ABSTRACT: There are many differences in the aims of studying the genetics of complex traits in farm animals and human populations. Very broadly, animal breeding is the science and application of making predictions, whereas human complex trait genetics is about understanding the path from genome to phenome.

Genetic studies of complex traits in human populations are usually performed to better understand individual differences in risk to diseases. Sample sizes of 100,000s of individuals each with millions of genotyped or imputed SNPs and measured phenotypes are now routinely available for analysis and in the near future these will be augmented with full whole-genome sequence data. In addition to gene and gene variant discovery (association between genetic variants and outcomes), such new data also allow the re-visitation of old questions in quantitative genetics, such as the joint distribution of allele frequency and effect size, estimation of genetic (co)variance using new experimental designs, estimation of non-additive genetic variation, genetic variation for phenotypic variance and genetic variation in gene expression and the genetics of epigenetic marks.

Large-scale genome-wide association studies (GWAS) continue to find robustly associated additional variants in biologically meaningful pathways, with ~100 loci identified for schizophrenia and ~700 for height. Only five years ago, there were no robustly associated gene variants with schizophrenia and a handful of Mendelian loci were known to affect human stature. Cumulatively, the newly discovered variants explain approximately 5-15% of phenotypic trait variation the population. What remains unresolved is the frequency distribution of the causal polymorphism/mutations that are responsible for the association signals. We have approached this question by estimating a genetic correlation of trait values across populations using all common SNPs in the genome. If causal variants are themselves common then their linkage disequilibrium (LD) with SNPs will not erode when comparing populations that differ in allele frequencies and LD. We have applied this to data on height and body-mass-index on a combined sample size of up to 50,000 individuals from multiple populations of European descent, all imputed to the 1000 Genomes reference panel. Estimates of genetic correlations are high, ranging from 0.8 to 1.0, for populations with Fst values of 0.01 or smaller.

Multivariate GREML analyses (analogous to GBLUP, but with an emphasis on estimation of (co)variance components) has become routine in human complex trait analysis, and has revealed genetic correlations between diseases that were not known to be genetically related and has confirmed genetic correlations between diseases that were either suspected or inferred from the observed risk to relatives (e.g., schizophrenia and bipolar disorder). Human pedigree studies, unlike those in livestock, suffer from confounding due to common environmental effects, and therefore the experimental design with conventionally unrelated individuals and people not measured for the same traits is appealing.

The advance of genomic technologies also allows genome-wide interrogation of gene expression and methylation at CpG sites. These genomic data are just like other complex traits, and the standard arsenal of quantitative genetic analysis can be applied. We will show results on the heritability of gene expression and gene methylation, the detecting of DNA variants that are associated with variation in these genomic traits, their repeatability across the human lifespan and their correlation with ageing.

Prediction of complex traits from DNA variants, for potential utility in ‘personalised medicine’, is at an early stage of development and application, because firstly not all genetic variation is captured by the SNP arrays used to date and secondly because the effect sizes at individual variants are so small that predictors suffer from too much prediction error. The former might be overcome by imputation to larger reference samples and the latter by having larger experimental sample size for effect size estimation. Nevertheless, polygenic predictors are increasingly used in human genetics research because they allow new hypotheses to be tested and validated, for example pleiotropy between risk of cancers and other traits and the design of experiments based upon selected groups of individuals who are at high and low predicted risk. Interestingly, quantitative genetics has started to make inroads in the social sciences, for example by estimating genetic parameters for educational attainment.