



# MGI: The power of AI and genetic sequencing

The illnesses a person experiences in their lifetime are to a large extent determined by genetics. But today, the predictive power of gene sequencing technology is becoming invaluable in preempting disease. MGI Tech's mission is to benefit mankind through genomic technologies. The BGI subsidiary focuses on developing instruments and equipment for the medical and health sectors, providing a range of real-time, full-scenario, and full-life-cycle digital devices. However, the resulting rise in data requires the right digital infrastructure to make these services a reality.

By Xu Shenglan, Xue Hua

**D**r. Jiang Hui is Chief Operating Officer of MGI and vice president of BGI Research. She explained to us that gene sequencing technology can help solve practical issues in precision medicine, describing how innovations in digital technology – such as AI, 5G, and cloud computing – can help transform the medical

industry.

## Genome sequencing is within reach

The human genome defines most characteristics of the human body, including the risk of suffering from



Everyone will be able to have a record of their own genetic data in full. ‘All-in-one-day’ is within reach – acquisition, analysis, and interpretation of genetic data completed in one day.



certain diseases. Knowing our genomes’ features – or the flaws in our genes – can let us understand these traits and the odds of getting a given disease. Gene sequencing and data analysis can aid research on the relationship between diseases and specific gene forms and enable early detection and treatment.

Even in cases where certain gene-related diseases have already developed, gene sequencing can help find the genes that correspond to mutations, which can aid targeted therapy. In fact, this is an area of research in precision medicine.

“The concept of ‘precision medical treatment’ first appeared in the medical community in the US in 2011. The focus of precision medicine is not ‘medicine’ but ‘accuracy,’” says Jiang Hui. “Gene sequencing technology can be used to predict what diseases might develop in the future and thus better prevent them. We can carry out early diagnosis after a disease develops so diagnostic drugs can be used in a more targeted way.” As a result, she says, doctors can provide patients with the most appropriate treatment and drugs at optimal doses and times with minimal side effects, which in turn can help with post-treatment care.

In January 2015, US President Barack Obama

announced the Precision Medical Initiative in his State of the Union Address, increasing its profile and prompting projects to start popping up, including in China. In 2016, precision medicine was included in China’s 13th Five-Year Plan on Scientific and Technological Innovation, sparking the rapid development of gene technology in the country. The main areas of focus are cohort studies of diseases, molecular classification and staging of diseases, personalized treatments, and big data collection and mining. In October 2017, MGI Tech launched two new high-throughput genetic sequencers, MGISEQ-2000 and MGISEQ-200, at the 12th International Conference on Genomics (ICG-12) in Shenzhen. At the same time, BGI launched the Genome Decode Program (GDP) and the Chinese Millionome Database (CMDB).

Jiang Hui says, “With these magical tools and the help of scientific research, next-gen gene sequencing technology has revolutionized cost. In the near future we’ll be able to complete rapid testing on 24 tumor samples per day. In three years, we’ll have 100 yuan (US\$15.6) genome sequencing. Everyone will be able to have a record of their own genetic data in full. ‘All-in-one-day’ is within reach – acquisition, analysis, and interpretation of genetic data completed in one day.”

After costs fall, ordinary people will be able to enjoy the benefits of genome sequencing. According to Jiang Hui, “China has a relatively high birth defect rate at around 5.6 percent. People can carry out various genetic tests before and during pregnancy and after birth to reduce and prevent birth defects. High-throughput sequencers can be used to perform embryo detection and post-pregnancy neonatal testing.” She believes that in the future, genome sequencing will offer infinite possibilities in both scientific research and clinical medicine, and even in fields such as agriculture, forestry, animal husbandry, and fishing.

## Harnessing AI

The closer integration of AI and gene sequencing will help develop the industry. Many Internet giants have been crossing sectors to enter the genetics market. Google, in collaboration with DNAnexus, has set up a vast open genome database that is free to access. And AWS’s Cloud Database also provides similar open databases on its public data platform. Going forward, we will see more algorithms that can be used in genomic data analysis.

The global gene sequencing market is projected to reach US\$13.8 billion by 2020 at a compound annual growth rate of 18.7 percent. MGI has been successfully exploring gene sequencing and AI, with achievements in big data deep mining and applications. Jiang points out that, “With AI, we can develop more accurate disease and phenotype prediction models through the in-depth analysis of desensitized gene data.” This, she says, will help us build a more detailed genetic structure of the population and hopefully find new drug targets. For example, non-invasive prenatal genetic testing was originally used for screening fetal chromosomal disorders such as Down’s syndrome. Although it was not designed for detecting cancer, MGI’s big data and AI makes it possible to identify tumor signals in non-



“ With AI, we can develop more accurate disease and phenotype prediction models through the in-depth analysis of desensitized gene data.

”

— Dr. Jiang Hui, Chief Operating Officer of MGI and vice president of BGI Research



The rapid rise in demand for healthcare and serious shortages on the supply side have driven the integration of technologies such as AI in the healthcare industry.



invasive gene detection data by analyzing millions of samples and genetic characteristics of tumors. “There have been more than 40 cases of gestational cancers discovered early through non-invasive prenatal genetic testing data, and dozens of families have benefited from early detection of the disease,” Jiang says.

By the end of 2017, BGI had run genetic tests for pregnancy on almost 5 million women, including more than 2.8 million non-invasive prenatal genetic tests, resulting in a detection and specificity rate of over 99 percent, including more than 38,000 fetuses with abnormal chromosomes. More than 1.53 million newborns and pregnant women have received genetic screening for deafness, which has prevented approximately 80,000 carriers from developing hearing losses.

Jiang Hui also points out that, aside from gene sequencing, huge potential exists for AI in other applications in the healthcare industry. She says, “China has 22 percent of the world’s population but only 2 percent of the world’s healthcare resources – a serious shortage. Statistics show that 80 percent of healthcare institutions in China are located in urban areas, with just 20 percent in rural and remote areas.” Most of the new, high-tech equipment and best medical professionals are found in big cities and large hospitals. There isn’t enough equipment and

staff to meet the healthcare needs of everyone, so resources are strained in large hospitals. The rapid rise in demand for healthcare and serious shortages on the supply side have driven the integration of technologies such as AI in the healthcare industry.

“In terms of the collection and analysis of objective data, AI is much faster and more capable than humans,” says Jiang, which is especially the case in medical imaging. “Because medical image data is relatively standardized, this lends itself very well to machine recognition and deep learning. Intelligent auxiliary diagnosis systems can slash the time physicians spend looking at panels by an average of 4.25 hours and increase accuracy to over 90 percent,” Jiang says. This not only gives doctors more time to enhance service levels but also “returns” them to patients, giving doctors more time to spend on treating patients. In addition, AI can integrate different data, such as genetics and imaging, which, she says, allows doctors to more comprehensively analyze correlations in different stages of diseases.

## 5G and cloud computing

in 2017, Shenzhen’s Luohu Hospital Group began trialing MGI’s remote ultrasound diagnostic system MGIUS-R3. The device expands the range of

application scenarios for ultrasound equipment. It can play a crucial role in remote or extreme environments or in community centers. This device overcomes the limitations of traditional ultrasound diagnosis and treatment methods, and is a potential way to even out the distribution of medical resources. Integrating genetic data and imaging data will provide equal access to high-quality precision medical services for all.

The bandwidth and latency demands of real-time remote-control applications, such as remote ultrasonic inspection and remote surgery, are extremely high. Surgeons' hand movements, image transmission, and force feedback require high levels of synchronization. This is difficult to achieve using current networks. To solve this problem, Huawei Wireless X Labs and MGI have jointly looked at using 4.5G/5G networks to provide signal and data connections for remote ultrasound robots to help doctors remotely control patient-end ultrasound probes for diagnosis.

4.5G/5G networks can provide several megabits of transmission bandwidth for the real-time transmission of audio and video and b-scan ultrasonography imaging. 4.5G/5G networks also support low latency, allowing force feedback signals of a patient's body surface to be sent back to a physician's haptic device in just a few milliseconds. 4.5G/5G networks can also send ultrasound images to the cloud to provide real-time analysis and multi-level assistance to doctors.

According to Jiang Hui, "With a remote ultrasonic diagnosis system supported by 4.5G/5G networks, a high-sensitivity force feedback system and scan diameters of 850 mm are possible. This meets requirements for scanning the whole body including heart, neck, chest, and abdomen." Moreover, the system's HD video and voice system enables specialists to communicate with patients unimpeded.

Ultrasound image parameters can be adjusted without any latency, supporting rapid, real-time diagnosis and making diagnosis and treatment more efficient.

Gene sequencing is moving from individuals to larger populations and from research to broad commercial application. Moreover, computational analysis models are also starting to be used online in the cloud, whereas before they were restricted to local, offline use.

"The cost of gene sequencing is falling, while data is being produced at a greater and greater speed. We can't build a large-scale data center capable of keeping up with the massive amounts of data being produced on our own," says Jiang Hui. "We produce 300 TB of data every month and we now have a total of 22 PB of data. How can we deal with this explosive data growth?"

She also states that the hospital needs to enable data sharing of "our national gene bank with our partners."

"Cloud computing meets these needs well," she says. "We've improved our storage performance by 25 to 30 percent through long-term testing and using the Huawei OceanStor 9000 cloud storage system. The time it takes to deliver reports to users has dropped from 15 to 7 days."

OceanStor 9000's unified file system and hierarchical storage management are also of great help to MGI's data lifecycle management and support non-invasive prenatal genetic testing services. "In the era of integration of IT and biotechnology, we want to join forces with Huawei to tap into more opportunities to benefit humanity using genomic technology," Jiang says. 