Network approaches to understanding genome-wide association studies in breast cancer

Kersten Meyer*

kerstin.meyer@cruk.cam.ac.uk

Genome-wide association studies for breast cancer have identified nearly 100 risk loci, which individually are of small effect. An individual's risk for breast cancer is conferred by the combination of multiple of these small effect variants. To better understand how risk loci might combine, we examined whether risk-associated genes share regulatory mechanisms. We have modelled breast cancer as a gene regulatory network onto which the loci relating to risk can be mapped to identify key regulators. We have developed a computational pipeline (eQTL filtered variant set enrichment, EVSE) that identified 36 TFs that are frequently associated with transcriptional regulation of risk loci. The analysis encompasses several steps in which incomplete data sets are used to make inferences about causal variants and target genes that feed into the analysis. Despite the uncertainties associated with our analysis we were able to generate robust experimental data to validate our findings. The network-identified risk TFs were shown to bind at risk loci using ChIP-seq data and risk-TFs were frequently mutated in tumours, supporting a functional role in conferring risk. Within the regulatory network the risk-TFs lie in two opposing subgroups, which relate to ER+ luminal A/B and to ER- basal-like cancers and to different, luminal epithelial cell populations in the adult mammary gland. These computational results were confirmed in functional studies of risk TF activity. Our network approach provides a foundation to reveal the regulatory circuits governing breast cancer, to identify targets for intervention, and is transferable to other disease settings.

*Li Ka Shing Centre, CRUK Cambridge Institute, Robinson Way, Cambridge, CB2 0RE, UNITED KINGDOM