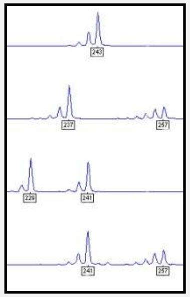
Farm animal genetic diversity: measuring neutral variation and detecting selection

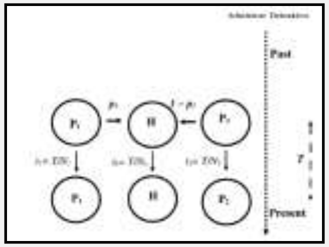
MW Bruford1, T Perez1, GA Juma1, A Sheppy2, L Chikhi3 and TC Bray1

1 Cardiff University, Cardiff, Wales, UK;
2 Cobthorne Trust, Somerset, England;
3 Paul Sabatier University, Toulouse, France

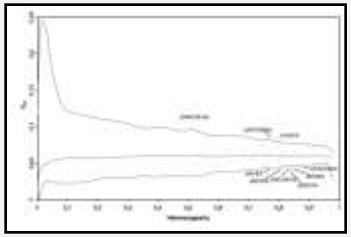
Introduction



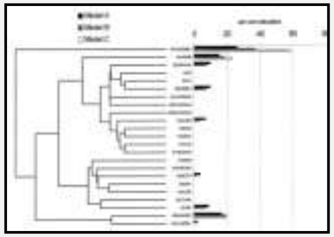
Recent history of molecular markers and FAnGR diversity studies



Livestock as a special case: evolutionary history, patterns of diversity and unusual population processes



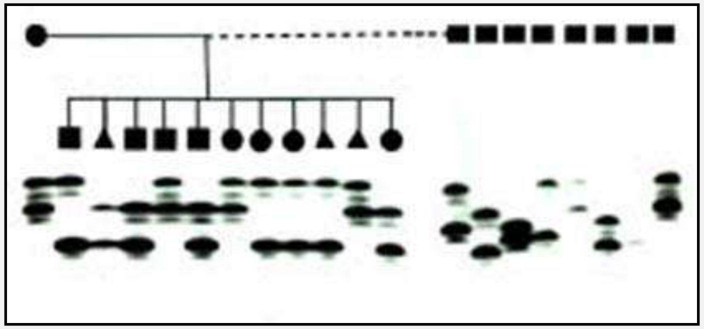
New prospects for conservation: detecting and weighting markers under selection



Current issues: data usage in conservation prioritisation – prospects and problems

Markers 1

DNA diversity studies (mid to late 80s) – kick started by DNA fingerprinting and STR markers generated for the genome projects: thousands to choose from…

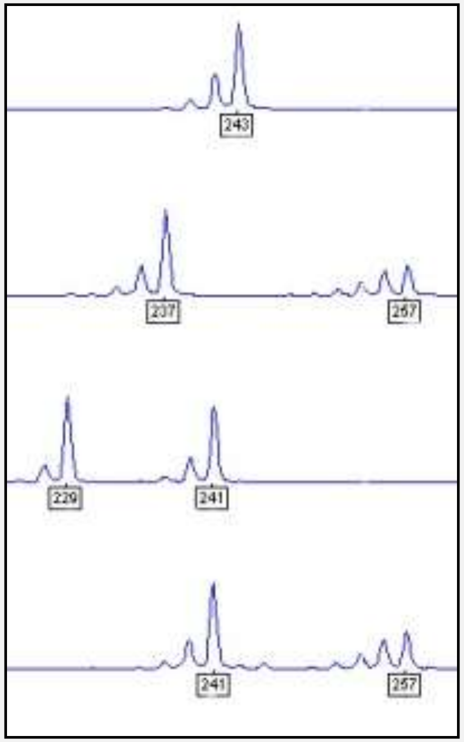


High polymorphism and informativeness

Could be used for studies of parentage, inbreeding, genetic diversity and distance and individual identity…

Microsatellites (STRs) became the most heavily used markers by early ‘90s

Markers 2



Early on the FAO realised that STRs would become the mainstay for AnGR diversity studies and quickly moved to produce recommended marker lists for the main domestic species.

Most of these are still used (updated 2004)…

Still many datasets have been produced using different subsets of these markers

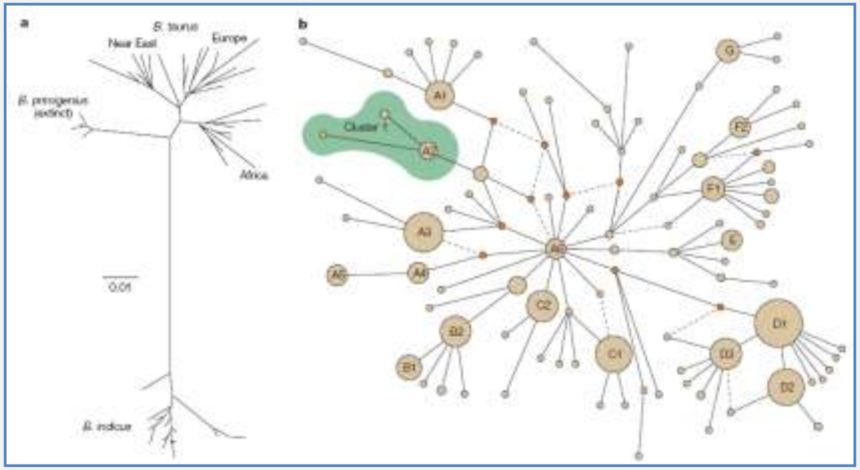
Ongoing EU-funded project (GLOBALDIV) aimed at normalising many of these datasets..

FAO also recommended a variety of analytical approaches which could be applied – most practitioners would recognise these as having been replaced…

Genealogical simulation and Bayesian clustering approaches are now more widely adopted than allele frequency-based methods

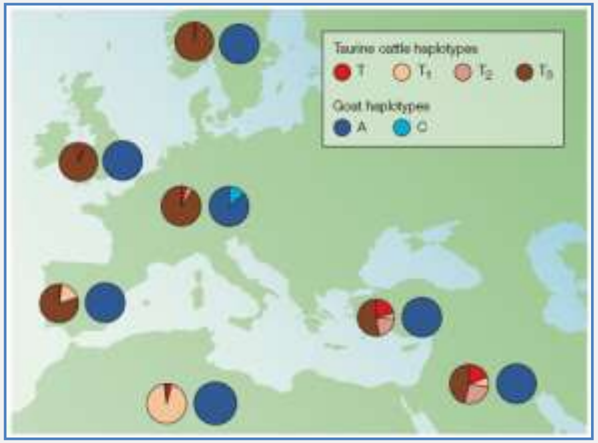
Markers 3

Meanwhile, understanding the origin of domestic livestock species led to the adoption of mitochondrial DNA sequencing – leading to surprising results..



Multiple mitochondrial lineages the rule – implying multiple domestication events…

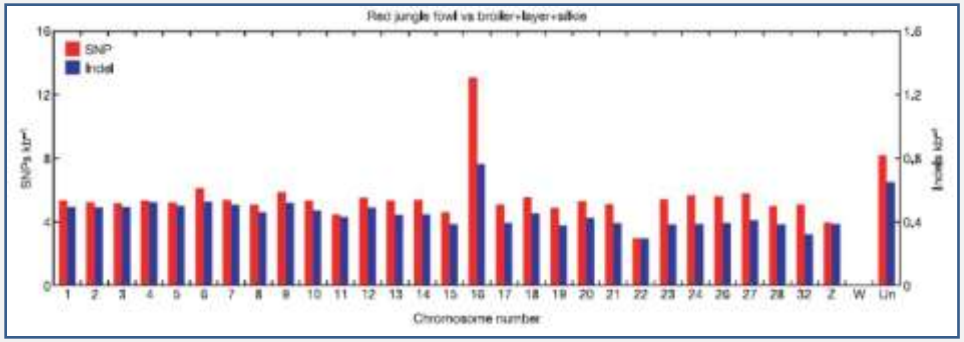
General (but not exclusive) pattern of lower diversity in regions farthest from the centre of domestication



Advances in ancient DNA analysis is enabling much more accurate inference of evolutionary origins of domestication (e.g. pigs, cattle)

Markers 4

Genome-wide analysis (e.g. via STR and AFLP data and huge SNP datasets) promise to yield much more accurate estimates of demographic parameters and fine-scale detection of selective sweeps.



Chicken SNPs by chromosome…

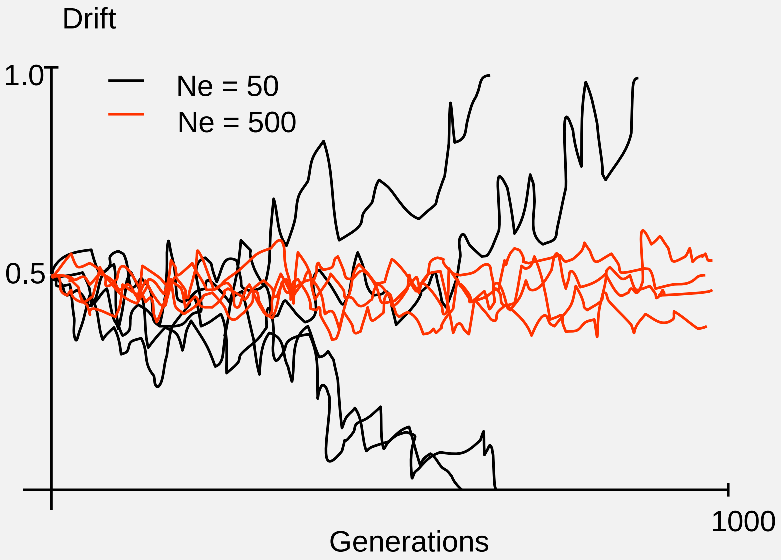
Single and multiple genes under direct selection such as TYRP-1 (right) are becoming more widely used…

Figure 1. Dark (left) and light (right) coat colour morphs in Soay sheep.

High throughput SNP genotyping (multiple platforms) likely to supercede other technologies

Livestock 1

Livestock evolution is highly unusual compared to most species: interplay between natural and artificial selection and how this has varied over time scales (10,000 BP - last few hundred years)

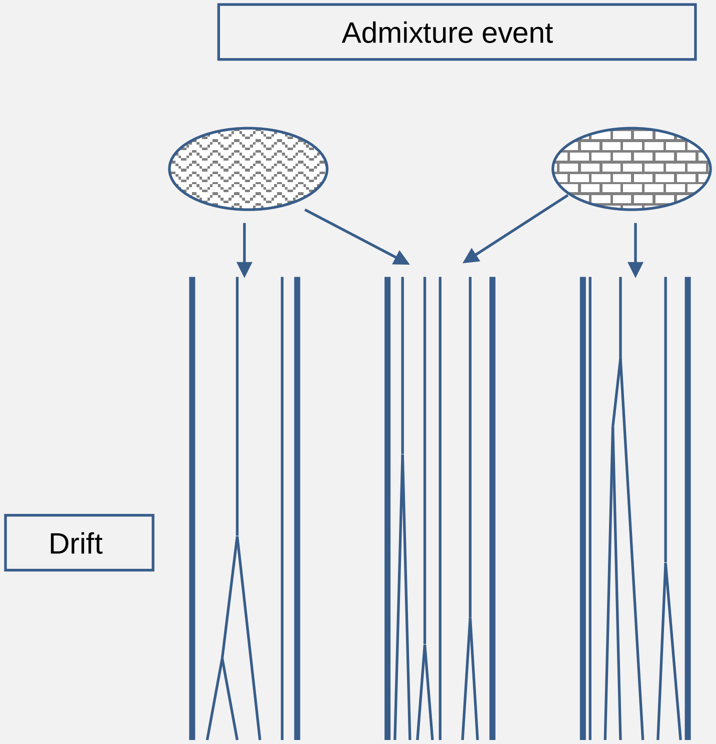


Founder effect, drift and admixture are important components (among others) of breed evolution

Within-breed diversity is often a goal for conservation programs, but how this diversity arrives is very important…

Livestock 2

Livestock breeds sometimes evolve in isolation but may be improved by specific upgrading programs involving selection of a few genes: poses problems for analysis

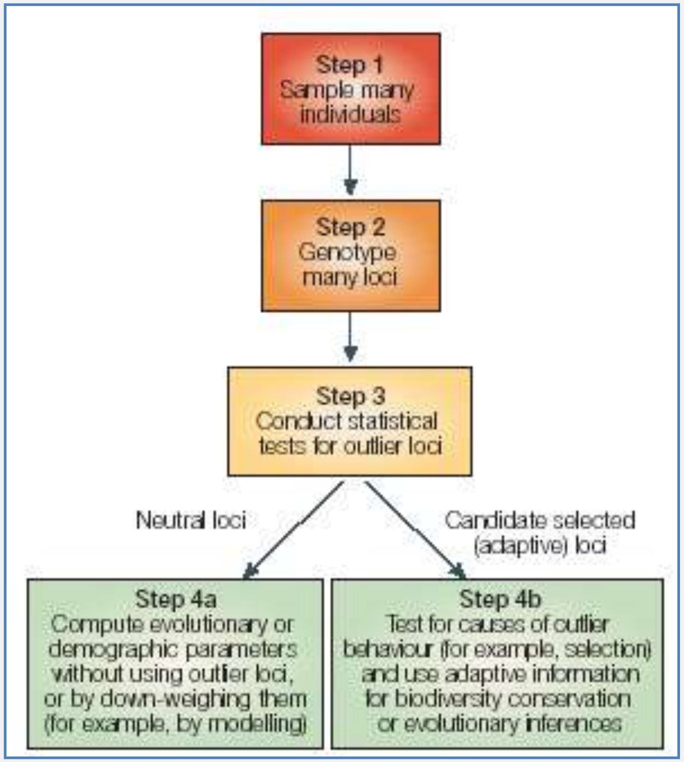


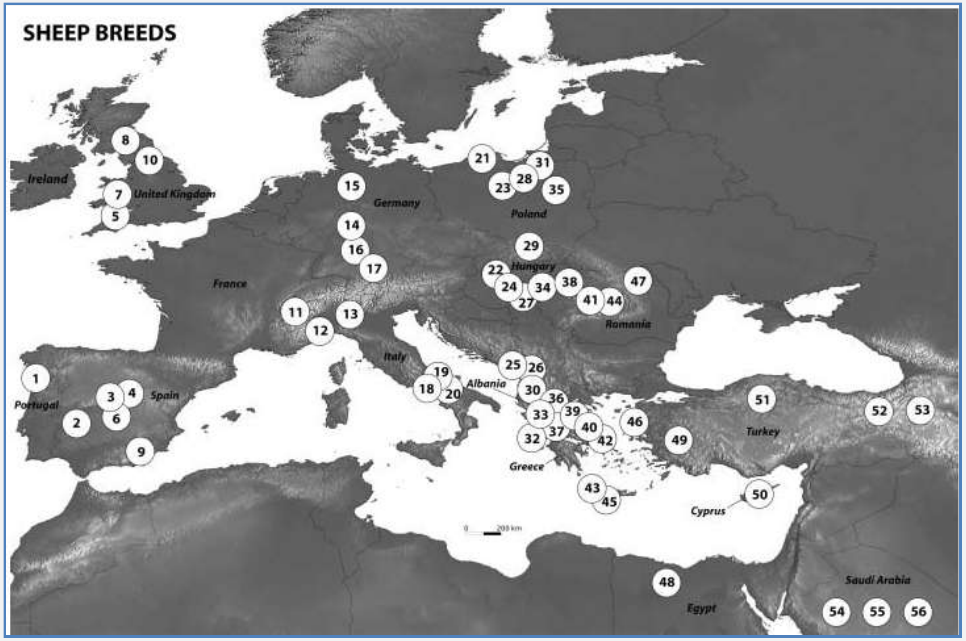
Recurrent admixture can complicate matters further…

Hence potential disassociation between genotype and phenotype and lack of correlation between uniqueness and utility...

Prospects

A number of promising methods have emerged over the last 5 years to identify and incorporate markers which show evidence of selection in diversity studies and conservation prioritisation...

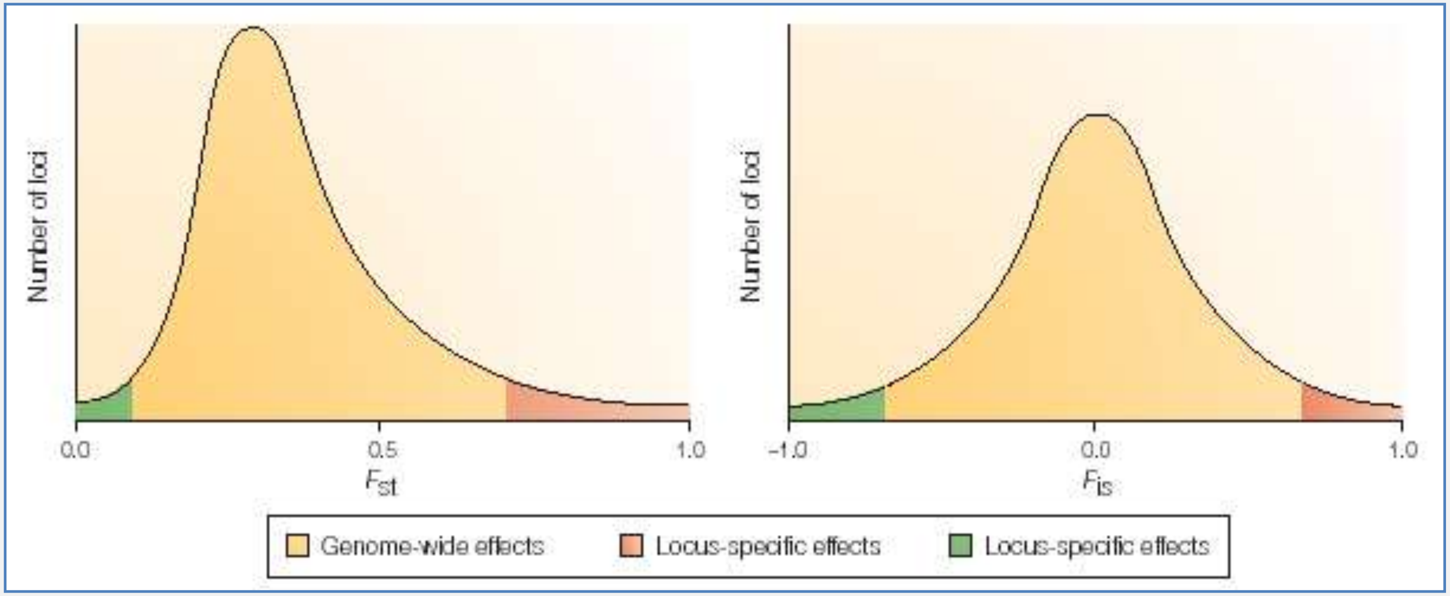


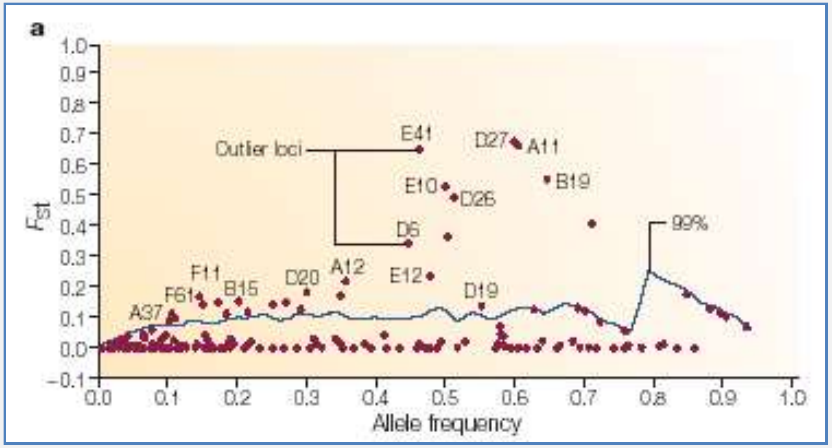


Population genomics (outlier analysis)

Landscape genomics

Outlier analysis

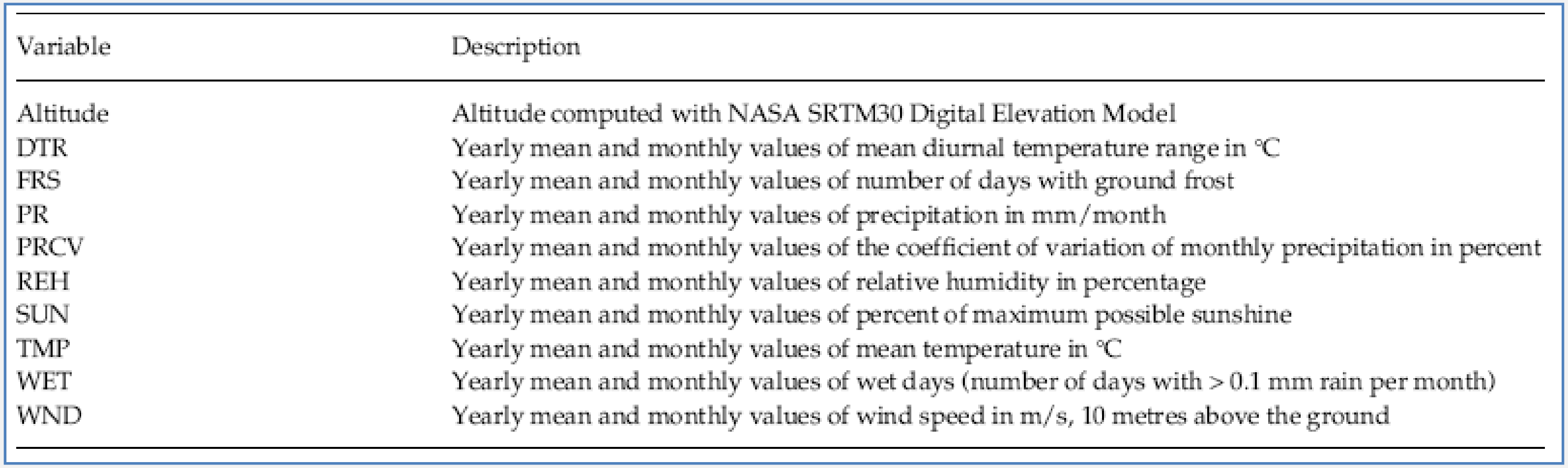




This approach has proved to be remarkably robust to different patterns of genetic structure...

Landscape genomics

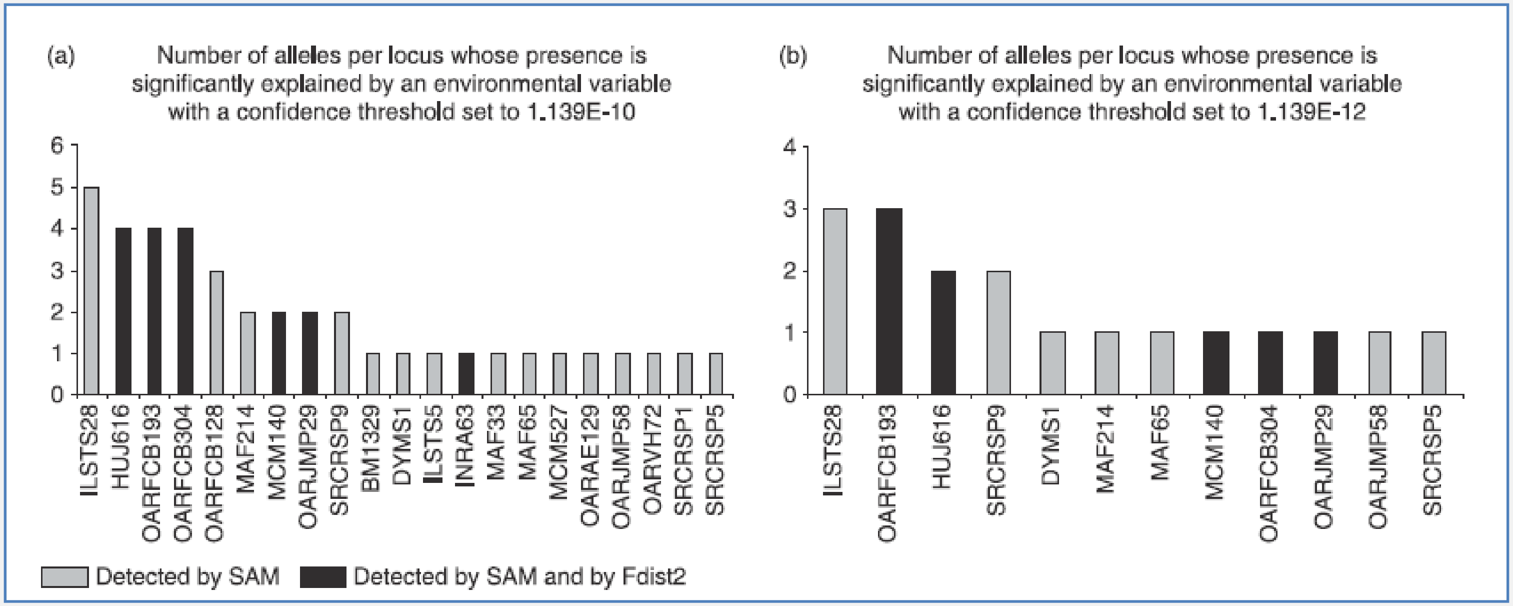
Landscape genomics uses multiple logistic regressions to test for association between marker frequencies and environmental (or other) variables.

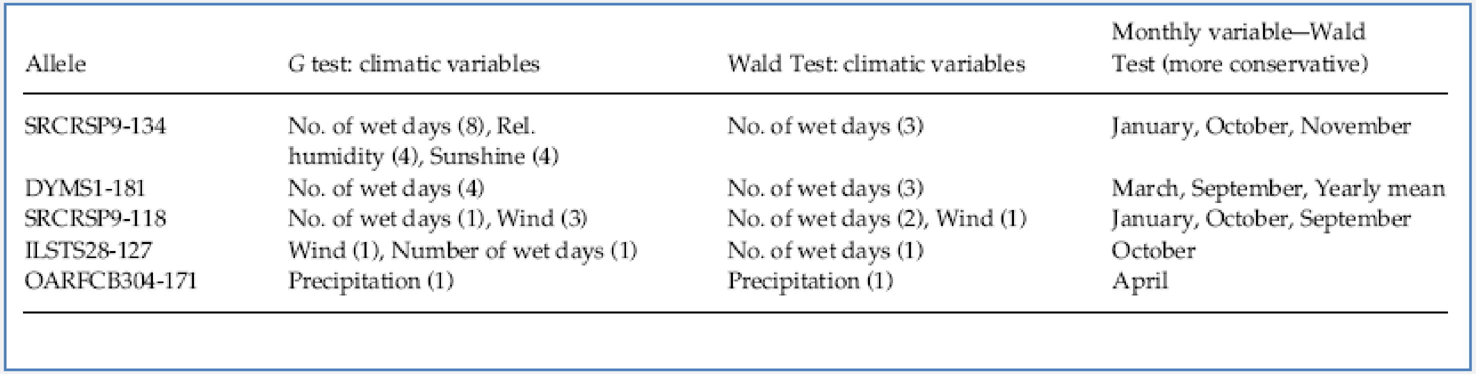


Variable	Description
Altitude	Altitude computed with NASA SRTM30 Digital Elevation Model
DTR	Yearly mean and monthly values of mean diurnal temperature range in °C
FRS	Yearly mean and monthly values of number of days with ground frost
PR	Yearly mean and monthly values of precipitation in mm/month
PRCV	Yearly mean and monthly values of the coefficient of variation of monthly precipitation in percent
REH	Yearly mean and monthly values of relative humidity in percentage
SUN	Yearly mean and monthly values of percent of maximum possible sunshine
TMP	Yearly mean and monthly values of mean temperature in °C
WET	Yearly mean and monthly values of wet days (number of days with > 0.1 mm rain per month)
WND	Yearly mean and monthly values of wind speed in m/s, 10 metres above the ground

This approach has potential to establish hypotheses about ecological factors that could exert the selection pressure responsible.

A comparison..





Allele	G test: climatic variables	Wald Test: climatic variables	Monthly variable—Wald Test (more conservative)
SRCRSP9-134	No. of wet days (8), Rel. humidity (4), Sunshine (4)	No. of wet days (3)	January, October, November
DYMS1-181	No. of wet days (4)	No. of wet days (3)	March, September, Yearly mean
SRCRSP9-118	No. of wet days (1), Wind (3)	No. of wet days (2), Wind (1)	January, October, September
ILSTS28-127	Wind (1), Number of wet days (1)	No. of wet days (1)	October
OARFCB304-171	Precipitation (1)	Precipitation (1)	April

Econogene sheep dataset: Joost et al *Molecular Ecology* (2007)

Alternate route to prioritisation..

Population adaptive index (PAI): a link between genome-scan data and conservation prioritisation. Not applied (at least *in lit*) to livestock yet.

Measures the % of divergent outlier loci per population which are significantly diffrentiated from others in >2 comparisons.

Levels of neutral and adaptive genetic diversity (measured through the Population Adaptive Index [PAI]) in common frog and Austrian dragonhead populations (from Bonin *et al.* 2007).

Population	Proportion of polymorphic loci	Nei's gene diversity	PAI
Common Frog			
AI	0.52	0.18	0.07
CO	0.69	0.23	0.21
PP	0.74	0.25	0.14
RM	0.68	0.23	0.43
TE	0.68	0.23	0.29
TI	0.66	0.22	0.29
Austrian dragonhead			
BE	0.58	0.21	0.43
CH	0.75	0.25	0.29
ES	0.61	0.22	0.21
FO	0.58	0.20	0.14
LA	0.60	0.21	0.14
RE	0.64	0.22	0.14
VA	0.69	0.21	0.21

Complementarity (as with Weitzman and Moritz) can in principle be used to identify those populations which possess different (or the same) outliers and compare these patterns of differentiation with 'neutral' data... This would seem to offer an efficient option for livestock..

Bonin et al. *Conservation Biology* 2007

Issues 1

Considerable debate on the best performing measures of diversity and differentiation to use for prioritisation analysis

Expectations for livestock in terms of statistical versus biological significance and its context without detailed knowledge of the demographic history of the breed (e.g. Dexters)

Global diversity patterns - at what geopolitical level is prioritisation appropriate?

(..*e.g. distance from centres of domestication and populations with human influence and transport*)

Weighting neutral versus adaptive diversity (e.g. PAI versus other molecular indices) and combination with other measures of uniqueness...

Comparison and combination of the above factors could be a complex process.. Detailed studies require information sometimes lacking... Starting point??