Research highlights

Early development

NR5A2 activates the zygotic genome

Shortly after fertilization, the zygotic genome remains in a transcriptionally inert state. In order for embryonic development to proceed, it needs to activated. Although some regulators of mammalian zygotic genome activation (ZGA), including DUX, have been identified in recent years, this process remains incompletely understood. Gassler et al. carried out DNA motif analysis of *cis*-regulatory regions near genes that are expressed in mouse two-cell embryos, and identified binding motifs for the orphan nuclear receptors NR5A2 and ESRRB. Chemical inhibition, targeted protein degradation or transcript depletion of NR5A2 showed that this factor is necessary for embryonic development past the two-cell stage and that it regulates the expression of a large fraction of ZGA-linked genes. CUT&Tag experiments indicate that, as predicted based on motif analysis, NR5A2 binds to SINE B1/Alu elements located in *cis*-regulatory regions of ZGA-associated genes, including transcription start sites, and also distal enhancers. Chemical inhibition of NR5A2 suggests that this receptor can increase chromatin accessibility, potentially working as a pioneer factor. Indeed, the authors demonstrate that NR5A2 can bind nucleosomal DNA in vitro. It will be interesting to test whether NR5A2 also regulates ZGA in other mammals, including humans.

Tiago Faial

Nature Genetics

Original reference: *Science* https://doi.org/10.1126/science.abn7478 (2022)

Disease genetics

Genetics of clonal hematopoiesis

Large-scale genetic studies have enhanced the understanding of germline causes of clonal hematopoiesis. Using exome-sequencing data from 628,388 individuals in the UK Biobank and the Geisinger MyCode Community Health Initiative, Kessler et al. identified 40,208 carriers of clonal hematopoiesis of indeterminate potential (CHIP). Using common-variant genome-wide association analyses, they identified 24 loci that contained variants associated with CHIP, 21 of which were newly identified associations. In addition, rare variant and gene burden exome-wide association analyses identified new rare variant associations with clonal hematopoiesis and telomere length. Cross-sectional association analyses of 5,041 traits in the UK Biobank revealed relationships between CHIP and a wide range of traits such as COVID-19 infection, cardiovascular disease and hematological malignancy. Mendelian randomization analyses showed an increased risk of developing solid cancers such as lung cancer among CHIP carriers. Overall, the study provides a valuable compendium of common and rare variant associations with CHIP and other clonal hematopoiesis phenotypes, which may facilitate further research on disease mechanisms and potential therapeutic agents.

Wei Li

Nature Genetics

Original reference: *Nature* **612**, 301–309 (2022)

Population genetics

Mapping dog behavior

Domestic dog breeds exhibit a diverse range of behavioral characteristics. To gain insights into the genetic basis of this diversity and how it was shaped by historical breeding practices, Dutrow et al. analyzed genomic datasets from a large collection of dogs, including purebred animals, pet dogs of mixed breed, semi-feral village dogs and wild canids. To account for the complex population structure of domestic dogs, they used a dimensionality reduction approach to define major lineage trajectories within their sample collection. They also used owner-reported questionnaire data from over 46,000 purebred dogs to calculate average behavioral metrics for each breed and then map these onto genetically defined lineage trajectories The strongest positive correlations were observed between the herder lineage and non-social fear and between the terrier lineage and predatory chasing, and the strongest negative correlation was between the scent hound lineage and trainability. Next, they performed genome-wide analyses to identify variants associated with each lineage and found that most lineage-associated variants were not lineage-specific but instead were present across the entire dataset, which suggests that selection during breed formation acted on pre-existing variation. Within sheepdogs, they observed enrichment of associated variants near genes implicated in axon guidance, providing clues into the neurobiology that underlies herding behavior.

Kyle Vogan

Nature Genetics

Original reference: *Cell* **185**, 4737–4755 (2022)

Cancer genomics

Fanconi anemia-associated signature in cancer

Fanconi anemia (FA) is caused by a loss of DNA interstrand crosslink repair, often as a result of endogenous or exogenous aldehydes. Many patients with FA develop head and neck squamous cell carcinoma (HNSCC), but how loss of the FA repair pathway shapes the cancer genome is unclear. Using various sequencing techniques, Webster et al. demonstrate that tumors showed high levels of *TP53* loss and structural variants including fold-back inversions, unbalanced translocations and small deletions. Structural variant breakpoints were prevalent at fragile sites and early-replicating regions, and were associated with the amplification of key HNSCC driver genes. Sporadic HNSCCs did not show these copy number changes, although tumors that displayed mutational signatures associated with tobacco and aldehyde exposure were notable exceptions. Loss of FA activity led to the upregulation of genes associated with epithelial-to-mesenchymal transition, and the activation of innate inflammatory signalling pathways. Together, these data highlight the role of aberrant crosslink repair in tumorigenesis.

Safia Danovi

Nature Genetics

Original reference: *Nature* **612**, 495–502 (2022)