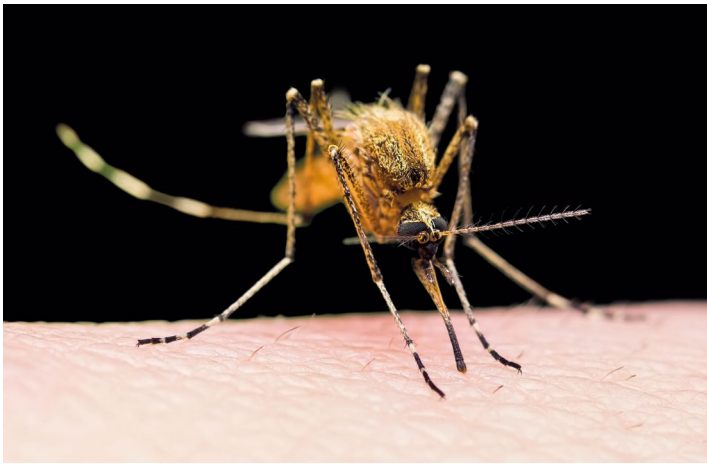
VIRAL EVOLUTION

Scanning for determinants of Zika virus host tropism



Credit: Konstantin Nechaev/Alamy

Zika virus (ZIKV) is a medically important member of the Flaviviridae family that cycles between insect (*Aedes* mosquitoes) and mammalian (primate) hosts. Understanding what governs viral fitness in each host is crucial for developing approaches to prevent disease. Now, a study in *Nature Microbiology* describes how deep mutational scanning (DMS), by speeding up virus evolution, can rapidly identify key determinants of host tropism and potentially aid vaccine development.

In DMS, a cDNA library is generated that encodes every possible amino acid (or codon) variant for a protein (or protein region) of interest. Variants are then selected for in a given environment and can be identified by deep sequencing and computational analysis. Setoh et al. applied DMS to the 204 C-terminal residues of the ZIKV E protein, which is involved in receptor-mediated infection, and the resulting library was transfected into mosquito cells (C3/36) or primate cells (Vero). Two highly enriched variants were detected in C3/36 cells (K316Q and S461G) and three in Vero cells (Q350L, T397S and a synonymous substitution at R416).

To investigate how the mosquito-selected substitutions affected viral replication in mammalian cells, the authors created a double mutant virus strain, 316Q/461G. Whereas the replication efficiency of 316Q/461G was similar to that of wild-type (WT) virus in C3/36 cells (which grow at 28 °C), it was substantially lower than WT virus in Vero cells and all other mammalian cell lines tested (which grow at 37 °C). Subsequent analysis revealed that 316Q/461G could bind to and infect mammalian cells but a temperature-sensitive defect in virion formation prevented efficient viral replication, secretion and spread. Furthermore, 316Q/461G replicated less efficiently in induced pluripotent stem cell (iPSC)-derived human brain organoids, which developed similarly to mock-infected organoids. By contrast, organoids infected with WT virus displayed abnormalities consistent with previous studies.

Next, Setoh et al. showed that 316Q/461G replication was attenuated in IFNAR$^{-/-}$ mice, which

COMPLEX TRAITS

Keeping score with obesity

An increasing knowledge of the genetic underpinnings of human diseases provides opportunities to leverage an individual's genetics to predict future disease risk. A new study in *Cell* reports a genome-wide polygenic score (GPS) that is predictive of obesity, thus providing opportunities for early health interventions.

Unlike monogenic Mendelian diseases, where a single genetic variant can greatly affect the risk of disease occurrence, complex traits and common diseases are typically underpinned by many genetic variants that individually make only a modest contribution to disease risk. However, as data accumulate from well-powered genome-wide association studies (GWAS), the concept of GPS has emerged — often termed a polygenic risk score (PRS) when used in the context of disease — whereby the cumulative small effect sizes across many genetic variants can be leveraged to provide an overall score for predicting traits.

In their obesity-focused study, Khera at al. used data from a GWAS that examined the association between 2.1 million common genetic variants and body mass index (BMI) in >300,000 individuals. They used a computational algorithm to generate a GPS to predict BMI based on an individual's genotype.

The GPS showed promising predictive power, with a correlation coefficient (predicted versus actual BMI) of 0.292 in a validation data set of nearly 120,000 individuals in the UK Biobank. This degree of correlation is consistent with BMI being only partially genetically determined.

Importantly, the GPS based on 2.1 million genetic variants showed greater predictive power than a score based on the top 141 variants from the GWAS that reached genome-wide statistical significance (correlation coefficient 0.133). This result emphasizes that there is substantial useful information among small-effect genetic variants that are below the threshold of genome-wide significance.

Khera at al. applied their GPS to additional validation data sets, demonstrating that individuals with a GPS in the top 10% of the population have a substantially elevated risk of obesity, bariatric surgery, other types of cardiometabolic disease and death. Analyses of cohorts where childhood development was tracked revealed that high GPS had only a minor effect on birth weight but that statistically significant differences in weight emerged throughout childhood.

> “individuals with extremely high GPS could represent a much larger high-risk population than those harbouring pathogenic obesity-related mutations”

Credit: Dimitri Otis/Getty Images Plus

lack the innate antiviral response and are therefore susceptible to viral infection. However, 316Q/461G infection elicited a protective immune response against a lethal dose of the African ZIKV MR766 strain. Taken together, these observations suggest that 316Q/461G has the potential to be developed as a vaccine, even for pregnant or immunocompromised individuals. Importantly, the substitutions were stably maintained in both mosquito and mammalian cells after 5 passages and the virus can be grown to high titres in C3/36 cells, further supporting its vaccine potential.

This study demonstrates that DMS can — within a few days — identify not only amino acids with host-specific roles in replication but also virus strains with the potential to be developed as vaccines.

Dorothy Clyde

ORIGINAL ARTICLE Setoh, Y. X. et al. Determinants of Zika virus host tropism uncovered by deep mutational scanning. *Nat. Microbiol.* https://doi.org/10.1038/s41564-019-0399-4 (2019)

Last, highlighting the cumulatively large effect of the variants in the GPS, for the 1.6% of the population with the highest GPS, the increase in BMI was equivalent to large-effect pathogenic mutations in the melanocortin 4 receptor (*MC4R*) gene. As these *MC4R* variants are present in only 0.14% of the population, individuals with extremely high GPS could represent a much larger high-risk population than those harbouring pathogenic obesity-related mutations.

It will be interesting to determine what more we can learn from obesity-related predictive scores, such as whether they shed light on molecular networks underlying obesity, and the degree of clinical value they add beyond a genetically naive approach of targeting health interventions to individuals already displaying elevated BMI.

Darren J. Burgess

ORIGINAL ARTICLE Khera, A. V. et al. Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell* **177**, 587–596 (2019)
FURTHER READING Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* **19**, 581–590 (2018)

Credit: Federico Rostagno/Alamy

REPROGRAMMING

When the elite compete

A new study in *Science* reports the existence of a subpopulation of somatic cells from which 'elite' clones emerge that outperform other clones to drive reprogramming, the transformation of differentiated cells into pluripotent stem cells. Indeed, competitive interactions — either indirect due to competition for limited nutrients or space, or direct competition between reprogramming clones — drive clonal dynamics in a population setting in such a way that somatic cells differ in their reprogramming potential.

Previous analyses of isolated single cells undergoing reprogramming suggested that every somatic cell has the ability to reprogramme, a concept known as clonal equipotency. However, the random nature of clonal reprogramming has been hard to gel with findings from population analyses, which suggest that cells undergo deterministic steps during reprogramming. Shakiba et al. thus set out to relate cell population outcomes and single-cell reprogramming events.

By tagging mouse embryonic fibroblasts (MEFs) with a unique cellular barcode, the team were able to track the number of clones present in a cell population after induction of reprogramming and assess clonal selection. A rapid drop in the number of clones was observed, coinciding with increased cell death, and it continued to the extent that less than 10% of clones survived 30 days of reprogramming.

In contrast to the previous findings that suggested all cells have reprogramming capabilities, the analysis of surface marker profiles showed that not all clones contributed to the final cell fraction, comprising induced pluripotent stem cells. The differing reprogramming potential of cells in population analyses versus isolated conditions suggests that competitive interactions between clones have a role in shaping population dynamics.

Analysing the surviving cells after 30 days of reprogramming, the authors found that two dominant clones had emerged and made up more than 50% of the culture. To determine whether these clones exhibited selective advantages over others, the investigators analysed clone size distributions. The observed bimodal distribution of dominant versus non-dominant clones challenged the assumption of clonal equipotency, which would be expected to yield a unimodal clone size distribution.

To test whether the dynamics observed experimentally could arise owing to differences in clonal fitness, the authors looked at the outcomes of the barcoding study based on clone size. Clones that were largest after 8 days and 14 days of reprogramming exhibited greater dominance than smaller clones. However, clones that were larger at the start of reprogramming — that is, owing to random differences in barcoding — did not become dominant. Further experiments indicated that clonal dominance was present already in the MEF state.

To identify the subpopulation of 'elite' MEFs, the researchers developed a lineage tracing strategy to retrospectively identify MEFs based on *Wnt1* expression, which marks MEFs derived from neural crest cells, a cell population responsible for the generation of a diverse array of cell and tissue types during vertebrate development, including skin and smooth muscle. After 3 weeks of reprogramming, MEFs derived from the neural crest formed 100% of the cell population, and these cells also had a higher probability of initiating reprogramming than non-neural crest-derived MEFs in both population analyses and in isolated clonal studies.

The implications of this study range from applications in synthetic biology to regenerative medicine. For example, a deeper understanding of the competitive ability of injected cells, that is, how these cells interact with endogenous cells and survive in a patient, may help to not only predict outcomes but also control cell competition.

Linda Koch

ORIGINAL ARTICLE Shakiba, N. et al. Cell competition during reprogramming gives rise to dominant clones. *Science* https://doi.org/10.1126/science.aan0925 (2019)