100 years of quantitative genetics theory and its applications: celebrating the centenary of Fisher 1918

Nick Barton

Institute for Science and Technology, Klosterneuburg, Austria

The infinitesimal model

In his 1918 paper, Fisher introduced the analysis of variance, and reconciled Mendelian genetics with continuous biometric variation. He also showed that when variation is due to very many additive genetic loci, selection hardly changes the genetic variance. This has come to be known as the “infinitesimal model”, and is the foundation for practical animal breeding. Yet, organisms are plainly not simply the sum of their genes, and so it is not clear when the use of the infinitesimal model can be justified. I review the history of the infinitesimal model, and show how it can be extended to include dominance, epistasis, and linkage.

Sharon Browning

University of Washington, Seattle, Washington, USA

Identity by descent and the correlation between distant relatives

Identity by descent and the correlation between distant relatives
Related individuals have correlated trait values because they share parts of their genomes identical by descent. Because the genome is inherited in large chunks, punctuated by crossovers, the identity by descent (IBD) occurs in long segments. IBD segments inherited from more distant ancestors tend to be shorter due to more opportunities for crossing over. So-called “unrelated” individuals are in fact distant relatives, and share short segments of IBD. I will describe the detection of segments of IBD in population samples using genome-wide genetic data, and the application of these segments to IBD mapping and to estimating heritability.

Edward Buckler
USDA-ARS, Ithaca, New York, USA

How to get to Plant Breeding 4.0, given that Fisher was right?

As global population continues to grow and standards of living rise, the world food supply needs to almost double protein production. This needs to be accomplished on less land and with a more variable climate. One of the key tools for addressing this challenge is the quantitative genetics of our crops. The synergy of classical quantitative genetics, as developed by Fisher, with low-cost genome-wide genotyping is already laying the groundwork for how to respond, using genomic selection (Breeding 3.0). While these models are extremely successful and useful, they fail to use the incredible wealth of quantitative genomic biology that is being discovered at the nucleotide level. We provide a route to cost effectively identify functional nucleotides, by combining standard quantitative genomics with evolution, chromatin structure analysis, mRNA expression profiling and machine learning. This provides new opportunities to shift from rapid smart selection to the designing of crops that meet the world's future needs.

Heather J. Cordell,
Newcastle University, Newcastle, UK

Regional IBD Analysis (RIA): linkage analysis in extended pedigrees using genome-wide SNP data

Exact calculations for traditional linkage analysis are computationally impractical in large, extended pedigrees. Although simulation-based methods can be used, they require significant computational work. Here I describe Regional IBD Analysis (RIA), a non-parametric linkage method based on comparison of locally and globally estimated identity by descent (IBD) sharing in affected relative pairs (ARPs), for these circumstances. In RIA, genome-wide SNP data are used to calculate "global" expected

IBD sharing probabilities specific to each affected relative pair, against which “local” IBD sharing probabilities, estimated using SNP data within a window of pre-specified width, can be compared. The global and local IBD sharing probabilities are then used to construct a non-parametric test of linkage in each window.
The proposed method is illustrated through application to real nuclear family data from a study of primary vesicoureteric reflux, and to real and simulated data based on large extended pedigrees from a Brazilian study of visceral leishmaniasis. RIA successfully detected the known or simulated linkage signals, with a significant reduction in computational time and resources compared with traditional methods. Our method is particularly useful in studies involving large extended families, in which traditional linkage analysis is not feasible. Additionally, because it does not require prior knowledge about familial relatedness, our method has the additional advantage of being robust to pedigree misspecification and can be used even in absence of known pedigree information.

Michael E. Goddard

University of Melbourne and Agriculture Victoria, Melbourne, Australia

The Fisher Memorial Lecture: The Genetic Architecture of Complex Traits

Despite the complete lack of knowledge of the physical nature of genes, a century ago Fisher showed how Mendelian inheritance of many loci could explain the observed statistical properties of quantitative or complex traits, such as the resemblance between relatives. As the number of loci increases, we obtain the infinitesimal model, which has been the basis for our understanding of the genetics of complex traits and very successful practical applications in animal and plant breeding. However, there was little knowledge of the genes causing variation, or even of their number, and limited understanding of the evolutionary forces that controlled the ubiquitous genetic variation we observed. In the last decade our knowledge of the genetics of complex traits has been revolutionized by the availability of data on thousands to millions of single nucleotide polymorphisms (SNPs). The purpose of this lecture is to summarise what we

have learnt about the architecture of complex traits, and the evolutionary forces that bring this about.

A simple but surprising result is that most quantitative genetic variation is due to very many polymorphisms, each with a tiny effect on the trait, and segregating in the population at moderate allele frequencies. For instance, there are approximately 10,000,000 sites in the genome where a mutation can affect a typical quantitative trait. Each generation, mutation generates new variation; some of the new mutant alleles have a large effect, but selection keeps them very rare. Most of the variation is caused by mutations of small effect that are almost neutral, and hence segregate at moderate allele frequencies. However, occasionally a mutation is favoured by selection and while it segregates generates a large variance. This is most common when the environment changes greatly so that the direction of selection on some mutations reverses. This new understanding explains many previously puzzling results, such as the linear response to artificial selection and the failure to find the genes causing variation in complex traits.

The ability to genotype thousands of SNPs at moderate cost has been utilised in methods of genomic selection or genomic prediction, which predict the genetic value of individuals for a trait based on SNP genotypes. This method is revolutionising animal and plant breeding and will be important in human medicine, for instance in personalised medicine.

Jarrod Hadfield
University of Edinburgh, Edinburgh, UK

Hamilton's rule in multiple dimensions

Two frameworks that exist for modelling the evolution of traits involved in social interactions are kin selection models and indirect genetic models. In kin selection models, an actor directly affects the fitness of the recipient; in indirect genetic models, an actor affects the phenotype of the recipient and (as an indirect consequence) the fitness of the recipient. Using causal analysis, I show how the parameters of indirect genetic models relate to the costs and benefits in Hamilton's rule, and derive a multi-trait version of Hamilton's rule that has close connections to the Lande equation for selection on a multivariate quantitative trait.

When social interactions involve multiple traits, Hamilton's simple rule is shown to fail, unless the costs and benefits of all other traits balance so that they are at evolutionary equilibrium. When they are out of equilibrium, the evolutionary process cannot be understood without considering the system of traits as a whole and the underlying genetic architecture. Nevertheless, compact expressions can be derived using matrix algebra that shed light on how the system as a whole should evolve when social interactions exist.

Richard Mott

University College London, London, UK

Genomic rearrangements in *Arabidopsis* considered as quantitative traits

To understand the population genetics of structural variants and their effects on phenotypes, we developed an approach to mapping structural variants that segregate in a population sequenced at low coverage. We avoid calling structural variants directly. Instead, the evidence for a potential structural variant at a locus is indicated by variation in the counts of short-reads that map anomalously to that locus. These structural variant traits are treated as quantitative traits and mapped genetically, analogously to a gene expression study. Association between a structural variant trait at one locus, and genotypes at a distant locus indicate the origin and target of a transposition.
Using ultra-low-coverage (0.33x) population sequence data from 488 recombinant inbred *Arabidopsis thaliana* genomes, we identified 6502 segregating structural variants. Remarkably, 25% of these involved transpositions. While many structural variants could not be delineated precisely, we validated 83% of 44 predicted transposition breakpoints by PCR. We showed that specific structural variants may be causative for quantitative trait loci for germination, and resistance to infection by the fungus *Albugo laibachii*. Furthermore, we showed that the heritability attributable to read-mapping anomalies differs from, and can exceed, that due to standard genetic variation. Genes within structural variants are also more likely to be silenced or dysregulated. This approach complements the prevalent strategy for structural variant discovery in fewer individuals sequenced at high coverage. It is generally applicable to large populations sequenced at low coverage, and is particularly suited to mapping transpositions.

Josephine Pemberton
University of Edinburgh, Edinburgh, UK

Quantitative genetics of free-living populations: successes and challenges

The heritability of quantitative traits and genetic correlations between them are vital parameters for understanding evolutionary potential and constraints. Estimating these quantities *in situ* in the wild is important, since trait variance due to environmental heterogeneity is generally large. Since the arrival of marker-based parentage inference and the animal model, this field has burgeoned, but it now faces a downstream challenge: predictions and estimates for the response to selection rarely match. This is indicative of some kind of missing information. I will discuss the various possible sources for these mismatches and how they may be resolved.