There have been great strides in our understanding of cancer genomics – the mutational events that lead to high cancer risks and those that drive progression in cancer patients. Using vast databases such [The Cancer Genome Atlas](#) (TCGA) from the United States, which has molecularly characterised over 20,000 primary cancer and matched normal samples, in 33 cancer types, patterns of genetic mutations and signatures specific to different types of cancers are now known.

But in the last few years it has become apparent that the predictive strength of the approach does not always work for patients who do not have European ancestry. In 2020 a consortium of researchers analysed data from 10,678 patients across all 33 cancer types in the TCGA. They found [significant ancestry-related differences](#) in genetic mutation rates, DNA methylation and RNA expression, which would impact the identification of cancer subtypes and mean predictions based on data analysed from the complete Atlas may not hold for those with African or Asian ancestry.

This could be important when it comes to identifying biomarkers for treatment options. It also has implications for the accuracy of polygenic risk scores, which assess the combined effect of multiple genes to predict the aggressiveness of a cancer. These models may not work across different ethnic groups if the data used in their creation is predominantly from patients with European ancestry.

A 2023 meta-analysis of 10 studies of data from prostate cancer patients, consisting of almost 20,000 data sets from patients with African ancestry, found [nine new areas of genetic susceptibility](#) for prostate cancer, seven of which were found only – or principally – in men of African ancestry. A similar study combined Asian and European [genome wide associations with colorectal cancer](#) to improve prediction across racial and ethnic populations.

Director of the Center for Clinical Cancer Genetics and Global Health at the University of Chicago, Olufunmilayo Olopade, has also studied genetic variance for breast cancer, particularly HR+/HER2– and triple negative breast cancer (TNBC). Patients of African ancestry experience higher rates of TNBC and higher rates of mortality than those of European ancestry. In an analysis of data taken from the records of the US Tempus database of breast cancer patients with TNBC or HR+/HER2– disease, Olopade and colleagues [found more than 8,000 genes were differently expressed](#) between patients in the two ancestral groups, with the two groups showing different patterns of dysregulated pathways.

Weiwei Zhai, now at the Institute of Zoology, at the Chinese Academy of Sciences in Beijing, has been studying this issue with former colleagues at the Genome Institute of Singapore, at the National Cancer Centre Singapore. In their 2020 study they demonstrated ancestry-based differences in the molecular profiles of patients with lung adenocarcinoma (LUAD) – a highly heterogeneous cancer that currently has few distinct biomarkers.

Comparing a large genomic and transcriptomic dataset of East Asians (EAS LUAD, $n$=305) with LUAD profiles in the TCGA, they found that the East Asian LUAD genomes [had fewer mutations and copy number alterations](#) – commonly referred to as greater 'genome stability' – than the TCGA genomes. They also had better predictive accuracy for East Asian patients with lung adenocarcinomas than using molecular phenotypes from the TGCA for prognostic stratification. That study also identified a new Asian-specific LUAD transcriptomic subgroup with upregulated immune-related genes, opening up the possibility that patients with these profiles would respond better to immunotherapy-based approaches.

In parallel, Zhai and colleagues also [compared ancestry signatures for hepatocellular carcinoma](#), a liver cancer that is the third leading cause of cancer-related deaths worldwide. They found the factors affecting the molecular signatures of the disease differed between patients with European and Asian ancestries. The frequency of driver genes, by contrast, did not differ much across

ethnicity. In contrast to the findings of the lung adenocarcinoma study, the molecular profiles of hepatocellular carcinoma showed higher genome instability in Asians. The study also found a new transcriptomic subtype, known as P2, that differentiated the two ethnicities, with P2 being a clinically aggressive Asian subtype associated with deletions on chromosome 16.

So far these studies are just scratching the surface. What they tell us is that precision approaches to cancer treatment and risk-stratified prevention need to be adapted to the different cancer genomic signatures characteristic of different ethnic ancestries. *Cancerworld* spoke to Weiwei Zhai about the challenges to make that happen, and how to overcome them.



Weiwei Zhai, Chinese Academy of Sciences, Beijing

### CW: How difficult is it to study the molecular cancer profiles of patients with different ancestries across the world?

**WZ**: Ethnic comparisons are not so easy because of the way the clinical practice is done [across the world]. There are a lot of differences in the way we handle and collect the data, so a rigorous comparison is actually very challenging.

Of the 33 cancer types being molecularly profiled in the TCGA study of about 10,000 cancer patients, in only one was there roughly equal representation of patients of European and Asian ancestry. This was for HCC [hepatocellular carcinoma], because it's very difficult to collect a large number of HCC patients in the US, so they used an Asian cohort, collected from Vietnam, that had been processed and sequenced using the same approach.

HCC is, in general, a disease of poor or underdeveloped countries. In Asia, and especially in China, historically there's been a lot of hepatitis B virus infection, which is a high risk factor for HCC. In many European countries and Japan, there's a lot of hepatitis C virus infection leading to HCC, and also alcohol-related HCCs. So there seems to be many aetiological differences between two ethnic backgrounds, which provided us with a good opportunity to do a comparison.

### CW: Are all ethnic disparities in molecular cancer signatures connected to ancestral differences or could other factors be at play?

**WZ**: They are some environmental factors involved. For example, in Asia, we use some traditional Chinese medicine that can cause genome lesions. It's described in our paper as the 'aristolochic acid' signature. We see these signatures in liver cancer patients, and this is quite prevalent. If you look at HCC patients in Taiwan, 60 to 70% have the aristolochic acid signature. In mainland China it's somewhere between 30 and 60%.

Lung adenocarcinoma has a very peculiar profile in Asia. There are a lot of female non-smokers prone to getting this type of lung adenocarcinoma. We found the 'flagship' mutated driver gene – the epidermal growth factor receptor (EGFR) gene – is very common in this type of cancer in Asia, especially in female non-smokers. It is mutated in 40–60% of patients in Asia with lung

adenocarcinomas, but only in 10–15% in Europeans. This is because in a lot of European countries, people smoke and carry mutations linked to this risk factor.

Asian non-smokers tend to have higher EGFR mutations. For example, in the lung cancer story, people with Latino heritage, are a mixture of European and Asian ancestries, and colleagues find that EGFR mutation frequency correlates with the Asian Ancestry in Latinos as well. We don't yet have the full machinery or architecture or cohorts to really dissect which components are due to genetics, and which are due to environments… but this is becoming a very active area of research.

**CW: What have we learnt about ethnic differences in the stability of our genomes, and what are the implications?**

**WZ**: In our lung adenocarcinoma comparison, genome instability was lower in the Asian cohort. You have so few mutation events in the Asian genome, but then those events are more likely to be the driver event. Because the number of events is so much smaller, their importance is higher.

Because Europeans smoke a lot, there's a lot of mutation events happening in their genome. Some of them are driving events, some of them are what we call 'passenger' events – they're not really driving tumorigenesis. This means the background is very noisy, it's not as easy to see which events are really important.

But with hepatocellular carcinoma we see the opposite trend. We see higher genome instability in Asians. We don't yet know the exact reason, and we will need to research into this further.

**CW: How will these studies help us build better predictive models of cancer progression?**

**WZ**: You can build a model that combines multiple features to predict patient survival. We find that the model we build using Asian data is much better at predicting Asian patients' survival than those that are built using the European ancestry cohort. That's not saying that Asian or European patients survive better or worse compared to each other – it's the predictive accuracy of the models that is different for different ethnicities. In the case of lung adenocarcinoma, the predictive accuracy of our model, based on East Asian data, is better in Asians.

**CW: How can we improve our understanding of cancer in patients of all ancestries?**

**WZ**: In cancer research a lot of knowledge and the way things are done have developed from Western countries… I think there needs to be a parallel story or architecture built based on other ethnic backgrounds. Diseases like hepatocellular carcinoma and lung cancer have a lot of ethnic disparities, so we need more than just one single view of these diseases.

A lot of researchers in Asia are busy catching up with the West, rather than looking at the bigger picture of how to build something in Asia. A lot of Asian countries, including Japan, China, and Korea are working in a more local manner, but this work needs to be done via an academic medical centre, where scientists and clinicians collaborate. We have not reached that point yet, which means we are still relying on the Western paradigm.

There's a lot of narrow thinking going on everywhere, but in science there should be no boundaries to anything. Understanding ethnic differences in cancer requires collaboration across countries – we need to look at the same object from different angles with a very open and collaborative spirit across ethnic backgrounds. I hope we can find a way to work together to benefit all humankind.