on pedigree and mixed-breed cats estimates over 17 million SNPs across different feline populations.[2] Most SNPs are typically found in intergenic regions and have no effect on health or the physical manifestation of traits.[1] However, some can serve as markers, helping the identification of genes and mutations associated with different diseases or traits. SNPs associated with an increased disease risk comprise the largest category of health markers and can be two main types[1,3,4]:

- **Linked SNPs (indicative SNPs):** These polymorphisms do not reside in the gene body and do not affect protein function. However, they are predictive of disease risk due to their proximity to disease-causing mutations.

- **Causative SNPs:** If an SNP occurs within the coding region of a gene (coding SNP) or in a regulatory element proximal to a gene (non-coding SNP), this could indicate that the SNP is disrupting the gene's function, and thus has a direct role in disease.

The same concept of linked and causative SNPs applies to genetic markers associated with physical traits. Some SNPs can have a passive association with a trait, while others could be linked to a functional change in a protein, which directly results in a particular phenotypic presentation.

In addition to SNPs, genetic markers can also be regions of abnormally inserted, deleted, inverted or duplicated genomic sequence (named insertions, deletions, inversions, and duplications, respectively).[5,6] Unlike SNPs, these markers can span across multiple nucleotides.[5]

Genetic markers can be used to track disease and trait inheritance patterns.[1] Generally, diseases and traits with a genetic component fall into four groups (Figure 1):

- **Autosomal dominant:** inheriting one mutated copy of a gene results in developing the disease or exhibiting the trait (e.g., Polycystic Kidney Disease).[7]
- **Autosomal recessive:** inheriting two mutated copies of a gene results in developing the disease or exhibiting the trait; inheriting just one mutated copy is associated with a carrier status where the individual does not have the disease or trait themselves, but can pass down the mutation to their offspring (e.g., Myotonia Congenita).[8]
- **X-linked**: inheriting a disease- or a trait-associated mutated copy of a gene located on the X chromosome results in males developing the disease or manifesting the trait and females being carriers (e.g., Haemophilia B).[9]
- **Polygenic**: inheriting multiple genetic variants (each associated with only a slightly elevated disease risk or likelihood of exhibiting a trait) together contributing to an increased likelihood of developing the disease or manifesting the trait.[10,11]

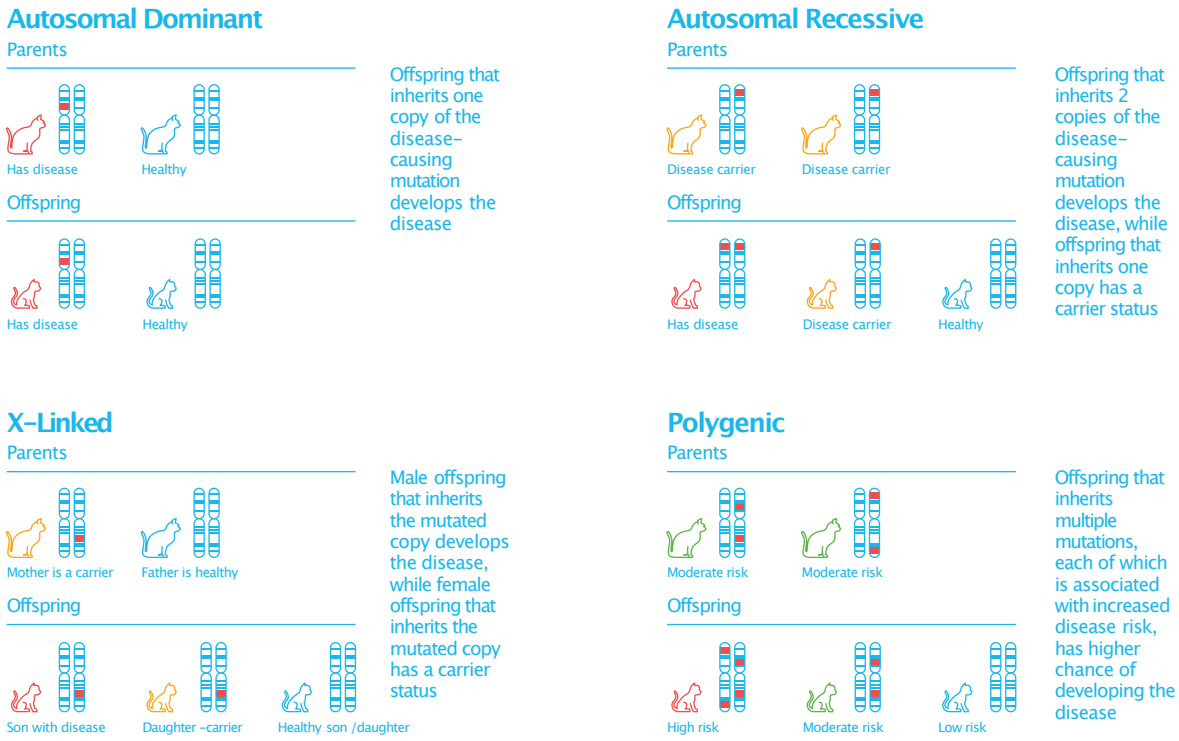


Figure 1. Inheritance patterns of genetic diseases.

basepaws

Genetic Marker Identification (Genotyping): Platform Comparison

Microarrays

Microarrays are a popular tool for genetic marker testing. Thousands or even hundreds of thousands of genetic markers can be assessed through a single microarray hybridization–based reaction. Microarrays consist of millions of short, synthetic, single–stranded DNA sequences attached to a solid surface, such as a chip or a bead. Some of these synthetic DNA sequences are complementary to the known normal (wild type) sequence of the genomic region of interest, while others are complementary to the known disease–related mutations occuring in the same genomic region. The DNA sample to be genotyped is denatured, fluorescently labeled, and applied to the chip (or bead) so that a hybridization reaction between the sample DNA and the synthetic DNA on the chip can occur. Results are interpreted on the basis of hybridization–dependent fluorescent patterns.[11,12]

Microarrays are a well–established, relatively inexpensive and easy to use technology, which makes them a popular choice for routine diagnostic testing. However, procedures can sometimes generate high levels of non–specific 'noise' fluorescence due to cross–hybridization (hybridization between only partially complementary DNA fragments and synthetic DNA probes on the microarray).[13] This phenomenon can confound result interpretation, making it difficult to determine whether the sample is heterozygous or homozygous for a particular marker. In addition, updating microarray genetic marker panels with new markers requires manufacturing new hardware and can become costly. Lastly, DNA microarrays are only useful for detecting markers associated with SNPs or small insertions, inversions, and duplications, since the length of the single–stranded synthetic DNA sequences attached to the microarray is shorter than 100 base pairs.[14] This makes the technology an inappropriate choice for identifying large insertions, inversions or duplications.

Sanger Sequencing

DNA sequencing allows the identification of the order of nucleotides comprising a genomic region of interest. The technology works on the basis of polymerase chain reaction (PCR) and the incorporation of fluorescently labeled nucleotides. Sanger sequencing has become the gold standard in clinical DNA sequencing and is appropriate for the identification of SNPs, as well as a wide variety of sizes of insertions, inversions, and duplications (as long as they are shorter than 1000 base pairs). Even though Sanger sequencing was used in the Human Genome Project, it is currently impractical and expensive to use the technology for querying more than a few genes at a time[11].

Next Generation Sequencing

Next generation sequencing (NGS) is a much faster and more easily scalable type of DNA sequencing compared to the Sanger method. Unlike Sanger, NGS allows the sequencing of millions of small DNA fragments in parallel, rather than sequentially.[12] NGS–mediated genotyping can be performed by either of the two methods below:

- **Targeted sequencing**: pre–selected regions of the genome are amplified through PCR and sequenced.
- **Whole genome sequencing:** sample DNA representative of the whole genome is fragmented, and PCR amplified and sequenced.and sequenced.

While high–depthwhole genome sequencing can be expensive, and is thus used mostly for exploratory purposes, targeted sequencing is suitable for screening multiple/hundreds of health markers at an affordable cost.

PROs and CONs of Genotyping Platforms

Microarrays, Sanger sequencing, and NGS all have advantages and disadvantages when it comes to health marker screening. A factor that can potentially skew the accuracy of sequencing results is the misrepresentation of allele frequency, due to PCR bias where a random PCR error can get propagated (discussed in more detail later).[15] This can lead to incorrect interpretation of zygosity. While microarrays generally do not require PCR, they can suffer from cross-hybridization bias,[13] which can also have a confounding effect on zygosity interpretation.

The accuracy of NGS is further affected by a phenomenon known as "barcode hopping". When multiple samples are sequenced together, each individual sample is tagged with a unique barcode sequence. Sometimes, depending on the sequencing library complexity, barcode sequences can be misassigned to the wrong sample.

With Illumina sequencing machines (the most widely used NGS instruments), this happens approximately 0.1–2% of the time.[16] Sanger sequencing's accuracy, on the other hand, can be affected by the purity of the DNA template being sequenced (PCR artefacts such as primer dimers can affect the results).[17] However, the resulting chromatogram produced with Sanger sequencing makes it relatively easy to identify this problem.

Sanger sequencing and NGS are superior to microarray genotyping when it comes to screening for health markers associated with insertions, inversions, and duplications longer than 100 base pairs.[14] However, when it comes to the ability to screen for multiple health markers at once, NGS and microarrays are at an advantage compared to Sanger sequencing, where scaling can become expensive. One caveat to microarrays' scalability is the significant cost and time associated with updating the platform to include new health markers.

Criteria	Microarray	Sanger Sequencing	Next Generation Sequencing
PCR bias	No potential for PCR bias	Potential for PCR bias	Potential for PCR bias
Accuracy	Potential for cross-hybridization	Accuracy depends on template purity	Potential for barcode hopping
Detection of long variants	Cannot detect long variants	Can detect long variants	Can detect long variants
Scalability	Scalable (but difficult to add new targets)	Not scalable	Scalable

Advantage Neither an advantage, nor a disadvantage Disadvantage

Basepaws' Feline Health and Trait Marker Selection

We used the OMIA (Online Mendelian Inheritance in Animals) database[18] and literature search to identify 65 feline genetic markers associated with 43 genetic disorders, and 50 genetic markers associated with 25 traits. We manually curated and rated all genetic markers, according to the amount and strength of supporting scientific evidence available for each one, using a five-star system. The evidence grading criteria we used are shown in Table 1. Tables 2 and 3 contain detailed information on all health and trait markers included in the Basepaws report.

Rating	Meaning
★ ★ ★ ★ ★	Consistent results reported across multiple scientific studies. More than 100 cats across studies confirmed to have the expected correlation between the genetic marker and physical presentation. Multiple scientific tools used to confirm the mechanism behind the mutation. No discordant results reported in literature.
★ ★ ★ ★	Results based on one or more studies reporting findings from at least 30 cats confirming the expected correlation between the genetic marker and physical presentation. Multiple scientific tools used to confirm the mechanism behind the mutation. No discordant results reported in literature.
★ ★ ★	Results based on one or more studies reporting findings from at least 10 cats confirming the expected correlation between the genetic marker and physical presentation. Multiple scientific tools may or may not have been used to confirm the mechanism behind the mutation.
★ ★	Results based on a study reporting findings from fewer than 10 cats confirming the expected correlation between the genetic marker and physical presentation. Multiple scientific tools may or may not have been used to confirm the mechanism behind the mutation. Potentially inconclusive inheritance pattern. No discordant results reported in literature.
★	Preliminary results to be interpreted with extreme caution. Based on a scientific study that focused on a single cat, or on only a few studies that report contradictory results.

Table 1. Scientific evidence strength grading criteria.

Table 2. Basepaws' health marker panel.

Condition type	Condition	Gene	Mutation	Evidence grading	Seen in
Metabolic Disorders	Mucopolysaccharidosis I	IDUA	del(TCG)	*	No specific breeds
		IDUA	del(GTC)	*	No specific breeds
	Mucopolysaccharidosis VI	ARSB	A > G	*****	Birman, Siamese
		ARSB	C >T	****	Birman, Siamese
	Mucopolysaccharidosis VII	GUSB	G >A	***	No specific breeds
		GUSB	T >G	*	No specific breeds
		GUSB	C >T	*	No specific breeds
	Neuronal Ceroid Lipofuscinosis (NCL)	CLN6	G >A	*	No specific breeds
	Porphyria, acute intermittent (AIP)	HMBS	ins (T)	*	Birman, Siamese
		HMBS	G >A	**	No specific breeds
		HMBS	C >T	*	No specific breeds
		HMBS	G >A	*	No specific breeds
		HMBS	del (GAG)	*	No specific breeds
	Porphyria, congenital erythropoietic (CEP)	UROS	C >T	*	No specific breeds
	Alpha-mannosidosis	MAN2B	del (GAGG)	**	Persian, Domestic long-haired
	Niemann-Pick disease, type A	SMPD1	G >A	*	No specific breeds
	Niemann-Pick disease, type C1	NPC1	C >G	***	No specific breeds
		NPC1	T >G	*	No specific breeds
	Niemann-Pick disease, type C2	NPC2	C >T	**	No specific breeds
	Gangliosidosis, type GM2 (variant 0)	HEXB	inv (TA{...}AC)	***	Korat, Domestic Shorthair
		HEXB	del (TA{...}CA)	***	Burmese, Korat, Domestic Shorthair
		HEXB	C >T	**	No specific breeds

Condition type	Condition	Gene	Mutation	Evidence grading	Seen in
Metabolic Disorders	Gangliosidosis, type GM1	GLB1	C >G	****	No specific breeds
	Gangliosidosis, type GM2 (variant AB)	GM2A	del (GACC)	**	No specific breeds
	Wilson's Disease	ATP7B	C >G	*	No specific breeds
	Dihydropyrimidinuria	DPYS	C >T	*	No specific breeds
	Pyruvate Kinase (PK) Deficiency	PKLR	G >A	*****	Abyssinian, Somali
	Hyperlipoproteinaemia	LPL	C >T	****	No specific breeds
Musculoskeletal and Connective Tissue Disorders	Hypokalemic Periodic Paralysis (HYPP)	WNK4	C >T	****	Burmese, Burmese-derived
	Vitamin D-deficiency rickets, type IA	CYP27B1	C >A	*	No specific breeds
		CYP27B1	del (C)	*	No specific breeds
	Vitamin D-deficiency rickets, type IB	CYP2R1	del (T)	*	No specific breeds
	Myotonia Congenita	CLCN1	G >T	**	No specific breeds
	Ehlers-Danlos Syndrome (classic type)	COL5A1	del (G)	*	No specific breeds
	Fibrodysplasia Ossificans	ACVR1	C >T	**	No specific breeds
Renal Disorders	Cystinuria, type B	SLC7A9	T >A	**	No specific breeds
		SLC7A9	A >G	*	No specific breeds
		SLC7A9	C >T	*	No specific breeds
		SLC7A9	G >A	*	No specific breeds
		SLC7A9	G >A	*	No specific breeds
	Cystinuria, type IA	NPC2	C >T	*	No specific breeds
	Polycystic Kidney Disease (PKD)	HEXB	inv (TA{...}AC)	*****	Persian, Persian-derived
	Primary Hyperoxaluria, type II (Oxalosis II)	HEXB	del (TA{...}CA)	**	No specific breeds

Condition type	Condition	Gene	Mutation	Evidence grading	Seen in
Cardiovascular Disorders	Hypertrophic Cardiomyopathy (HCM)	MYBPC3	G >A	*****	Ragdoll
		MYBPC3	C >G	*****	Maine Coon
		MYH7	C >T	*	No specific breeds
		TNNT2	C >T	**	Maine Coon
Blood Disorders	Factor XII Deficiency	F12	del (C)	****	No specific breeds
		F12	G >C	****	No specific breeds
	Methemoglobinemia	CYB5R3	C >T	*	No specific breeds
		CYB5R3	C >G	*	No specific breeds
	Haemophilia B	F9	C >T	*	No specific breeds
		F9	G >A	*	No specific breeds
	Glanzmann Thrombasthenia	ITGA2B	del (G)	*	No specific breeds
Eye Disorders	Glaucoma, primary congenital	LTBP2	del (CTCC)	***	Siamese
	Progressive Retinal Atrophy (AIPL1-related)	AIPL1	C >T	****	Persian, Persian-derived
	Late-Onset Photoreceptor Degeneration (rdAc)	CEP290	A >C	***	Abyssinian, Somali
	Progressive Retinal Atrophy (KIF3B-related)	KIF3B	C >T	**	Bengal
Endocrine Disorders	Congenital Adrenal Hyperplasia (CAH)	CYP11B1	G >A	*	No specific breeds
	Hypogonadotropic Hypogonadism, congenital (TAC3-related)	TAC3	C >T	*	No specific breeds
	Hypothyroidism, congenital	TPO	C >T	*	No specific breeds
Skin Disorders	Inflammatory Linear Verrucous Epidermal Nevus (ILVEN)	NPC2	C >T	*	No specific breeds
	Epidermolysis Bullosa Simplex	HEXB	inv (TA{...}AC)	*	No specific breeds
Autoimmune Disorders	Leukocyte Adhesion Deficiency (LAD)	ITGB2	del (GC[...]GC)	*	No specific breeds
	Autoimmune Lymphoproliferative Syndrome (ALPS)	FASLG	ins(A)	***	British Shorthair

Table 3. Basepaws' trait marker panel.

Condition type	Condition	Gene	Mutation	Evidence grading	Seen in
Blood Type	Blood Type	CMAH	G >T	****	No specific breeds
		CMAH	T >A	****	No specific breeds
		CMAH	C >T	****	No specific breeds
		CMAH	del (T)	****	No specific breeds
Coat Color and Pattern	Amber coat color	MC1R	G >A	***	Norwegian Forest Cat
	Russet coat color	MC1R	del (TCT)	***	Burmese
	Copal coat color	MC1R	del (GC[...]GG)	**	Kurilian Bobtail
	Cinnamon (light brown) coat color	TYRP1	C >T	*	No specific breeds
	Charcoal coat color	ASIP	G >T	****	Bengal
		ASIP	T >C	****	Bengal
		ASIP	A >G	****	Bengal
		ASIP	A >T	****	Bengal
		ASIP	T >C	****	Bengal
		ASIP	C >T	****	Bengal
		ASIP	A >G	****	Bengal
		ASIP	del (CA)	****	Bengal
	Chocolate (dark brown) coat color	TYRP!	G >A	*	No specific breeds
		TYRP!	C >G	*	No specific breeds
	Black coat color (melanism)	ASIP	del (CA)	****	No specific breeds
	Blotched tabby coat color	LVRN	G >A	*	No specific breeds
		LVRN	A >G	*	No specific breeds
		LVRN	C >A	*	No specific breeds

Condition type	Condition	Gene	Mutation	Evidence grading	Seen in
Coat Color and Pattern	Mackerel tabby coat color	LVRN	A > G	*	No specific breeds
		LVRN	C >A	*	No specific breeds
	Dilute coat color	MLPH	del (T)	****	No specific breeds
	Albinism	TYR	del (C)	***	Eastern breeds
		TYR	G >A	*	Eastern breeds
	Siamese coat color (pointed)	TYR	C >T	****	Siamese, Birman, Himalayan
		TYR	del (C)	*	Eastern breeds
		TYR	G >A	*	Eastern breeds
	Burmese coat color	TYR	C >A	**	Burmese, Singapura
		TYR	del (C)	*	Eastern breeds
		TYR	G >A	*	Eastern breeds
Coat Texture	Lykoi coat	HR	dup (GT)	****	Lykoi
		HR	ins (GACA)	****	Lykoi
		HR	compound	****	Lykoi
		HR	G >A	****	Lykoi
		HR	C >T	****	Lykoi
	Curly coat (KRT71-related)	KRT71	C >G	****	Selkirk rex
		KRT71	compound	***	Devon rex, Selkirk rex
		KRT71	C >T	*	Selkirk rex
	Sphynx coat	KRT71	C >T	****	Sphynx, Kohana
		KRT71	compound	*	Sphynx, Kohana
	Hypotrichosis (with short life expectancy)	FOXN1	del (ACAG)	***	Birman
	Curly coat (LPAR6-related)	LPAR6	del (GTTT)	****	Cornish rex, German rex

Condition type	Condition	Gene	Mutation	Evidence grading	Seen in
Coat Length	Long-haired coat	FGF5	A >T	*	Norwegian Forest Cat
		FGF5	G >T	*	Maine Coon, Persian, Ragdoll, Somali, cross-bred cats
		FGF5	T >G	**	No specific breeds
		FGF5	ins (A)	**	Ragdoll
		FGF5	G >A	**	Norwegian Forest Cat
Body Morphology	Osteochondrodysplasia and folded ears	TRPV4	C >T	****	Siamese, Birman, Himalayan
	Short tail	T	del (A)	****	Eastern breeds
		T	del (G)	***	Eastern breeds
		T	del (G)	*	Burmese, Singapura
	Short and kinked tail	HES7	A >G	***	Eastern breeds
	Polydactyly	ZRS	C >G	**	Eastern breeds
		ZRS	T >C	**	Lykoi
		ZRS	T >A	**	Lykoi
Susceptibility to Viral Infection	Likely resistance to FIV	APOBEC3Z3	compound	*	No specific breeds

basepaws

Sample Preparation and Sequencing

Genetic marker tests require a sample from the body, such as blood, urine, saliva, tissue, hair, or bone marrow.[11] At Basepaws, we work with buccal swab samples collected from the cat's mouth, at least one hour after food or drink. The samples are transported in DNA storage liquid to prevent DNA degradation that may result from temperature fluctuations during transport. Once the sample is delivered to the laboratory, the DNA from the swab is extracted and quantified. Next, a multiplex PCR with primers for our 115 genetic markers is performed on the DNA sample. In a subsequent round of PCR, adaptor and barcode index sequences are added to each sample in order to make sequencing on an Illumina platform possible. Finally, individually-barcoded samples from hundreds of cats are pooled together, and a battery of quality checks are performed on the pool by Basepaws and by our sequencing partner. If the sequencing library passes the quality checks, the samples are sequenced **(Figure 2)**.

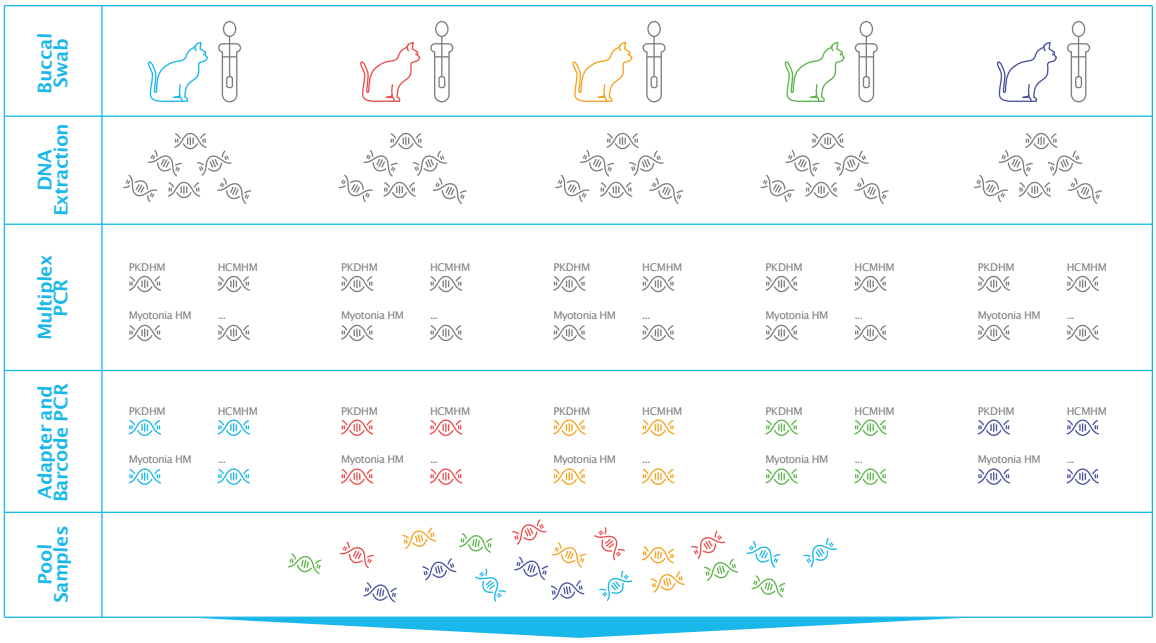


Figure 2. Sample preparation and sequencing process.

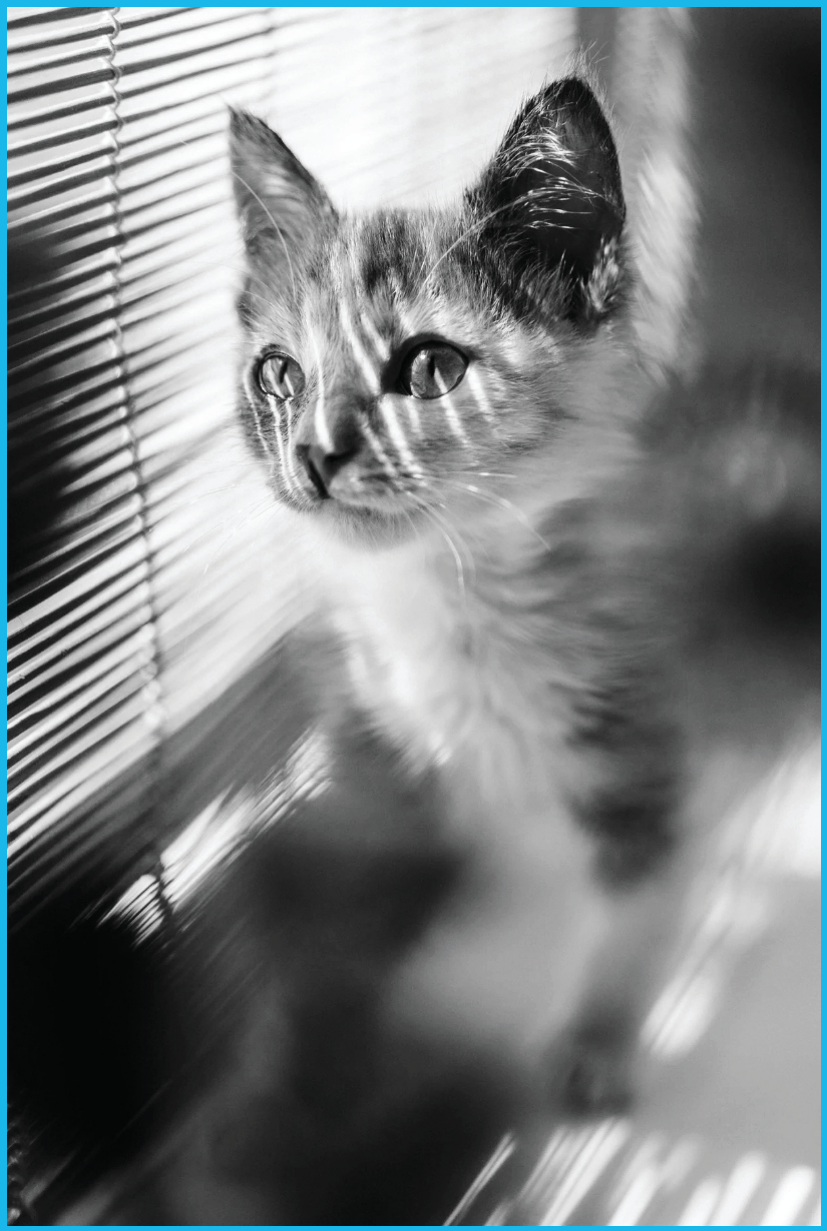
Barcode Hopping Mitigation Strategy

In order to reduce the already low probability of barcode hopping, we follow Illumina's guidelines[16] and:

- Remove free adapters and adapter dimers from our libraries prior to sequencing
- Use unique dual indexing combinations for our individual libraries
- Store libraries from each cat individually and only pool them prior to sequencing
- Buy an entire sequencing lane for our samples, avoiding the possibility of third-party low-complexity libraries mixing with our samples and increasing the chance of barcode hopping

basepaws

Analysis and Interpretation



Analysis Steps

Once the sequencing run is completed, we obtain raw sequencing reads data in a text format (FASTQ files). Our FASTQ file analysis pipeline performs the following steps:

- Assignment of barcoded sequencing reads to their corresponding original samples
- Removal of low quality sequencing data (usually, less than 2% of reads)
- Mapping/aligning of sequencing reads to the known genomic loci for our 115 genetic markers –this step allows us to see discrepancies between the sequencing reads and the known genomic loci and identify mutations
- Determination of sample zygosity
- Spot-checking a fraction of our results with Sanger sequencing as a final quality check

PCR Bias Mitigation Strategy

Like the golden standard for health marker testing (Sanger sequencing), NGS also suffers from potential PCR amplification bias. This happens when the polymerase incorporates the wrong nucleotide during template elongation and the error gets amplified in subsequent PCR cycles, resulting in the illusion of the presence of another allele.[15] To mitigate the effect of PCR-associated sequencing errors:

- We use high-fidelity Taq polymerase in our PCR reactions (high-fidelity Taq is optimized for accurate amplification, introducing minimal errors).
- We use a genotype quality score (GQ) >60 as a threshold. This typically means at least 30 reads cover the genomic region of interest.
- For all of our markers, we make sure that we obtain both forward and reverse sequencing reads covering the region of interest, which further mitigates the effect of PCR bias and ensures correct interpretation.

With our current procedure, the error rate is approximately 0.37% for SNPs. This means that for every ~270 reads mapping to the same genomic locus, one read has an error. This error rate was generated by observing the frequency of reads showing false alternate alleles at a homozygous locus. Aiming for >30X read coverage per genomic locus of interest makes the identification of errors easy, allowing us to remove erroneous reads prior to further analyses. On average, we get 60,897 sequencing reads covering each locus, which makes sequencing errors impossible to miss.

Determination of Zygosity

An ideal example of a heterozygous sample would be one where 50% of the sequencing reads map to one allele of a fictional Gene 1 and the remaining 50% map to another allele of Gene 1. **Figure 3** illustrates this concept with a real example from our data for the C>A mutation in the PKD1 gene.

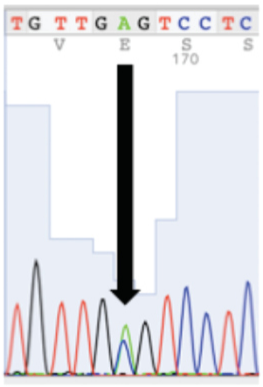
Known Genomic Locus – PLD1; Mutation of Interest: C>A

TGTTGCGTCCTC

Mapped Illumina Sequencing Reads for Sample Z (30X)

TGTTGAGTCCTC
TGTTGAGTCCTC
TGTTGAGTCCTC
TGTTGAGTCCTC
TGTTGAGTCCTC
TGTTGCGTCCTC
TGTTGCGTCCTC
TGTTGCGTCCTC
TGTTGCGTCCTC
TGTTGCGTCCTC

Sanger Sequencing Quality Check for Sample Z



- Allele A : Allele C = 1:1 in Illumina reads
- Sanger sequencing shows a double peak for A and C

Conclusion:
Sample Z is heterozygous for the C>A mutation in PKD1

Figure 3. An ideal example of heterozygosity. Note that the figure shows only the region around the mutation of interest and Illumina reads coverage has been abbreviated for illustration purposes (original coverage in this sample was >30X).

Unfortunately, NGS sequencing data does not always result in a perfect split in heterozygous allele coverage (mostly due to PCR bias). This can potentially result in instances where allele A of Gene 1 makes up 70% of all the sequencing reads covering Gene 1's genomic locus, while allele B of Gene 1 comprises only 30% of the reads. Such situations can lead to ambiguity in interpretation, with two potential ground truths:

- The sample is homozygous for allele A; allele B reads are simply a byproduct of propagated PCR error.
- The sample is heterozygous and has one copy of allele A and one copy of allele B; however, allele A was preferentially amplified in the PCR reaction.

In order to perform a stringent analysis and determine unequivocally whether the sample is homozygous or heterozygous, we calculate a likelihood for each possible diplotype according to alleles observed. Then we use GATK's[2] Genotype Quality calculation to obtain the best genotype and produce a GQ score. A cut-off of GQ >60 is then used as a threshold. According to our data, a GQ>60 usually means at least a depth of 30 reads.

Due to the high read coverage we aim to obtain for each genomic region of interest, samples with an 'undetermined' zygosity status are extremely rare. For those samples, we perform another NGS run and, if the results are still inconclusive, we exclude them from our customer report.

Report Results Interpretation: Health Markers

As noted above, in our report we only include results where we can conclude beyond doubt whether the sample is homozygous or heterozygous for a particular allele. There are four health marker status designations that we use:

Clear – The cat is negative (i.e., has zero copies) for all of the markers for which we tested that are known to be associated with a particular disease. This result, however, should not rule out the need to seek a professional diagnosis of this disease by a veterinarian, should the cat develop symptoms. It is still possible that the cat is positive for markers that are yet to be discovered that could be associated with a disease, or in some cases, environmental factors could contribute to a cat's potential to develop the disease.

Carrier – In most cases, this means that the cat has one copy of an autosomal recessive disease–associated marker. For diseases with a recessive inheritance pattern, the cat will develop the disease only if it has two two copies of the marker. Alternatively, the 'Carrier' result may mean that a female cat has one copy of a marker associated with a disease that follows an X–linked inheritance pattern. As a carrier, your cat is not at risk for developing the disease, but its offspring may be at risk.

At Risk – The cat has one copy of a marker for a disease where there is a dosage component to disease presentation (e.g., Hypertrophic Cardiomyopathy, Mucopolysaccharidosis VI, Hyperlipoproteinaemia). Gene dosage refers to the number of copies of an allele present in a cat's genome. A higher gene dosage results in an increased expression of a disease. This means that having one copy of the disease–associated marker could result in a milder disease presentation, as compared to having two copies where the disease presentation may be more severe.

At High Risk – This designation indicates that there is a very strong chance that the cat will develop the disease. It can mean one of four things: (a) the cat has one or two copies of a marker associated with an autosomal dominant disease (a single copy is enough to cause the disease); (b) the cat has two copies of a marker associated with an autosomal recessive disease; (c) the cat has two copies of a marker associated with a disease, where having these two copies results in a more severe disease presentation than if there was only one copy (i.e., the gene dosage effect); (d) a male cat has one copy of a disease with an X–linked (sex–linked and therefore non–autosomal) inheritance pattern (i.e., they are affected because they only have a single copy of the X chromosome that carries the mutation).

If a cat is classified as 'at risk' or 'at high risk' for a particular disease, we strongly encourage veterinarians to perform additional physiological exams prior to developing a treatment plan.

Report Results Interpretation: Trait Markers

We use slightly different marker status designations for our trait marker panel. See below:

Carrier – The cat has one copy of a marker associated with a specific physical trait, however, it is unlikely to be physically manifesting this trait. This could be because the trait has an autosomal recessive pattern of inheritance (i.e., needs two copies to manifest physically) or because the physical presentation of the trait is associated with a specific combination of markers, of which your cat only has one.

Likely to have – The cat is positive for a marker (or markers) linked to a specific trait, and is likely to exhibit this trait. This could be a result of the cat having one copy of a trait marker with an autosomal dominant pattern of inheritance, or the cat having two copies of a marker with an autosomal recessive pattern of inheritance. Alternatively, your cat could have the specific allelic series (combination of markers) that are likely to result in a specific trait.

Not likely to have – Based on the cat's genotype, it is unlikely that it is exhibiting this particular trait.

Report Results Interpretation: Markers for Blood Type

There have been seven genetic variants with association to blood type reported in the literature. However, due to some discrepancies in the available literature, we only report on four markers that have the most robust association to blood type.[20,21] Cats that have zero or one copy of mutations G>T, T>A or del(T) in the gene CMAH are type A, while those that have two copies of these mutations are type B. Cats that have two copies of the C>T mutation or have one copy of the C>T mutation in combination with one of the other three mutations characteristic of type B, are blood type AB.

The Basepaws report also provides blood transfusion risk assessment (low, medium or high), based on the cat's blood type.

Due to our still evolving understanding of the role of genetics in feline blood type determination, these results should be interpreted as highly likely rather than 100% definitive. If your cat has a life-threatening emergency where a blood transfusion is required, we strongly advise you to test your cat at your vet's office using an immunochemical assay.

About Basepaws

Basepaws is a pet health company that specialises in genetics, genomics, and microbiome research. In 2018,we launched the world's first at-home consumer cat DNA test for delivering actionable health-and breed-related insights to pet owners. We are committed to providing a cutting-edge service and continue to update our health and trait marker panel as new discoveries are made in the field of feline genetics. Basepaws is actively advancing feline genetics with its own internal research projects as well as through collaborations with animal pharmaceutical companies. Our mission is to improve the health and wellbeing of every pet.

References

1)NIH US National Library of Medicine –Genetics Home Reference: https://ghr.nlm.nih.gov/primer/genomicresearch/snp

2) College of Veterinary Medicine (University of Missouri):
http://felinegenetics.missouri.edu/feline-genome-project-2/cat-genomic-resources-strs-snps

3) Aitman TJ. DNA microarrays in medical practice. BMJ. 2001; 323(7313):611-615.doi:10.1136/bmj.323.7313.611

4) Genetic Science Learning Center (University of Utah): https://learn.genetics.utah.edu/content/precision/snips/

5) Lehman DJ et al. Large Meta-Analysis Establishes the ACE Insertion-Deletion Polymorphism as a Marker of Alzheimer's Disease. American
Journal of Epidemiology. 2005; 162(4):305-317

6) Clancy S & Shaw K. DNA deletion and duplication and the associated genetic disorders. Nature Education. 2008; 1(1):23

7) NIH US National Library of Medicine –Genetics Home Reference:
https://ghr.nlm.nih.gov/condition/polycystic-kidney-disease

8) NIH US National Library of Medicine –Genetics Home Reference:
https://ghr.nlm.nih.gov/condition/myotonia-congenita#inheritance

9) NIH US National Library of Medicine –Genetics Home Reference:
https://ghr.nlm.nih.gov/condition/hemophilia#inheritance

10)Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet. 2018; 50(9):1219-1224.doi:10.1038/s41588-018-0183-z

11)American Association for Clinical Chemistry: https://labtestsonline.org/genetic-testing-techniques

12)NIH National Human Genome Research Institute –DNA Microarray Technology Fact Sheet:
https://www.genome.gov/about-genomics/fact-sheets/DNA-Microarray-Technology

13)European Bioinformatics Institute –Limitations of Microarrays:
https://www.ebi.ac.uk/training/online/course/functional-genomics-ii-common-technologies-and-data-analysis-methods/limitations-microarrays

14)Yee AJ & Ramaswamy S. DNA microarrays in biological discovery and patient care. Essentials of Genomic and Personalized Medicine.
Academic Press. 2010. 73-88

15)Aird D et al. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. 2011;12(2):R18

16)Illumina –Minimize Index Hopping in Multiplexed Runs:
https://www.illumina.com/science/education/minimizing-index-hopping.html

17)The Gene Pool (University of Edinburgh): Sanger Sequencing Troubleshooting Guide:
http://genepool.bio.ed.ac.uk/sanger/Sanger_troubleshooting_guide_v1.pdf

18)Online Mendelian Inheritance in Animals (University of Sydney): https://omia.org/home/

19)Meurs KM et al. A cardiac myosin binding protein C mutation in the Maine Coon cat with familial hypertrophic cardiomyopathy. Human Molecular Genetics. 2005; 14(23):3587-3593

20) Gandolfi, B., Grahn, R.A., Gustafson, N.A., Proverbio, D., Spada, E., Adhikari, B., Cheng, J., Andrews, G., Lyons, L.A. and Helps, C.R., 2016.A novel variant in CMAH is associated with blood type AB in Ragdoll cats. PloS one, 11(5), p.e0154973.

21)Kehl, A., Heimberger, K., Langbein-Detsch, I., Boehmer, S., Raj, K., Mueller, E. and Giger, U., 2018. Molecular characterization of blood type A, B, and C (AB) in domestic cats and a CMAH genotyping scheme. PloS one, 13(9), p.e0204287.

22) https://gatk.broadinstitute.org/hc/en-us

basepaws

Thank you
for your attention