IN THE EARLY DAYS of the pandemic, personal and professional aspects of Carlos D. Bustamante’s life were converging. His daughter had just fallen ill with a post-infectious condition at a time when he was partnering with a team at Stanford University to study why some children recover better from strep infections than others. So when the coronavirus came on the scene, Bustamante, a population geneticist by training, became curious about what patterns could emerge from individuals who contracted the virus and how they recovered. Very quickly, he asked his Stanford colleagues to keep all the coronavirus swabs they collected at the clinic so that he could study the valuable genetic material they held. “I viewed this COVID epidemic as a once-in-a-generation opportunity to follow people longitudinally,” he says. “We could sequence that material and link it to health record data to prospectively follow a person’s progress.”

When it came to collecting genetic samples, Bustamante prioritized diversity. As a Venezuelan American, he recognized the importance of representing Hispanic Latinos adequately in genetic datasets.

“We’ve known that we can do genetic mapping in Latin American populations for over 40 years,” he says, in a nod to the fact that the gene for Huntington disease was discovered in Venezuela. “I wanted to build that map, both for Latin America and for patients suffering from neuropsychiatric, cardiological, or metabolic conditions.”

Later in 2020, Bustamante cofounded Galatea Bio with Nicholas Katsanis and Alexander Ioannidis with the goal of expanding the diversity of genetic datasets beyond European communities to understand health disparities. The Florida-based company partners with organizations across the Americas to access genetic materials from Latin and South American populations to power research. It also offers ancestry-adjusted clinical genetic tests based on proprietary generative AI technology in-licensed from his work as a professor at Stanford and investigator at Chan Zuckerberg Biohub.

When Bustamante started the company, less than 1% of biobank samples that existed worldwide were from Latin American populations. The overwhelming majority—about 90%—were from individuals of European ancestry. And considering that over 650 million Latin Americans...
represent more than 700 ethnic groups across 33 countries, Bustamante saw the potential to understand and improve their health outcomes.

The genetics of those populations and individuals are recorded in the Biobank of the Americas—a key part of Galatea’s business and vision. The CAP-accredited biobank has the capacity to house 10 million DNA samples. Galatea uses Illumina NovaSeq 6000 and NextSeq 2000 Systems to sequence genomes. In the near future, Bustamante is looking forward to sequencing genomes on Illumina’s newly released NovaSeq X Plus System.

Galatea worked with hospitals that were collecting nasal swabs from patients coming in for COVID-19. These so-called COVID discards were invaluable for Bustamante’s team. The swabs were already stored in a medium that readily allowed researchers to extract and sequence the DNA. The original research team at Stanford and Chan Zuckerberg built an AI/machine learning model that included clinical, genomic, and epidemiological features to predict COVID severity.

Galatea has since expanded to partner with blood banks and hospitals in Colombia, Paraguay, Mexico, Puerto Rico, and throughout Latin America to collect genetic samples. Bustamante is especially interested in using this genetic information to understand how responsive a patient will be to a particular drug. For instance, Puerto Ricans and Mexicans bookend the spectrum when it comes to asthma rates in the Americas: Puerto Ricans have the highest asthma rates while Mexicans have the lowest. Puerto Ricans also have the lowest response to albuterol, which is the most prescribed medication for asthma. Bustamante says that a person’s response to the drug depends on a combination of environment, genes, and possibly epigenetics: “We’re trying to work with groups that are interested in these kinds of problems, and eventually improve the compendium of medications that can benefit everybody.”

In addition to sampling from relevant populations, he’s also looking to give back to the individuals and communities who contribute to the Biobank of the Americas. Currently, participants who submit their samples can access their own ancestry test results—or they can always opt out of the project if they later change their minds. “It’s all part of the right way of doing research, engaging the community, and being really upfront about what we’re doing, why we’re doing it, and what people can learn,” he says. Galatea Bio also offers clinical genetic laboratory-developed tests, including whole-genome sequencing, whole-exome sequencing, and ancestry-adjusted panels for autism, hereditary cancer, immune dysregulation, cardiovascular disease, thyroid disease, and pharmacogenomics.

Bustamante hopes that the biobank Galatea is building can eventually facilitate an exchange of knowledge among investigators, individuals, and communities. Their goal is to have 10 million data points; the biobank currently has about half a million. “We’ll be at over a million by the end of this year, and we should be at 10 million by, say, 2026,” Bustamante says. “We already have one of the most diverse biobanks in the world. And we’re really looking forward to being able to make use of the health care data that’s been entrusted to us and do right by the communities we’re working with.”