Feline Social Media Star Gets Sequenced

Researchers from the Max Planck Institute and the University of Pennsylvania crowdfunded the genome sequencing of Lil Bub, using the NextSeq® 500 System to uncover the genetic causes of the cat's rare diseases.

Introduction

With more than 2 million Facebook friends and almost 900 thousand followers on Instagram, the tiny cat from Indiana known as Lil Bub[1] has a plush toy of her likeness and has been featured on more television shows than she can count on her polydactyly toes. There's not much this feline hasn't done.

Researchers from the Max Planck Institute for Molecular Genetics in Berlin and the University of Pennsylvania are bringing Lil Bub into another spotlight—the one of genomic research. They've crowdfunded a project to sequence her genome with the NextSeq 500 System and examine the potential causes of her polydactyly and what appears to be osteopetrosis, a disease that hardens bones. Filming each step of the project, Drs. Dario Garcia Lupiañez, Daniel Ibrahim (at Max Planck in Berlin), and Uschi Symmons (formerly of the ENS de Lyon and now at the University of Pennsylvania) are engaging Lil Bub's fans and educating the general public about what it takes to perform genomic research. They're also expanding their rare disease research studies in the hopes of identifying causes of similar diseases in humans.

iCommunity spoke with the team to learn about what inspired them to initiate this project on their own time, the challenges and benefits of funding the Lil Bub research, and how they hope to change the public perception of science.

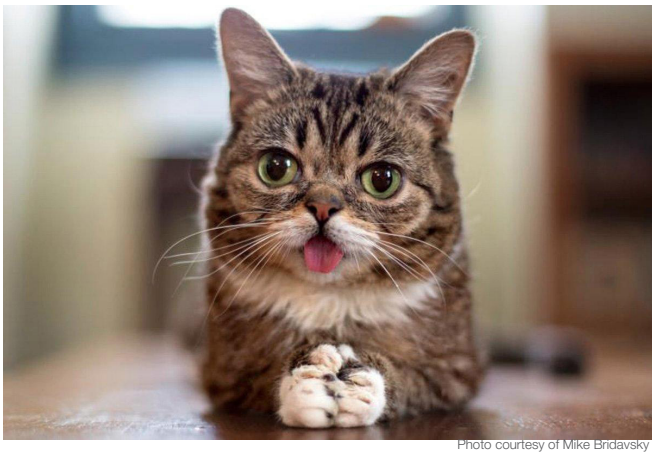


Photo courtesy of Mike Bridavsky

Lil Bub was born to a feral mother in 2011. She was the runt of the litter and possesses several genetic anomalies.

Q: What sparked your interest in studying the genetics behind rare diseases?

Daniel Ibrahim (DI): In some rare diseases, we see a direct genotype/phenotype link, with a mutation causing a visible, nonlife threatening effect. That enables us to learn a lot about how the genome works, how genetics can cause certain diseases, and decipher the genetic mechanisms responsible for certain rare diseases. The genetic data enables us to really address the cause of the rare disease, rather than being limited to analyzing symptoms.

Dario Garcia Lupiañez (DGL): It's exciting when we can study these rare diseases in other species. Most diseases are caused by the same genes no matter what the species, so this research can make a connection not only to human disease, but to evolution.

Uschi Symmons (US): As a basic researcher, rare diseases are really an amazing way to get into genetics. Studying the genetic links associated with rare disease allows us to actually help people.

Drs. Dario Garcia Lupiañez and Daniel Ibrahim are scientists at the Max Planck Institute for Molecular Genetics in Berlin, Germany and Dr. Uschi Symmons is a scientist at the University of Pennsylvania.

Q: How did you first hear about Lil Bub and what sparked your interest in studying her?
DI: About 2 years ago, we saw a documentary about Lil Bub by chance. When I saw her, I couldn't help but see the similarity with rare disease patients that we were studying in our lab. As it turned out, Lil Bub is suffering from a rare bone disease.

DGL: We thought that this case would be biologically interesting. We also knew that many people cared about Lil Bub. We thought they might be as interested as we were in knowing the genetics underlying her condition and if there were any treatment options available.

> "The feline colony still exists and it so happens that Lil Bub shares a mutation with the Hemingway cats."

Q: Describe the phenotypic issues that Lil Bub possesses?
US: She has polydactyly, which in her case means that she has an additional toe on each paw. She also has been diagnosed with something similar to osteopetrosis, which means that her bones are denser than they should be. This is why she is so petite. It's also responsible for her very small lower jaw, which is why her tongue is always hanging out. She's also toothless. We think that's because her bones grew too densely and quickly, and her teeth didn't have time to form completely.

Without looking at the genetic data, we don't know if Lil Bub is exhibiting 1 or 2 diseases. No one has ever described this complete spectrum of phenotypes. It could be that her polydactyly is a rare manifestation of osteopetrosis, or she might have 2 or more rare diseases.

Q: What developmental factors do you think would be responsible for these issues?
DI: We already identified a strong genetic candidate for her polydactyly. It's a mutation in the regulatory element of the *Sonic hedgehog* (*Shh*) gene that causes misexpression of the gene and has been described in other polydactyl cats. As for her bone phenotypes, we don't really know. I think it's a protein coding mutation, because all her bones show this effect.

US: What was amazing to us was that the *Shh* enhancer mutation has been described previously in a colony of polydactyly cats that was started by Ernest Hemingway on his estate in Key West. The feline colony still exists and it so happens that Lil Bub shares a mutation with the Hemingway cats. Lil Bub is an orphan cat, so it was interesting to give her a potential, very distant family association.

At the same time, it's curious that she possesses a polydactyly phenotype that isn't exactly the same as the Hemingway cats. Most Hemingway cats only exhibit polydactyly in their front paws, while Lil Bub exhibits it in all 4 paws. We might have found the driver mutation for the polydactyly, but by sequencing the entire genome we might find other mutations that influence the phenotype.

Q: Are there any similarities between Lil Bub's disorders and certain rare human developmental diseases?
DI: Even though the doctors diagnosed osteopetrosis in Lil Bub, it's not definitive because her bones don't look like bones with that condition. Rather, she has what are called Erlenmeyer flask bones, which are long and become thicker toward her paws. There are human patients with these bones who have mutations in a calcium channel, so that could be a very strong candidate for us to look at after we have the genome sequenced.

Q: Are there currently any therapies for these diseases?
DI: Osteopetrosis is a known rare disease and there are several therapies that mainly affect the osteoclasts. However, there are several different subtypes of osteopetrosis, and a therapy for one subtype does not always work for another. In Lil Bub's case, electro pulse therapy worked quickly after her disease became more severe. It has helped her to lead a rather normal life—as normal as it can be for an Internet sensation!

Q: Is there a precedent for sequencing animals with certain disorders to understand rare human diseases?
DI: One of our collaborators, Dr. Leslie Lyons at the University of Missouri, is doing this in a much bigger and more systematic way with her 99 Lives Project.[2] Because of the inbreeding found in cat breeds, feline genetic research is accessible. For a number of diseases, there is a connection between the phenotype or disease in cats and humans.

US: Our Lil Bub project is different from the 99 Lives project because we're sequencing an individual cat. In an era where everyone is talking about personalized medicine, we think Lil Bub's case resembles that of human sporadic patients where you don't always have access to family member genotypes. We hope that sequencing Lil Bub's genome, pinpointing the mutations that cause her phenotype, and describing that process will be valuable in teaching people how biologists use genetics to understand what causes disease.

> "In an era where everyone is talking about personalized medicine, we think Lil Bub's case resembles that of human sporadic patients where you don't always have access to family member genotypes."

Q: What was the crowdfunding experience like?
DGL: When we initiated the crowdfunding effort, we didn't know what to expect. We were pleasantly surprised to find that the people investing their money in the project had the same motivation that we have. They wanted to help us understand the genetics behind Lil Bub's maladies. Some people are well informed about what we are doing and ask us specific technical questions. It's a super rewarding project to work on.

DI: The project is fun, and is completely different from anything else that we do. The research itself isn't that much extra work, but the communication and interaction are. What is different is that we're using video to share what we're doing with our backers.[3] We want to communicate our work and explain our ideas as we progress without having to wait 2 years for a project to be finished and a paper to be published. This is not peer-reviewed work yet, so it is as free floating as science can get.

US: We feel a tremendous amount of responsibility because we know that people have put their faith and their money in us to perform the Lil Bub sequencing project. We feel obliged to share our progress with them.

> "If you need to sequence one mammalian genome quickly, the NextSeq 500 System is the best option. You can prepare everything at once and don't have to wait for other samples to perform a run."

Q: What are the benefits of communicating a project like this to a nonscience audience?
DLG: The biggest enemy that science faces is not being able to translate the science into something that the general public can understand. This project provides the perfect opportunity to do it right. We think we can capture people's imagination with this project and explain to them the value of genetics.

US: We know we're reaching people with the Lil Bub project who care about Lil Bub. It doesn't matter if they understand science. We can explain the science. When we get positive feedback telling us whether we're explaining things well, it is really motivating.

Q: What is it about genomic science in particular that is so important to explain to the public?
DI: Some people think that genome sequencing suddenly elucidates everything about your future medical conditions. That's not the case. It's important to communicate what sequencing can do and what it can't do, and how it can be used optimally. By performing whole-genome sequencing of Lil Bub, we will likely find thousands of mutations. That will give us the opportunity to explain that all of us possess thousands of mutations. It's what makes us unique. In studying disease, it's necessary to evaluate whether these mutations are causative for the disease or not. These are not difficult things to comprehend, but sometimes you need a really good example to explain simple things. We think the Lil Bub project does that.

DLG: In the future when genomic sequencing becomes more commonplace, it will be important that people understand the implications and how that information should be shared.

Q: Which library prep kits are you using?
DI: We used the TruSeq® DNA PCR-Free Library Preparation Kit. We prepared one 300 bp and one 500 bp library and ran them on the NextSeq 500 System. Having alternate size libraries can improve the subsequent data analysis and the identification of mutations. So we ended up with paired-end 150 bp and used the whole flow cell for it. We'll probably get 40x coverage.

Q: Why did you decide to use the NextSeq 500 System?
DGL: We had used the NextSeq 500 System before and knew that it was fast and offered a good solution for sequenicng a single genome. When something works, you don't need to change it.

US: Daniel and Dario's lab has a lot of experience with sequencing human samples. This being novel territory for us, we wanted to use a method that was well established just to make sure we'd have as few technical problems as possible and could really focus on the science and analysis.

DI: If you need to sequence a mammalian genome quickly, the NextSeq 500 System is a great option. You can prepare everything at once and don't have to wait for other samples to perform a run. Most of our sequencing is performed at a sequencing facility and you never know how long you will have to wait for a slot to open up.

US: The human factor is what is slowing everything down now in our project, because we want to document things and make sure that we perform a proper analysis. It's a nice added benefit to be able to sequence rapidly because it means that we're not adding to our own human slowness.

> "We are comparing Lil Bub's genome with the reference cat genome, and are filtering through millions of SNPs to identify the ones that might be associated with disease."

Q: How important is accuracy to your project and how did that factor into your decision to use the NextSeq 500 System?
DI: The cat genome is complete, but it's not the best annotated and the official sequence is based on a single individual. We have to be sure that whatever mutations we find are true and not sequencing artifacts. That's one reason why we chose to perform deep sequencing on the NextSeq 500 System. Paired-end 150 bp is something that also contributes to high accuracy.

Q: What results have you seen so far?
DI: We just received the sequencing results, and are now in the process of analyzing the data using our inhouse pipelines. So far, we have confirmed that a mutation at the *Shh* enhancer is causing the polydactyly. We are comparing Lil Bub's genome with the reference cat genome, and are filtering through millions of SNPs to

identify the ones that might be associated with disease. We have some interesting candidates, but we want to validate them and make sure that we don't overlook anything. This is not a trivial issue, so it might take some time.

Q: When you have the sequencing data for the LilBubome, what software programs will you use to analyze it?
DI: We will perform data analysis using custom-built bioinformatics software that our group has run on human genomes. The only difference is that we will use a different reference and, as mentioned earlier, will factor in publicly available cat genome data.

DLG: We will provide other researchers with the opportunity to analyze the data on their own. They're welcome to the data.

US: This is one of those cases where the more eyes you have on the data, the more mutations you might find. For example, we already have people following us who study osteopetrosis. They'll be able to bring their expertise to the project. We're excited about using this project as an opportunity to practice science in an open consortium.

> "It's great to be able to share this experience with all of Lil Bub's fans. It shows how science can connect people."

Q: Why is this project important enough to you that you're willing to contribute your own time and resources to it?
DI: This project is basic research at its core. It's an intriguing case that makes us curious, and we want to understand it. Hopefully this project will determine the mutation that causes Lil Bub's phenotype. However, there is no next step. We aren't saying, "If we find this out, then this and this will happen." We all would like to know the genetic causes of her maladies. We hope it will lead to identifying new therapeutic options for her and for humans that have similar diseases. At the same time, we have the chance to share some aspects of what geneticists do with the wider public.

DLG: The most frequent question we get on the project is, "why do you want to sequence a cat?" We're happy to explain that we're sequencing Lil Bub's genome because it's important and that we can learn something that can be applied.

US: We all got into science because we wanted to understand how the world around us works, and I think Lil Bub is an amazing example for this. Science is fun and it's linked to our everyday lives. It's great to be able to share this experience with all of Lil Bub's fans. It shows how science can connect people.

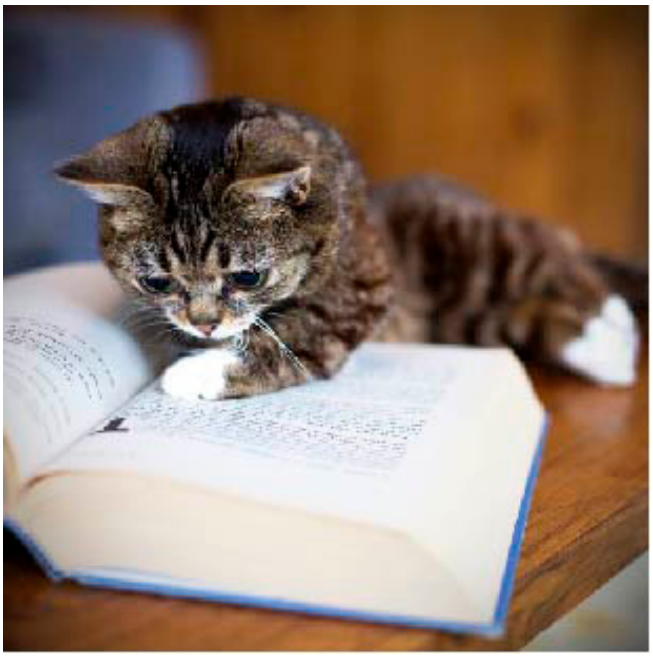
References:

1. Lil Bub. lilbub.com/. Accessed September 24, 2015.
2. 99 Lives Cat Genome Sequencing Initiative. felinegenetics.missouri.edu/99lives. Accessed September 24, 2015.
3. The LilBubome-Sequencing LilBub's Magical Genome. experiment.com/projects/the-lilbubome-sequencing-lilbub-s-magical-genome/labnotes. Accessed September 24, 2015.

Learn more about the Illumina technologies mentioned in this article:

NextSeq 500 System
www.illumina.com/systems/nextseq-sequencer.html

TruSeq DNA PCR-Free Library Preparation Kit
www.illumina.com/products/truseq-dna-pcr-free-library-prep-kits.html



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

illumina®