Statistical genetics, disease biology, and drug discovery

Yukinori Okada

Department of Statistical Genetics,

Osaka University Graduate School of Medicine

Statistical genetics is a research field that evaluates causality of human genetic variations on diseases, using statistical and bioinformatics approaches. Recent development of high-throughput genome sequencing and genotyping technologies, such as whole genome sequencing by next generation sequencers, have provided human disease genome data of hundreds of thousands of the subjects. Large scale disease risk association analyses successfully identified comprehensive catalogues of genetic susceptible loci that are linked to human diseases. However, little is known regarding how to develop methodology to integrate large-scale human genetic study results with diverse biological resources, to which statistical genetics should contribute. We have developed such methods and applied to a pioneering example of large-scale genetic association studies on a variety of human complex traits. In the case of rheumatoid arthritis (RA), an autoimmune disease characterized by inflammation and destruction of joints, our studies highlighted novel features of RA biology. (i) enrichment of microRNA-target gene networks, (ii) negative correlation of genetic risk with schizophrenia, (iii) significant risk of both classical HLA genes (*HLA-DRB1*, *HLA-DPB1*, *HLA-B*) and non-classical HLA genes (*HLA-DOA*), (iv) specific localizations of RA risk variants in the promoter regions of regulatory T cells, (v) overlap of risk genes with lymphoma and primary immunodeficiency, and (vi) epigenetic contribution of regulatory T cells and Th_{17} cells on disease biology. Further, we demonstrated that the RA risk genes were significantly enriched in overlap with the target genes of the drugs currently used for clinical treatment of RA, and that *in silico* analysis of the network between the disease risk genes and the drug target genes could identify candidate of drug repositioning (e.g. CDK4/6 inhibitors). These results should empirically show the value of statistical genetics to dissect disease biology and novel drug discovery.