version of a technique called angular analysis that was originally developed by the CMS Collaboration to characterize the Higgs boson. The *J*, *P* and *C* quantum numbers of the Higgs boson were established from the distribution of scattering angles of the particle's decay products[7].

To probe the structure of the all-charm tetraquark, the researchers applied this version of angular analysis to observations of all-charm tetraquarks that had been made at the CMS experiment between 2016 and 2018. Their analysis indicated that the *J* quantum number is 2 and the *P* and *C* quantum numbers are both +1 — this set of properties is expressed as 2^{++}. Although it does not rule out either theory, the 2^{++} result favours the compact tetraquark rather than the loosely bound pair of mesons. It also supports previous theoretical calculations[8,9].

Quarks are a foundational part of the standard model of physics, the conventional theory of fundamental particles and forces. Studying exotic hadrons might improve scientists' knowledge of the strong force. This knowledge, in turn, could help to explain how hadrons formed a few microseconds after the Big Bang; aid the development of accurate models of neutron stars; and help scientists to search for physics not described in the standard model.

Upcoming and recently completed upgrades to existing particle-collider facilities and experiments that have observed exotic hadrons, such as the High-Luminosity LHC, the Beijing Spectrometer III (BESIII) in China and Belle II in Japan, promise better detectors and more particle collisions — hopefully resulting in more observations of exotic states. Several future facilities and experiments are being designed to generate exotic states. These include the Super Tau-Charm factory in China, the Electron-Ion Collider (EIC) in the United States and the PANDA proton–antiproton collider in Germany. Alongside some of the planned upgrades, these will not just improve existing methods, but will also explore new pathways for generating exotic hadrons.

With this next generation of experiments, physicists hope to generate exotic hadrons that have been predicted but never observed, and to develop more ways to study the structure of these particles. Each exotic state might be a particular case that can teach physicists something new about the strong force. It might be that some exotic states are compact and others loosely bound — some might even be a mixture of both. This could be the beginning of a golden age of exotic-state research that will deepen understanding of a fundamental force that shapes the Universe.

Elena Santopinto is in the Sezione di Genova at the Istituto Nazionale di Fisica Nucleare (INFN), Genoa 16146, Italy.
e-mail: elena.santopinto@ge.infn.it

1. The CMS Collaboration. *Nature* **648**, 58–63 (2025).
2. LHCb Collaboration. *Sci. Bull.* **65**, 1983–1993 (2020).
3. ATLAS Collaboration. *Phys. Rev. Lett.* **131**, 151902 (2023).
4. CMS Collaboration. *Phys. Rev. Lett.* **132**, 111901 (2024).
5. Gell-Mann, M. *Phys. Lett.* **8**, 214–215 (1964).
6. Zweig, G. Report No. CERN-TH-412 (CERN, 1964).
7. Bolognesi, S. *et al. Phys. Rev. D* **86**, 095031 (2012).
8. Becchi, C., Ferretti, J., Giachino, A., Maiani, L. & Santopinto, E. *Phys. Lett. B* **811**, 13592 (2020).
9. Belov, I., Giachino, A. & Santopinto, E. *J. High Energy Phys.* **2025**, 93 (2025).

The author declares no competing interests.

Genomics

Taiwan invests in genetic resource for health

Genevieve L. Wojcik

A large biobanking effort captures genetic and health data from half a million people in Taiwan, widening the diversity of genomes used to predict disease risk. **See p.117 & p.128**

Discovering genetic variants that are involved in health-related traits relies on analysing genetic and health information from large cohorts of people. However, there has long been an imbalance of genetic ancestries represented in this kind of research, which has excluded large groups of humanity from the downstream benefits[1]. Millions of genetic variants differ in frequency around the world owing to complex demographic processes, such as migration, that depend on geographic proximity and shared histories. Therefore, variants found to be associated with a trait in one population could be less frequent or completely absent in another. Including under-studied populations in genetic resources enhances scientists' ability to discover and precisely map variants that are potentially responsible for health traits and diseases[2]. Two articles in *Nature*[3,4] demonstrate the power of a new large-scale study of an under-studied population: the Taiwan Precision Medicine Initiative (TPMI), which consists of genetic and electronic medical-record data from about half a million participants across Taiwan.

Yang *et al.*[3] (page 117) describe the TPMI resource itself. More than 560,000 Taiwanese residents (Fig. 1) were recruited between 2019 and 2023 from 16 academic centres affiliated with 33 hospitals, which provide care for about 40% of the population. This is a notable increase in sample size compared with other large-scale studies in Taiwan, such as the Taiwan Biobank (around 200,000 participants), and matches the size of biobanks in other parts of the world, for example the UK Biobank. With a larger sample size, greater statistical power can be achieved. This enables the discovery of disease-associated variants and increases precision when modelling the genetics of the population.

The authors first examined the population structure of TPMI participants. Using several computational approaches, they could determine how similar the participants' genomes were to each other and to external genetic reference groups. The vast majority of participants were of Han Chinese ethnicity, which constitutes more than 95% of the contemporary Taiwanese population. However, the study population also included individuals with ancestries resembling those of Indigenous Atayal and Ami peoples included in a previously formed data set of diverse human genomes called the Simons Genome Diversity Project[5].

The age distribution of the TPMI, with an average age of about 56 years and including many centenarians, enabled the authors to interrogate trends over the past century, during which Taiwan saw a rapid population increase. Comparing people born in the years before and after 1950 — representing a period of massive immigration into Taiwan after the end of the Chinese Civil War — the authors found that the genomes of participants born after 1950 showed much more mixing (admixture) between major ethnic groups than did those of people born before 1950. This has ramifications for how genetic medicine is applied to different age groups in this population, because the frequency of genetic variants will have changed with the shifts in population.

The second paper, by Chen *et al.*[4] (page 128), describes the power of the TPMI to characterize the role of genetic variation in a large set of traits, also known as phenotypes. Focusing on the roughly 460,000 individuals who were genetically inferred to be of Han Chinese ancestry, the authors evaluated genome-wide associations between variants and 695 binary phenotypes (for example, having heart disease or not) and 24 continuous traits (such as blood

ALAMY

Figure 1 | A crowd in Taipei, the capital city of Taiwan.

pressure or body-mass index). These analyses represent a massive increase in sample size compared with previously published studies in Han Chinese populations[6]. For example, it nearly triples the sample size for type 2 diabetes cases, leading to the discovery of five previously unknown variants in genomic regions that were identified in smaller studies to be involved in diabetes risk[6].

Using a customized array-based technology to capture more population-specific genetic variation, Chen *et al.* used the TPMI to investigate biological markers of disease that were unobserved in other studies. They replicated many previously published variant–trait associations from genome-wide association studies (GWASs) of East Asian populations[7], but also found 95 new associations across 57 health outcomes. Most notably, 33 of these variants had a frequency of less than 1% in populations with European ancestries, partly explaining their absence in previous work despite similar sample sizes. These population-specific findings in turn enabled better accuracy when developing 'polygenic score' models. In these models, the individual effects of genetic variants are summarized across an individual's genome to predict the individual's relative risk of a particular health trait.

Given that humans are more than 99% genetically identical, it is worth interrogating why diversity in participants is necessary for better understanding the genome's role in health. One consideration is that the frequency of variants can differ across populations, as the current findings demonstrate. Another is that the epidemiological patterns of traits, such as their distribution and causes (both genetic and non-genetic), can differ globally.

An example of this involves hepatitis B virus, which is endemic in Taiwan and has a high public health burden. It is present in almost 10% of unvaccinated TPMI participants born before 1984, which is when a vaccination campaign in Taiwan began. Previous work has shown that susceptibility to and severity of infection with hepatitis B virus is partly influenced by genetics, but it is not well-known which genes are involved. In the TPMI, a sample size of nearly 24,000 cases empowered a GWAS of hepatitis-B-virus infection, identifying previously unobserved genomic locations associated with this under-studied trait. This GWAS would not be as well-powered using data from other studies, such as the All of Us Research Program, which includes only 2,300 cases of hepatitis B virus as of October 2025 (ref. 8).

"The two studies establish the TPMI as a valuable resource for genetic research into human health."

One feature of the TPMI, as with all biobanks, is that it is not a representative sample of the general population. Because recruitment was hospital-based, the average age of participants was higher than that of the wider population and the cohort was enriched for people with chronic and severe illnesses. These biases limit the utility of the resource for investigating childhood-onset conditions, as demonstrated by the under-representation of respiratory diseases, such as asthma. Other biobanks have similar issues. The average age at recruitment for the UK Biobank, for example, is slightly higher (56 years old) than the average UK age[9]. Although efforts to increase the representation of children and younger adults in biomedical research are under way, including the Environmental Influences on Child Health Outcomes (ECHO) Program[10] at the National Institutes of Health in the United States, further resources are needed to better understand people's health in the first few decades of life.

The two studies establish the TPMI as a valuable resource for genetic research into human health that will demonstrate the importance of such an investment for decades to come. Biobanks around the globe have reshaped how scientists and clinicians do research[11] and communicate those findings back to the participants and their families. The TPMI has already developed a platform for returning results to participants, in conjunction with genetic consultation, for 83 genetic conditions with variants that affect disease risk or responses to drugs.

These innovations in Taiwan have the potential to improve the health not only of the local population but also of other individuals of Han Chinese descent, who comprise nearly 20% of the global population. More broadly, the findings could improve health care for all of us by complementing existing studies and filling gaps to ensure that everybody benefits.

Genevieve L. Wojcik is in the Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205, USA.
e-mail: gwojcik1@jhu.edu

1. Wojcik, G. L. *Cell* **185**, 4256–4258 (2022).
2. Wojcik, G. L. *et al. Nature* **570**, 514–518 (2019).
3. Yang, H.-C. *et al. Nature* **648**. 117–127 (2025).
4. Chen, H.-H. *et al. Nature* **648**, 128–137 (2025).
5. Mallick, S. *et al. Nature* **538**, 201–206 (2016).
6. Spracklen, C. N. *et al. Nature* **582**, 240–245 (2020).
7. MacArthur, J. *et al. Nucleic Acids Res.* **45**, 896–901 (2017).
8. Mayo, K. R. *et al. Annu. Rev. Biomed. Data Sci.* **10**, 443–464 (2023).
9. Mutz, J., Roscoe, C. J. & Lewis, C. M. *BMC Med.* **19**, 240 (2021).
10. Knapp, E. A. *et al. Am. J. Epidemiol.* **192**, 1249–1263 (2023).
11. Abul-Husn, N. S. & Kenny, E. E. *Cell* **177**, 58–69 (2019).

The author declares no competing interests.
This article was published on 15 October 2025.