

DISEASE GENETICS

Genome editing for disease locus dissection

“ the presence of the risk haplotype resulted in defects in VSMC adhesion, traction and contractility, which were rescued by deletion of the risk haplotype ”

Genome-wide association studies (GWAS) have identified many genomic variants associated with human diseases, but biological interpretation of these associations typically requires detailed functional studies. A new investigation uses genome editing to dissect mechanistic roles of the most influential common genetic risk factor for coronary artery disease (CAD).

The 60 kb risk locus at 9p21.3 accounts for 10–15% of CAD risk in non-African populations, but various features make it challenging to dissect functionally. First, it harbours no protein-coding genes, which impedes gene-based functional studies and weakens the bioinformatics-based interpretation of genetic variation. Second, it is poorly conserved outside primates, meaning that rodent models are of limited relevance. Third, each allele of the locus typically consists of one of two haplotypes (low-risk versus high-risk) for which blocks of 50–100 single-nucleotide polymorphisms (SNPs) are co-inherited, meaning that there is limited focal variation for high-resolution analysis within the region.

Lo Sardo, Chubukov et al. sought to study the 9p21.3 locus in vascular

smooth muscle cells (VSMCs), as dysfunction of these cells is implicated in CAD pathology. Primary VSMCs are challenging to study, as they suffer growth arrest in culture, and many healthy and patient samples would be required to disentangle the effect of the locus from the effects of the variable genetic backgrounds. Therefore, the authors turned to genome editing of induced pluripotent stem cells (iPSCs). They designed a transcription activator-like effector nuclease (TALEN)-based system to delete the 60 kb locus homozygously in human iPSCs. Initially, the authors generated 19 independent iPSC lines consisting of wild-type and knockout pairs derived from 1 individual homozygous for the high-risk haplotype and 1 individual homozygous for the normal haplotype.

To study the effect of the haplotypes in relevant cell types, the investigators differentiated the iPSC lines into VSMCs and characterized their phenotypes. RNA sequencing (RNA-seq) revealed that deletion of the risk haplotype had not only a substantially greater effect on gene expression in VSMCs than did deletion of the non-risk haplotype but also restored gene expression profiles to resemble non-risk patterns. These results indicate that the risk haplotype exerts a gain-of-function pathological effect relative to the presence of the non-risk haplotype or the absence of either haplotype. Such a model was supported by functional studies, whereby the presence of the risk haplotype resulted in defects in VSMC adhesion, traction and contractility, which were rescued by deletion of the risk haplotype. Such effects are consistent with the

mechanisms of CAD pathology, whereby these VSMC defects are thought to contribute to atherosclerotic plaque formation.

What genomic features within the 9p21.3 locus contribute to CAD risk? Despite not harbouring protein-coding genes, the region encodes terminal exons for the *ANRIL* long non-coding RNA (lncRNA) gene, and the RNA-seq data revealed that the risk haplotype results in expression of shorter *ANRIL* splicing isoforms than the non-risk haplotype. Lentiviral expression of short *ANRIL* isoforms in non-risk haplotype VSMCs partially converted their gene expression profiles and cellular properties to resemble risk-haplotype VSMCs. As *ANRIL* only partially explains the phenotypic consequences of the risk haplotype, focal genome editing within the 9p21.3 region is likely to be required to localize the genomic features underlying the remainder of the risk effects.

As confirmation that the results are not specific to the genetic backgrounds of the two human donors used for the iPSCs, the authors recapitulated various results in a replication set of iPSCs from five additional donors.

Finally, the authors performed various network analyses on the genome-wide gene expression data, identifying modules of altered gene expression patterns in risk-haplotype VSMCs, with several altered genes overlapping previously identified CAD risk loci from large-scale GWAS meta-analyses.

Overall, this study demonstrates how genome editing can enable the study of otherwise poorly tractable biological systems, and the identified CAD-associated gene networks may provide future opportunities for biological understanding and therapeutic intervention in CAD.

Darren J. Burgess

ORIGINAL ARTICLE Lo Sardo, V., Chubukov, P. et al. Unveiling the role of the most impactful cardiovascular risk locus through haplotype editing. *Cell* **175**, 1796–1810 (2018)



Credit: Glow Images/Getty