

Pioneering feline genetics with next-generation sequencing

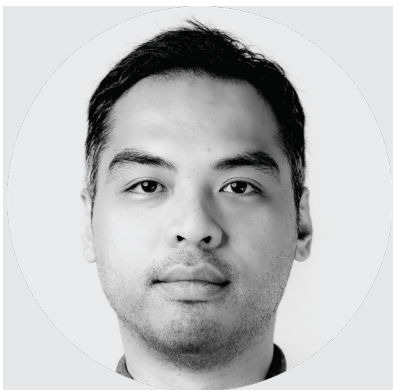
Basepaws leverages the Illumina DRAGEN™ Bio-IT Platform to develop direct-to-consumer DNA kits that provide cat owners with insights about their pets' health



ANNA SKAYA, FOUNDER AND CHIEF EXECUTIVE OFFICER, BASEPAWS

There are cat people and there are dog people. That divide even spills over into the direct-to-consumer (DTC) DNA testing market. Many products are available to help dog owners access genetic data about their pets. But cat owners have been left in the dark. Anna Skaya, founder and CEO of Basepaws, loves cats and saw that discrepancy as an opportunity.

Basepaws is a pioneering DTC company that has developed unique feline genetic test kits. Through these tests, Basepaws seeks to connect people with their pets and improve cat health. They are building the world's largest feline genomics database, currently containing tens of thousands of cat genomes. With their ever-growing genomic database, they are constantly learning new information about cat breeds, traits, and health.



DAMIAN KAO, PHD, CHIEF OPERATING OFFICER, BASEPAWS

Whole-genome sequencing for pets

Most DTC DNA testing companies use Illumina microarrays to read DNA. While dog DNA testing companies have taken advantage of canine microarrays, no feline microarray product was available. "It was a really difficult problem to solve and nobody was touching it," Skaya recalls. Basepaws decided to use next-generation sequencing (NGS) to develop genetic testing products for cats.

"Until a few years ago, NGS was too expensive to use for cats and dogs. Basepaws hit that perfect time point in 2018. When we came on the scene, Illumina had new systems that allowed us to do sequencing at scale and at a cost that was affordable to the pet parent. Our \$129 product includes 1x genome coverage and targeted amplicon sequencing, which allows us to perform genomic breed similarity analysis and screen for genetic markers associated with a variety of diseases and traits.

"In our hands, using DRAGEN analysis, we have realized cost savings of approximately 50%. In terms of computational time it's almost double the speed compared to the manual method."

"The breed part identifies the parts of your pet's DNA that are most similar to known breeds. For example, it could be more Siamese or more Persian. It's fun to know. But the most insightful part has to do with health markers. We test for 65 different health markers, as well as 50 markers associated with physical traits and blood type. Results from the health marker screening can potentially identify disease-associated genes. We also use shotgun metagenomic sequencing and oral microbiome analysis to identify microbes associated with dental disease. With all its versatile uses, we believe that NGS is a much better technology for our needs."

Starting from scratch

Damian Kao, PhD, is COO and Head of Science at Basepaws. He explains that the biggest challenge was not having enough publicly available feline genomic data. "Canine research is probably 10 times larger than feline research. We had to do a lot of our own sequencing to build up our database before we could actually start producing results for our customers.

"With Illumina NGS technology, we're able to scan the entire genome. As our database grows and we sequence more cats, we can start making genetic associations with diseases that our customers' cats might have. We will be able to research new markers that will ultimately help pet parents everywhere."

Managing genome-scale data

The next challenge, Dr Kao said, was how to manage whole-genome sequencing (WGS) data analysis. "The data sets that we get back from Illumina sequencing machines are large, a lot larger than with your traditional microarrays. That makes the analysis process much slower and more costly. A big challenge is trying to cut down that cost and make things faster."

The Illumina DRAGEN Bio-IT Platform directly addresses that challenge. The DRAGEN platform uses reconfigurable field-programmable gate array (FPGA) technology to deliver accurate, comprehensive, and efficient secondary analysis of NGS data. DRAGEN Original Read Archive (ORA) is able to compress WGS data without loss. "Since the DRAGEN platform is built on this FPGA architecture, it is optimized for WGS data processing."

Dr Kao explains that the traditional method to process WGS data involved a four-step open-source software pipeline. The DRAGEN platform consolidates all those steps into one process. "The steps that the DRAGEN platform takes over would be four separate steps. With each step you have to worry about error handling, quality check, and random errors that might happen.

The DRAGEN platform takes care of all those little edge cases and gives out very detailed statistics and metrics that help us check quality.”

The benefits of DRAGEN analysis

“The three biggest advantages of the DRAGEN platform for us are speed, cost, and reliability,” says Dr Kao. “Speed and cost are pretty self-explanatory. Reliability addresses complexity. Because DRAGEN analysis is just one single step, it makes our process easier to run.” During promotions on kits, Basepaws can get 10,000 samples back at once, for example. “It quickly becomes an exercise in how you juggle multiple processes. The DRAGEN platform actually takes care of the most complicated part.”

Using the DRAGEN platform, Dr Kao says that Basepaws has minimized the time it takes to return reports to customers. “In our hands, using DRAGEN analysis, we have realized cost savings of approximately 50%. In terms of computational time it’s almost double the speed compared to the manual method. As soon as we get the data back from a sequencing facility, we can immediately process it with the DRAGEN platform and have the report out to our customers in less than a day, where it used to take two or three days. That’s significant.”

According to Dr Kao, “One huge advantage of sequencing technology is that it is organism agnostic. It doesn’t have to be human. It doesn’t have to be dog. Every animal on this planet has DNA. If you can sequence it, you can use DRAGEN analysis. We plugged the DRAGEN platform directly into our pipeline without much effort. The only thing we had to do was insert the feline reference genome and just start using it. It was as simple as that.”

One unanticipated benefit is the ability to use repetitive region data. “The DRAGEN platform deals with repetitive regions of the genome very well. Repetitive regions are usually difficult to deal with using open-source software, so we used to throw out a lot of data.” Dr Kao explains, “The repetitive region data can provide us with genotypic resolution, which actually informs us a lot about the breed of the cats. We were never able to utilize it effectively before, but with the DRAGEN platform, now we’re able to.”

Scaling for rapid growth

The volume of samples Basepaws has received for the past four years has doubled or tripled each year. Dr Kao explains, “As you scale a business that depends on sequencing, there are several bottlenecks, including engineering and data processing. We are constantly thinking: How long does it take to process all this data?

“With the DRAGEN platform, we have the report out to our customers in less than a day, where it used to take two or three days. That’s significant.”

“The repetitive region data actually informs us a lot about the breed of the cats. We were never able to utilize it effectively before, but with the DRAGEN platform, now we’re able to.”

How do we organize it? How do we store it? Data will never go away, but it's something that's easily solvable by using the DRAGEN platform."

Basepaws uses Amazon Web Services to run their genomics pipelines and access the DRAGEN platform. Dr Kao says, "The main advantage for a small company like ours is that you can just rent space on their cloud platform, without having to invest in all that hardware. You literally just have to requisition more servers on the Amazon cloud and you can scale."

Looking to the future

Basepaws has a vision to improve the lives of cats around the world by understanding, genetically, what makes each cat unique. Skaya says, "Basepaws has been growing year over year, phenomenally well. Sequencing cats was a great starting point, but it's definitely not the end. We're very interested in using our technology to go after the dog market as well. We're very proud of the fact that we are a sequencing-based company. It allows us to see a lot more detail. It allows us create a much better product.

"We really love our pets and we've got to take care of them. Our test is not just a DNA test, it's not just for fun. Your cat can't talk, but their DNA and their microbiome actually can give you answers that otherwise you wouldn't know."

Learn more

DRAGEN Bio-IT Platform, illumina.com/products/by-type/informatics-products/dragen-bio-it-platform.html

Basepaws, basepaws.com

illumina[®]

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel | techsupport@illumina.com | www.illumina.com

© 2022 Illumina, Inc. All rights reserved. All trademarks are the property of Illumina, Inc. or their respective owners.

For specific trademark information, see www.illumina.com/company/legal.html.

M-GL-00575 v1.0

"One huge advantage of sequencing technology is that it is organism agnostic. Every animal on this planet has DNA. If you can sequence it, you can use DRAGEN analysis."