POPULATION GENOMICS

Diverse human genomes

Credit: Mopic/Alamy

Most large-scale human genome sequencing projects to date have sampled either large, metropolitan populations or only a few individuals from more diverse groups. Now, a study in *Science* demonstrates that anthropologically informed genome sequencing can provide a fuller understanding of human genetic variation than previous approaches.

In the study, 929 whole genomes, representing 54 populations with different geographical locations, languages and cultures, were sequenced at an average of 35× coverage. Linked-read sequencing was used to phase 26 of these genomes (representing 13 populations). Millions of the identified 67.3 million SNPs, 8.8 million insertions and deletions (indels) and 40,763 copy number variants were not detected in other large-scale projects, and hundreds of thousands of these novel variants are common in at least 1 of the 54 populations. Additionally, the discovered SNPs provide a more nuanced view of shared ancestry, particularly among African populations, than do variants typically included on common genotyping arrays.

Populations from all geographical regions were found to have some private common variation, but such variants reach high frequencies only in African, American and Oceanian populations. In most populations, this variation arose largely by novel mutation. However, a substantial proportion of private variants in Oceanian populations is derived from admixture with Denisovans. Patterns of genetic variation indicate that population sizes expanded for most groups over the past 10,000 years, except for hunter–gatherers in Africa.

Analyses of haplotype variation indicate that present-day population structure formed gradually over the past ~250,000 years, with evidence for more recent gene flow in most populations but also more ancient interactions in a few populations. Patterns of diversity in archaic haplotypes suggest that a single episode of admixing occurred between Neanderthals and the ancestors of present-day humans. By contrast, Denisovans are likely to have admixed multiple times with geographically distinct ancestral human populations.

The genome sequences from this study are freely available and provide a valuable resource for further examining human genetic variation from a range of perspectives, from anthropology through to medicine. Understanding whether population-specific variation has medical relevance will be particularly important.

Dorothy Clyde

ORIGINAL ARTICLE Bergström, A. et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* **367**, eaay5012 (2020)

DEVELOPMENT

Early chromatin dynamics

Higher-order chromatin structure undergoes dynamic changes after fertilization, but whether specific features of chromosome organization in gametes are passed on to the developing embryo or form anew has remained unclear. Now, a study in *Nature* provides new insights into chromatin dynamics and allele-specific gene expression during early mouse embryogenesis.

Allele-specific single-cell Hi-C on preimplantation embryos at defined stages of development (encompassing the 1-cell to 64-cell stages) showed the progressive formation of domains, which differed in number between the maternal and paternal genomes. The authors quantified the number of contacts within domains for each parental genome at every stage and through unsupervised clustering identified two main domain types that in further analyses could be associated with specific chromatin states.

Integrating single-cell Hi-C and chromatin immunoprecipitation and sequencing (ChIP–seq) data revealed an early parent-of-origin-specific domain type associated with allele-specific enrichment of the repressive histone mark H3 lysine 27 trimethylation (H3K27me3), with the highest enrichment seen on the maternal genome. These parentally preformed domains, most of which dissolved by the 4-cell stage, showed interaction patterns between domains similar to those observed for A or B compartments.

Another category of domains formed de novo on both parental genomes at different stages of development, resembled topologically associating domains (TADs) and was associated with active chromatin.

RNA sequencing data showed that parentally preformed domains were associated with repressed genes, while the other parental allele was expressed, albeit only at low levels. These early transient domains correlated with different allelic expression outcomes and were associated with lower gene expression overall. Gene ontology analysis revealed an enrichment for genes active at later developmental stages, for example, those involved in tissue morphogenesis. By contrast, de novo domains typically comprised genes active in early development.

Focusing on the paternal X chromosome before and during X chromosome inactivation in female preimplantation embryos, the authors observed a loss of TADs following or concomitant with gene silencing, whereas TADs in genomic regions escaping X chromosome inactivation were maintained. This finding suggests that transcription or an active chromatin state may be needed to maintain local structures.

Taken together, this study provides new views on allelic gene regulation in mouse early development and highlights that higher-order chromatin dynamics seem to take the form of early, parent-specific repressive compartments that switch to the gradual formation of local domains.

Linda Koch

ORIGINAL ARTICLE Collombet, S. et al. Parental-to-embryo switch of chromosome organization in early embryogenesis. *Nature* **580**, 142–146 (2020)

Credit: Vasiliy Vishnevskiy/Alamy