Asia-wide genome mapping project reveals insights into Asian ancestry and genetic diversity

December 4 2019

(Left to right) NTU Assistant Professor Hie Lim Kim, NTU Professor Stephan Schuster and GenomeAsia 100K Executive Chairman Mahesh Pratapneni, with the analysis of the Asian genomes which have been sequenced by all members of the consortium using the same methodology. Credit: NTU Singapore

After a global genetic comparison, a team of international scientists has discovered that Asia has at least ten ancestral lineages, whereas northern Europe has a single ancestral lineage.

In their first study reported in *Nature* this week, the GenomeAsia 100K consortium analysed the genomes of 1,739 people, which represents the widest coverage of genetic diversity in Asia to date.

The study covers 64 different countries and provides what the authors call "the first comprehensive genetic map for Asia" that will guide scientists in studying diseases unique to Asians, improve precision medicine and identify drugs that may carry higher risk of adverse reactions for certain ethnic groups.

Despite forming over 40 per cent of the world's population, Asian people have previously accounted for only six per cent of the world's recorded genome sequences.

The goal of GenomeAsia 100K, which launched in 2016, is to better understand the genome diversity of Asian ethnicities by sequencing 100,000 genomes of people living in Asia. It is a non-profit consortium hosted by Nanyang Technological University, Singapore (NTU Singapore), the only academic member. Its three other members are Macrogen based in South Korea, Genentech, a member of the Roche Group in United States, and MedGenome from India/US.

NTU Professor Stephan C. Schuster, the consortium's scientific chairman and a co-leader of the study, explained the significance of GenomeAsia 100K's initial findings on the vast genomic diversity in Asia: "To put it into context, imagine that we looked at all people of European ancestry and based on the level of their genetic diversity, observed that they could all be grouped into just one ancestral lineage or population. Now, if we took that same approach with our new data from

people of Asian ancestry, then based on the much higher levels of genetic diversity observed, we would say that there are 10 different ancestral groups or lineages in Asia."

Prof Schuster added, "GenomeAsia 100K is a significant and far-reaching project that will affect the well-being and health of Asians worldwide, and it is a great honour for Singapore and NTU to be hosting it."

Executive Chairman of GenomeAsia 100K, Mahesh Pratapneni said, "The publication of this pilot study is a first milestone for GenomeAsia 100K, which is an unprecedented collaboration between academia and industry leaders in the field of genomics. We are certain more partners will join GenomeAsia 100K to accelerate medical breakthroughs for people of Asian heritage."

Chairman and CEO of MedGenome, the largest genomics and molecular diagnostics provider in South Asia with facilities in the US, Singapore and across India, Sam Santhosh, said, "We are excited that over 1000 whole genome sequence data from the Indian sub-continent will now be available to researchers; this is an initial step in covering the underrepresented geographies."

Prof Jeong-Sun Seo, at Seoul National University Bundang Hospital Consortium scientific co-chair and Chairman of Macrogen, said, "I hope this Asian-focused study serves as a stepping stone for the democratisation of health care and precision medicine in Asia."

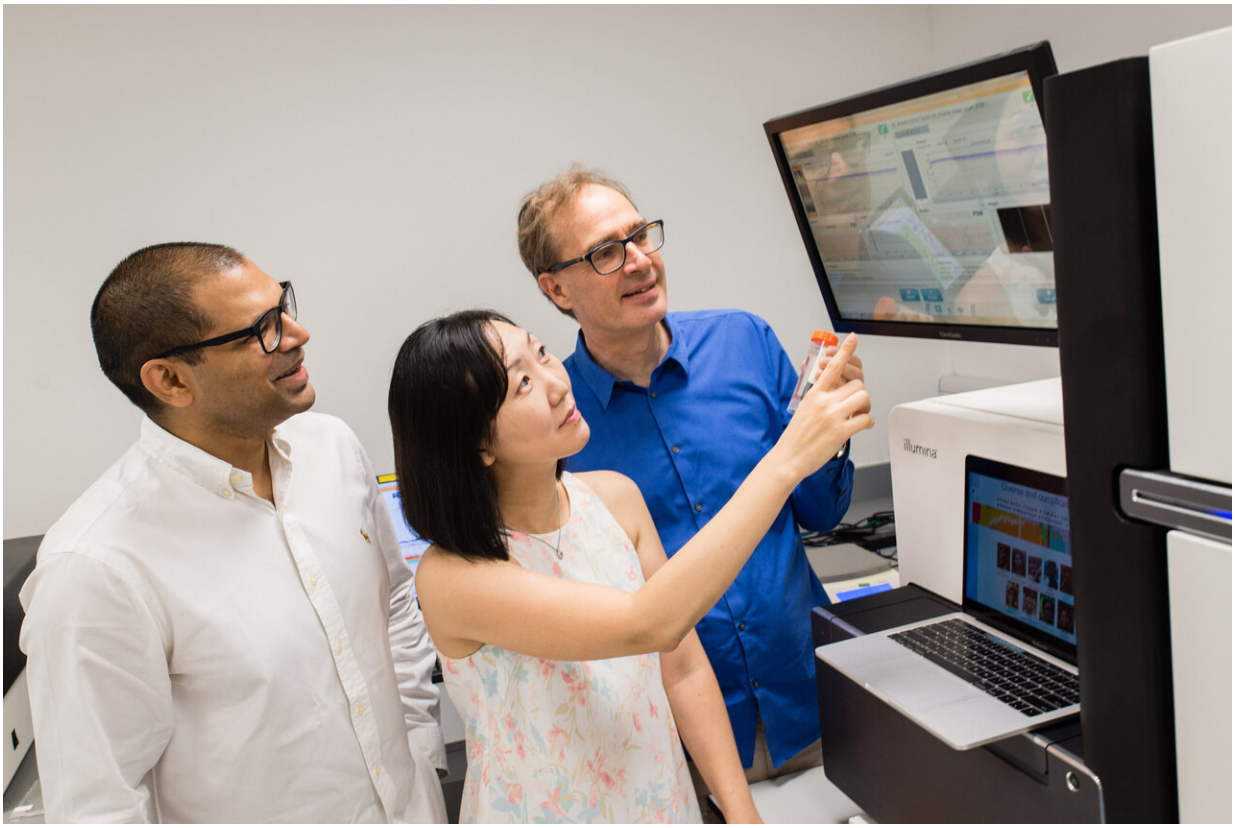
How the database of Asian genomes was formed

Over the course of the last three decades prior to the pilot project, thousands of blood and saliva samples were collected by scientists and anthropologists from donors across Asia in hopes that one day a deeper

analysis to gain insights into the Asian community could be done.

Of particular interest were participants from remote and isolated communities, who have long been the subjects of study by anthropologists but had not yet undergone genomic analysis, until the GenomeAsia 100K project was kickstarted.

The pilot study included 598 genomes from India, 156 from Malaysia, 152 from South Korea, 113 from Pakistan, 100 from Mongolia, 70 from China, 70 from Papua New Guinea, 68 from Indonesia, 52 from the Philippines, 35 from Japan, and 32 from Russia.



(from left) GenomeAsia 100K Executive Chairman Mr. Pratapneni, NTU Assistant Professor Hie Lim Kim and Professor Stephan Schuster discussing the results of the genome study involved the use of the high-throughput DNA

sequencers at NTU and the other companies. Credit: NTU Singapore

Genomic DNA extracted from the blood and saliva samples was then sequenced in laboratories of the four consortium members in the US, India, South Korea and Singapore. The digital sequencing data were subsequently sent to Singapore for processing and storage.

Singapore was selected by the consortium as the host, as the country offered good travel connections for collaborating scientists, strong supercomputing facilities to crunch the data, and the required cybersecurity standards in its data centre for handling sensitive genetic data.

The combined data was compiled and analysed by NTU scientists, including Asst Prof Hie Lim Kim, a population genomics expert at the Asian School of The Environment, with the help of the National Supercomputing Centre Singapore (NSCC) and international collaborators.

Different Asian ethnic groups respond differently to mainstream drugs

Every person has approximately 3.2 billion different nucleotides, or building blocks, in their genome, which form their DNA 'code.'

It's estimated that for the genomes of any two people, 99.9 per cent of this code is the same and on average, 0.1 per cent or three million nucleotides, are different between them.

This genetic variance helped humankind colonise the most diverse

environments on the planet and make it resilient to disease, but it also results in a differential response to many medicines.

"Genetic variance is the reason we are distinctively different from each other, including differences in the diseases that each of us suffer from during our lifetimes. Understanding these differences is the most important source of clues that we have for driving the discovery of innovative new medicines," said Dr. Andrew Peterson, an author of the paper and an expert in the use of genetics to drive drug discovery.

Dr. Peterson, who was head of Molecular Biology at Genentech while this work was being carried out, is now Chief Scientific Officer at MedGenome, where he leads drug discovery efforts at MedGenome's Seven Rivers Genomic Medicines division.

The frequencies of known genetic variants related to adverse drug response were analysed for the genomes collected in this study.

For example, warfarin, a common anticoagulant drug prescribed to treat cardiovascular diseases, likely has a higher than usual risk of adverse drug response for people carrying a certain genetic variant. This particular genetic variant has a higher frequency to appear in those with North Asian ancestry, such as those with Japanese, Korean, Mongolian or Chinese heritage.

Using data analysis, scientists can now screen populations to identify groups that are more likely to have a negative predisposition to a specific drug.

Knowing a person's population group and their predisposition to drugs is extremely important if personalised medicine is to work, stressed Prof Schuster: "For precision medicine to be precise, you need to know precisely who you are."

NTU Asst Prof Hie Lim Kim, who leads the project's efforts in population genetics, added: "Only by sequencing the entire genome of an individual can a person's ancestry and genetic background be known. Their genome explains why some people are afflicted by certain diseases while others aren't. Scientists know that there is no single drug that works well for everybody and our latest findings not only reinforce this, but suggest how specific groups could be harmed by specific medicines."

Moving forward, the GenomeAsia 100K will continue to collect and analyse up to 100,000 genomes from all of Asia's geographic regions, in order to fill in the gaps on the world's genetic map and to account for Asia's unexpected genetic diversity.

More information: The GenomeAsia 100K Project enables genetic discoveries across Asia, *Nature* (2019). DOI: 10.1038/s41586-019-1793-z , nature.com/articles/s41586-019-1793-z

Provided by Nanyang Technological University