

How to GWAS - ConGen 2022

Eryn McFarlane

2022-09-01

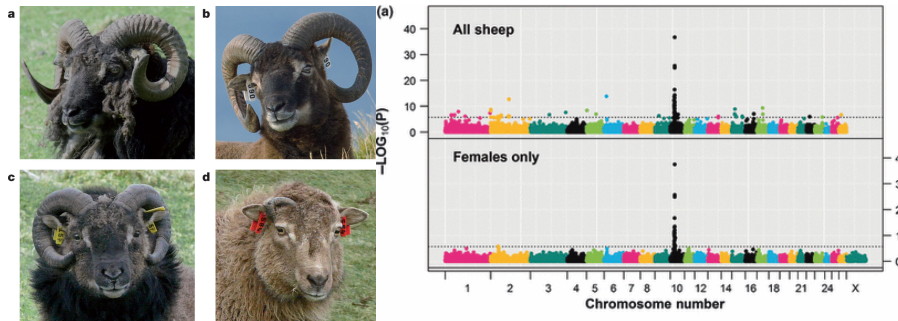
What is GWAS?

Genome Wide Association Study

- ▶ $y = a + bx + g + e$
- ▶ y: phenotype
- ▶ b: fixed effect
- ▶ x: the SNP predictor coded as 0, 1, 2
- ▶ g: the polygenic random effect (e.g. Kinship matrix)
- ▶ e: the error term

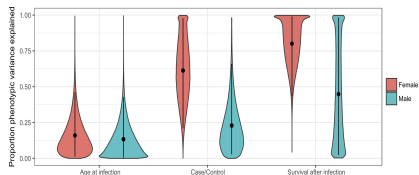
Soay Sheep Horns RXFP2 genotype

Johnson et al. 2011 Mol Ecol, Johnson et al. 2013 Nature



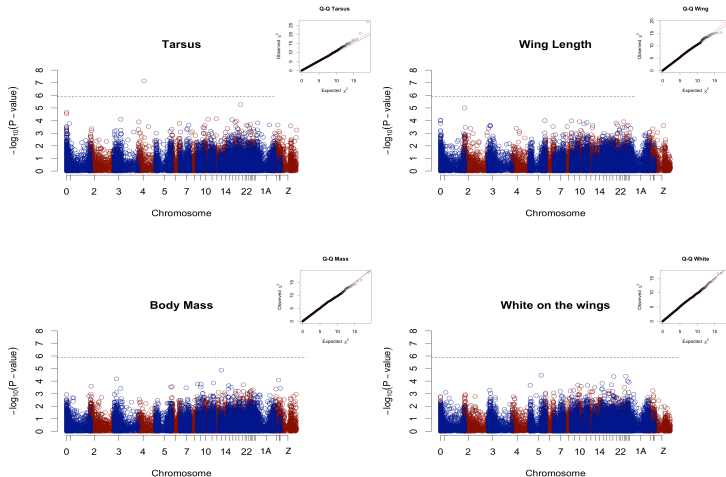
- ▶ 486 sheep
- ▶ 35 831 SNPs
- ▶ one huge effect locus, explains 76% of the variation in horn size

How often do people find huge effect loci using GWAS?



- Darwin's finches - sticklebacks - atlantic salmon - tasmanian devils face cancer 5 SNPs explain ~61% of variation in female survival

What are the expectations for a GWAS?



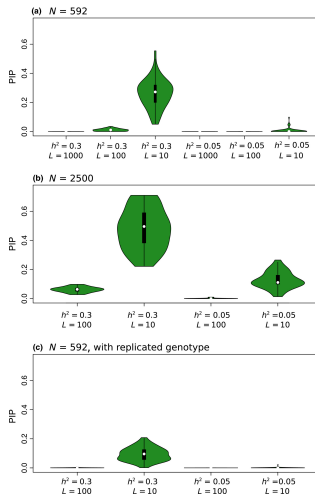
- ▶ No significant SNPs for morphological traits
- ▶ 3 SNPs explaining $\sim 3\%$ of the variation in clutch size

What are the expectations for a GWAS?

- ▶ Rockman 2012 - All that's gold does not glitter.

Some considerations for GWAS

- heritability of the trait
- genetic architecture of the trait (unknown!)
- Linkage Disequilibrium
- number of individuals sampled



Kardos 2015 MER, Gompert 2017 Mol Ecol

Some solutions

- ▶ Sample more individuals!
- ▶ replicated samples of individuals (RepeatABEL, RIP)
- ▶ Admixture mapping instead of within population GWAS
- ▶ sample in closed populations (but know this limits generalizability)

What is Admixture mapping?

Deer Case study

So what are we doing today?

- ▶ Simulated GWAS, using GEMMA