

Blog: The Cutting Edge

04 AUGUST 2022

Thirty years in genetics – A perspective from Brazil



Lygia Pereira

The course of academic research rarely runs smoothly, and through her 30-year career in genetics, Professor Lygia V. Pereira has certainly seen plenty of challenges – but also lots of successes. We talk to her about the importance of genetic diversity to better serve the health needs of countries like

Brazil but also to improve our understanding of disease and health across all populations, and why we need to make bioinformatics tools truly user-friendly. Can we start by talking about your career path – how did you get into genomics? My career path started off in a slightly unusual way, as I originally did a Bachelor's degree in Physics at the Pontifícia Universidade Católica in Rio de Janeiro – computer engineering was the latest thing, and I was keen to follow that route. But I had also really enjoyed genetics in high school, so while I was doing calculus and quantum mechanics, I kept reading about genetic engineering and molecular biology.

The more I read, the more interested I became, so towards the end of my physics course I asked if I could switch my course credits, and I started doing an internship in a molecular biology lab, studying the genetics of *Drosophila*. I loved that – the lab life, extracting DNA – and I realized that these tools could be used to study *any* living thing. That was when I saw that human genetics was the career for me.

And where did that internship lead to next?

Around that time, which was the late 1980s, the then Director of the Human Genetics Department at Mount Sinai medical Center in New York came to Brazil to give a talk about gene therapy. There was a lot of talk about how gene therapy was going to cure everything, and I was fascinated by that. I applied to Mount Sinai Graduate School, which at the time was affiliated to the City University of New York, and that meant that from 1989 to 1994 I was based in New York.

For my first rotation I worked on Niemann–Pick Disease, which involves a defect in lysosomal storage. Then in the second rotation I went to Francisco Ramirez’s lab, and at that time they had just found a little piece of the gene involved in Marfan syndrome, which affects the body’s connective tissues. I thought that was really interesting, so I stayed with that for my Ph.D. which involved cloning the full length of cDNA, determining the genomic structure of the gene, and ultimately developing a knockout mouse model for Marfan syndrome.

And then you returned to Brazil?

Yes, and despite my original intentions, I’ve never left! I was originally going to stay in Brazil for a couple of years and then go elsewhere for a postdoc, but in 1996 I ended up getting a four-year grant to work at the University of São Paulo, on establishing in the country the methodology to create knockout mice models, again for Marfan syndrome. The models I used with Ramirez’s group were great, but they had some limitations, so I worked on some different mutations. That work was eventually rewarded in 2001 when our group generated the first knockout mice here in Brazil. That was a proud moment – nationally it was a big contribution.

One of your passions has been improving the diversity of population diversity present in cell lines and biobanks. How did that come about?

That all arose when we developed the lines of human embryonic stem cells from blastocyst here in Brazil. We were thinking about compatibility, and asked

ourselves, are these cell lines a better match to the Brazilian population? We collected HLA (Human Leukocyte Antigen) profiles for different lines from elsewhere in the world, and we crossed those with our bone marrow registry, for which there was HLA typing on hundreds of thousands of Brazilians. And we found out that no, our line was not a better match!

To understand that we looked at the genomic ancestry of our lines, and we found that they had over 90% of European ancestry, not representative of the admixture of Indigenous, African and European genomic ancestries found in our general population. Our hypothesis was that the embryos were all donated in private clinics, and that the socioeconomic class that has access to those clinics isn't representative of the whole population.

As a result, we knew that if we wanted to capture the true genetic diversity of the Brazilian population, we needed to get out into the country and establish lines of induced pluripotent stem cells (iPS) instead. So that was what we did – we developed a collaboration with a longitudinal study of 15,000 Brazilians, who have been followed clinically for more than 10 years. And as expected, the genomic ancestry of these iPS cells were much more representative of the population – which of course should prove very useful when cell-based assays become robust enough to detect differential drug response between individuals, for example.

Did that lead on to your work on genetic diversity?

That's right. In about 2016–7, the lack of diversity in human genomic data started to gain a lot more attention, with an article that showed that about 80% of the genomic data that's out there related to populations of European descent. And we figured, if Brazil has anything to contribute to improving that figure, it's with our population diversity. In fact, since 2014 there had been much discussion about doing a Brazilian genome project, but this global discussion seemed to me to provide the push that was needed.

I joined forces with two colleagues, Alexandre Pereira and Tábita Hünemeier, who were also at the University of São Paulo, and we sought funding to put this project together. Our plan was to take the 15,000-person longitudinal study I mentioned earlier, and basically add genomic data to that, to kind of jumpstart the program. And it took quite a while, but eventually we got the Ministry of Health on our side,

and in the end, they were willing to fund sequencing of the whole cohort. We launched the project in December 2019 – it was my happiest Christmas ever!

And then Covid came along...

Indeed, it did, and priorities shifted completely. To cut a long story short, as I'm talking to you now in July 2022, we've sequenced 4000 genomes, and a contract should be on the way to do 6000 more. Our hope is that the program will now start to reaccelerate – we will get there!

What do you see as the ultimate benefits of this regional genomic database?

First and foremost, it should benefit the Brazilian population, because you should be able to develop polygenic risk scores that better reflect an admixed population like ours, rather than a European one. And at the same time, we may find novel variants that tell us things about human biology that can eventually be translated into new medicines or therapies. And this isn't just the case for Brazil of course – there's such a lot to be gained from looking at different genetic lineages across the globe.

Our proposal is that for the first two years the principal investigators from the cohort studies would have exclusive access to the data. After that, we'd like it to be accessible to Brazilian-led groups for two years, and only then would we open it to the rest of the world. That decision is of course in the hands of the government, but what I'd like to avoid is a situation where we generate the data and then the discoveries are made mainly by groups outside Brazil, solely because we lack the pace of research that comes when you have a critical mass of researchers and suppliers.

You raise an interesting point there, about who gets to make the discoveries based on genomic data. Why do you think that's so important?

Partly it's about fostering skills and collaboration in the country where the data was generated – our thinking in Brazil is that if groups here have exclusive access for a few years, that will kickstart capabilities to use that data, and hopefully promote the growth of a well-developed genomics and bioinformatics 'ecosystem' in this country.

But for me, it's in part a more fundamental principle too. When I was at school, I was taught that as an 'underdeveloped' country, we export primary goods, and then we import the manufactured goods. This may be a feature of the Brazilian economy in many areas, but I'm determined not to see it replicated for human genome work too. I see this issue in some of the African initiatives too – why should they go to the effort of producing the primary data, only to watch the developed world make the discoveries and bring the innovations that stem from it? We need to strike a balance that benefits all parties equally.

And speaking about processing the data, what are your thoughts on the tools that are now available?

Well, as I've mentioned I'm not really a bioinformatician, so I'll give you an end-user's perspective.

I think there are several challenges – having a platform that allows access globally while remaining secure and affordable is high on my wish-list! But also, I would like to see bioinformatics tools made more user-friendly than they are currently. I like to make an analogy with smartphone apps – they do some amazingly complex processes and offer a wealth of capabilities, but on the whole, they are a lot easier to use than many of the bioinformatics software platforms that are out there. Why should that be?

I think bioinformatics software companies need to think about the end-users a little bit more. They need to make sure their tools are not needlessly complex, and that they're structured to allow the most popular tasks to be done easily and quickly. That way, I think we could get more scientists – and especially those without a bioinformatics background – using the almost infinite amount of data that's out there. If they can do that, then I see potential to accelerate genomic discoveries dramatically.

Thank you for your time, Lygia – it's been very interesting to talk to you and hear your inspiring story and unique perspective.

It's clear to us that what we perceive as the scientific challenges in nations where genomics is well-funded are not necessarily the same faced by researchers in other countries. It's renewed our focus on making sure that our tools truly help genomic

scientists with their everyday tasks, and that we strive to ensure that they're easy to use by those who aren't experts in bioinformatics or coding.

For more information about our suite of tools for genomics research, contact us at lifesciences@paradigm4.com.

Author Bio

Professor Lygia V. Pereira is Director of the National Laboratory of Embryonic Stem Cells (LaNCE) in the Department of Genetics and Evolutionary Biology, at the University of São Paulo, Brazil. Her previous training was in Physics (Bachelor's degree, 1989) and Molecular Biology (M.Sc., 1990), followed by a Ph.D. in human molecular genetics from the City University of New York (1994). Professor Pereira's current research interests include studying animal models of Marfan Syndrome, X-chromosome inactivation, establishing new lines of human embryonic stem cells, and identifying human knockouts within the Brazilian Genome project.

References

1. Moreira de Mello, J.C., Fernandes, G.R., Vibranovski, M.D. *et al.* Early X chromosome inactivation during human preimplantation development revealed by single-cell RNA-sequencing. *Sci Rep* 7, 10794 (2017). <https://doi.org/10.1038/s41598-017-11044-z>

For more information about how our analytics platform can help you to get answers more efficiently than before, [contact us today](#).

RECENT POSTS

Thirty years in genetics – A perspective from Brazil

Lygia Pereira

04 August 2022

Network science: big data, strong connections, powerful possibilities

Dr. Ahmed Hamed

07 March 2022

Tackling complexity: making sense of genomic data

Dr. John Quackenbush

10 February 2022

The role of computational biology in drug discovery: an inside perspective

Dr. David DeGraaf

10 January 2022

Find out more.



©2022 Paradigm4 [Legal/Privacy/IP](#) [Site Credits](#)