Classical selective sweeps revealed by massive sequencing in cattle

Saber Qanbari

Department of Animal Sciences
Animal Breeding and Genetics Group
Georg-August-University Göttingen, Germany

64th EAAP Meeting
August 26th-30th 2013, Nantes, France

Acknowledgements

Synbreed

In collaboration with:

H. Simianer

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

H. Pausch
S. Jansen
R. Fries

TUM
TECHNISCHE UNIVERSITÄT MÜNCHEN

M. Somel
R. Nielsen

Berkeley
UNIVERSITY OF CALIFORNIA

T.M. Strom, T. Maitinger, T. Wieland, S. Eck, E. Graf., A. Benet-Pages

HelmholtzZentrum münchen
German Research Center for Environmental Health

Outline

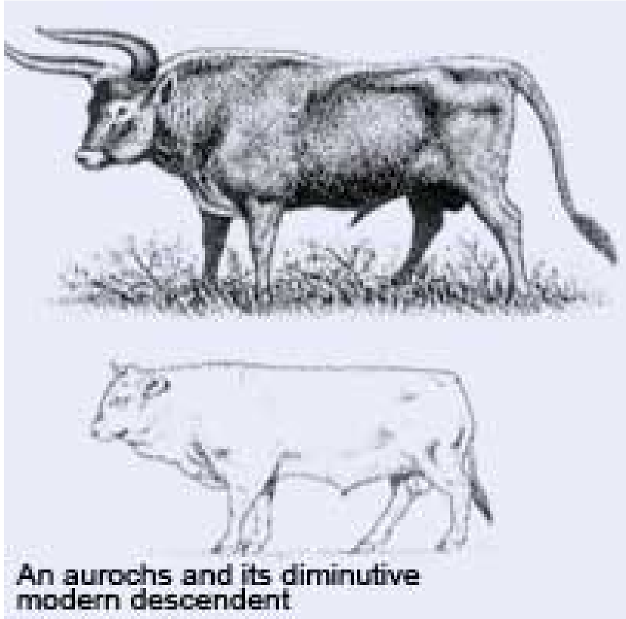
- Background
  - Methods to detect [positive] selection
    - Local variability
    - Allele frequency spectrum
    - Haplotype based approaches
- Case study in cattle
  - Linkage Disequilibrium
  - Selection signature analyses
  - GWAS
  - Conclusions

Background: domestication of cattle

- Cattle is domesticated from aurochs (*B. primigenius*) 10’000 BCE
- Major impact on human civilization
- Domestication/breed formation created > 800 breeds
- Modern cattle have different appearance and production level, milder temper, … .

Motivation:

- How human interaction has influenced the cattle genome?
- What proportion of the genome is involved in domestication/adaptation?
- Which traits are involved?
- …



An aurochs and its diminutive modern descendent

http://www.strideguides.com/unearthingthepast/ web-content/aurochs.html

Detecting positive selection

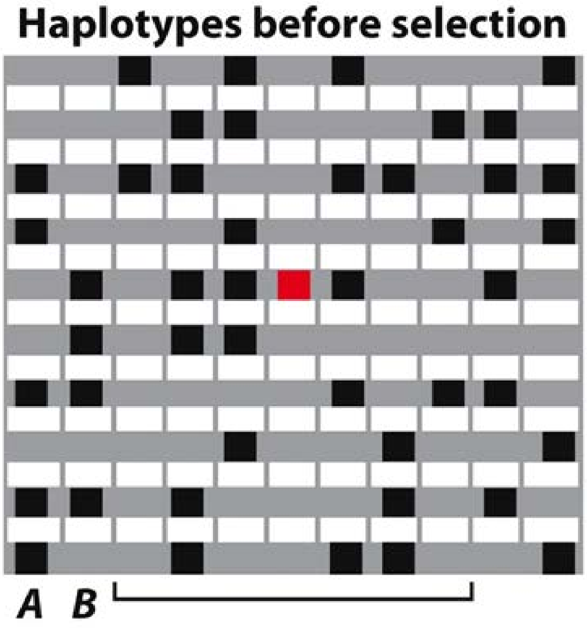
- Testing candidate genes (very few proven examples)
- GWAS for marker-trait associations (difficult for multi-genic complex trait and fixed variation, recording problems …)
- Looking for patterns of selection
  - Difference between species
    - High proportion of function altering mutations (dN/dS)

      Age: millions of years
  - Within-species variation
    - Differences between populations (e.g., Fst, XPEHH)
    - **Intra-population analysis (e.g., Heterozygosity, CLR and iHS)**

Signatures left by positive selection:

'selective sweep': Hitch-hiking theory of Maynard-Smith and Haigh (1974)

before



1) Low local variability
2) Excess of frequent derived and rare alleles
3) Long-range (unbroken) haplotypes

after

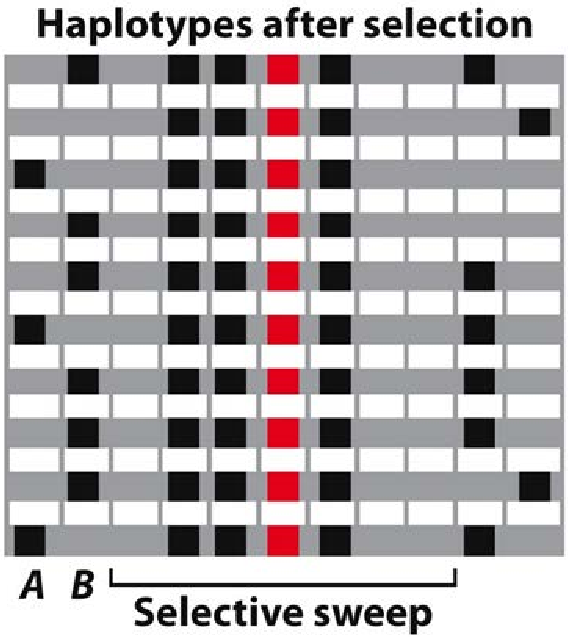
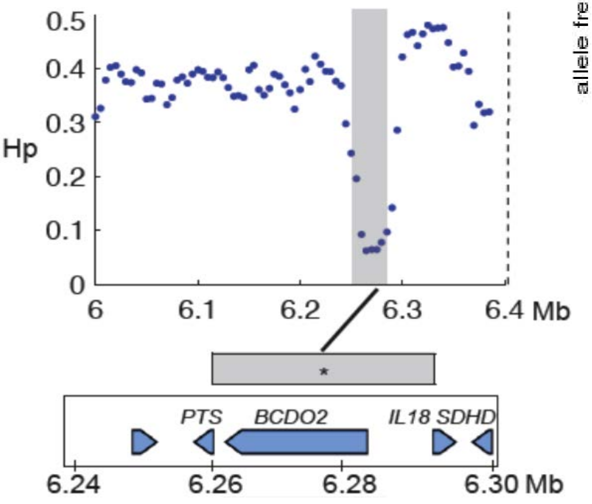


Figure 18-22
Introduction to Genetic Analysis, Tenth Edition

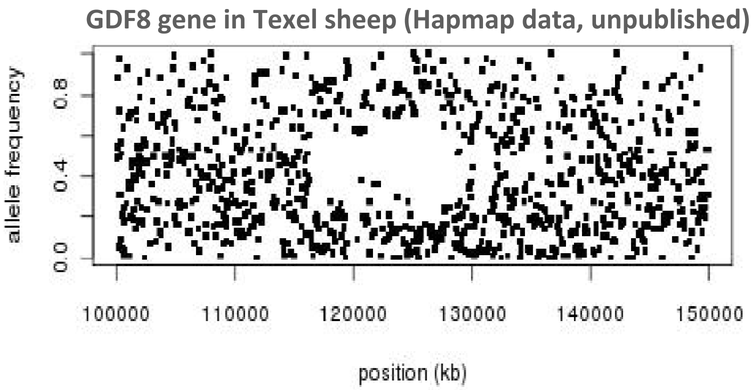
Testing positive selection:

I. Low diversity

e.g., Pooled Heterozygosity (Rubin et al. 2010)

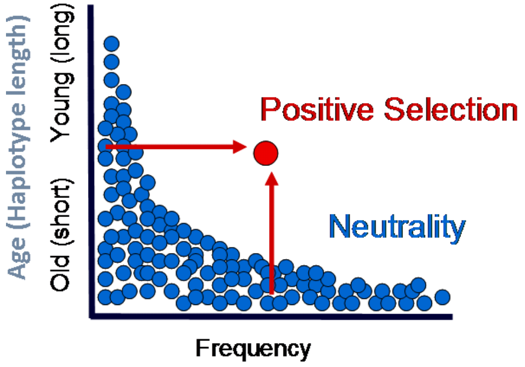


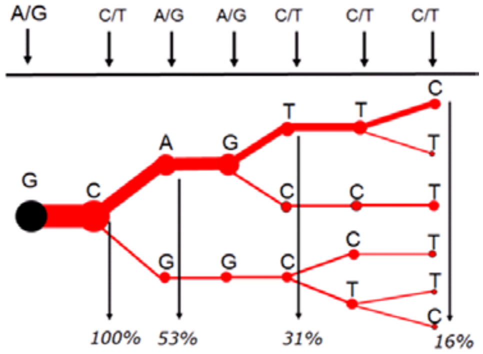
II. Allele frequency spectrum



e.g., Composite of likelihood ratio test (CLR) (Nielsen et al. 2005):

III. Haplotype analysis



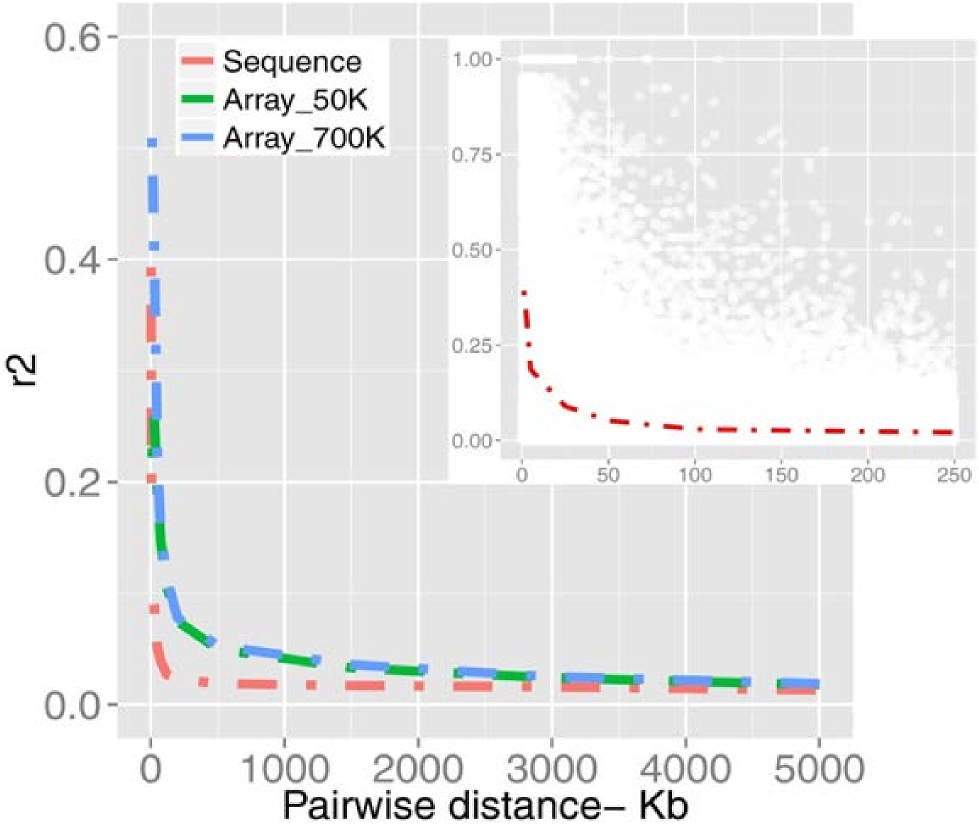


e.g., EHH/iHS, (Sabeti et al. 2002 & Voight et al. 2006)

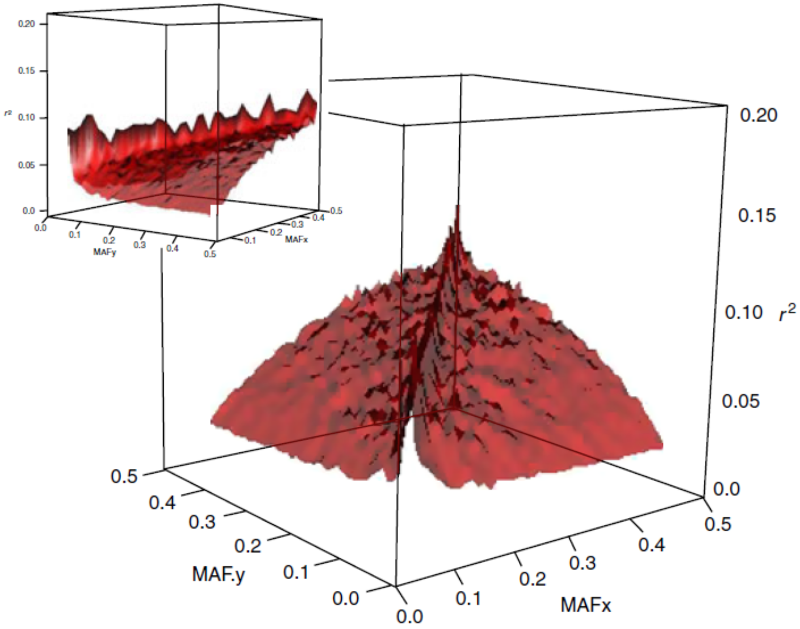
Genetic material: (from Jansen et al. 2013)

- 43 Fleckvieh animals were sequenced (coverage ≈ 7 folds)
- Mean accumulative genome coverage: 261.9 ± 56
- Number of SNPs: 15’060’623
- Mean Inter-marker space: 166.8 ± 393 bp
- Mean allele frequency: 0.27 ± 0.27

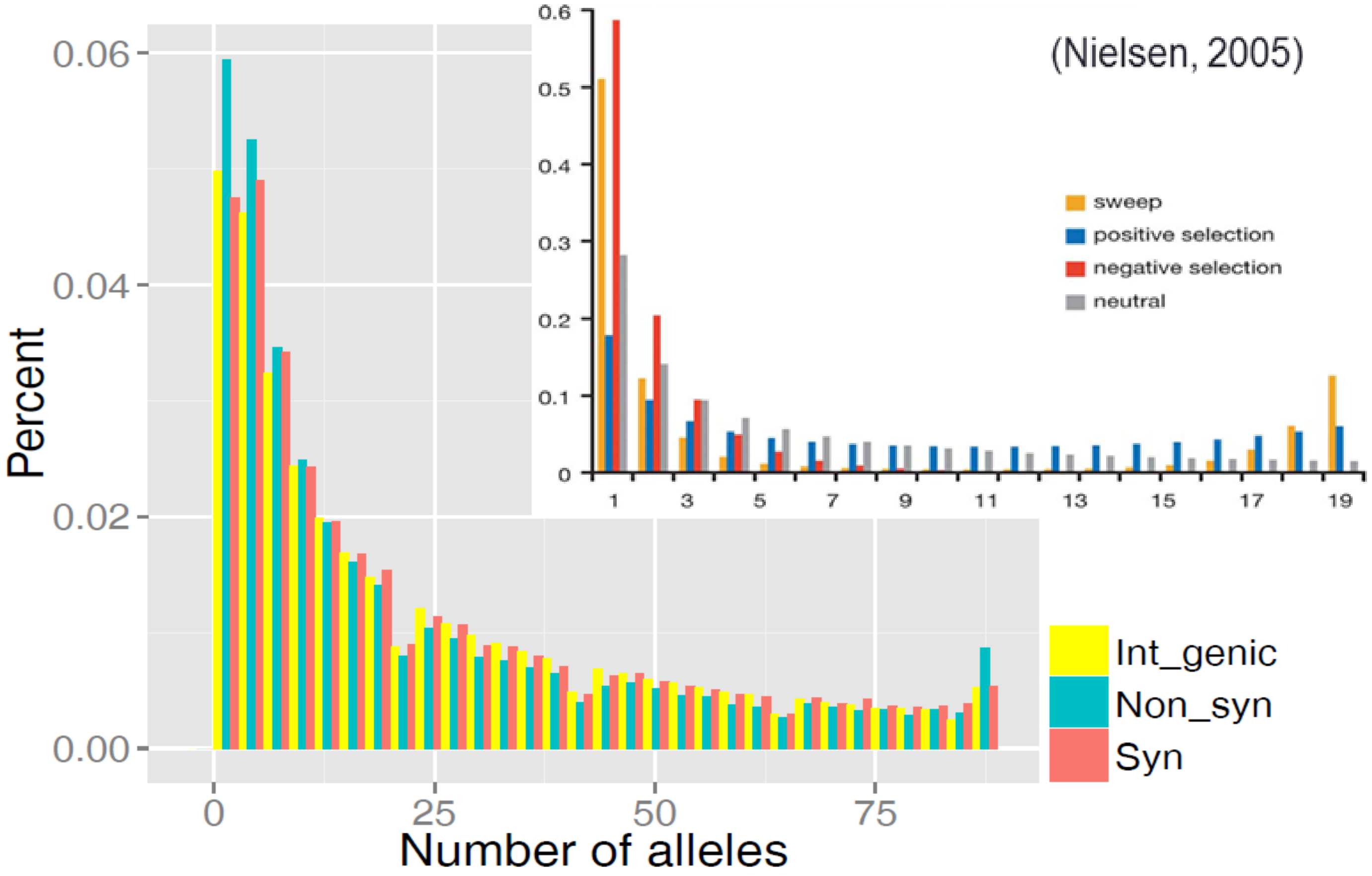
Linkage Disequilibrium



Qanbari *et al.* (2010)



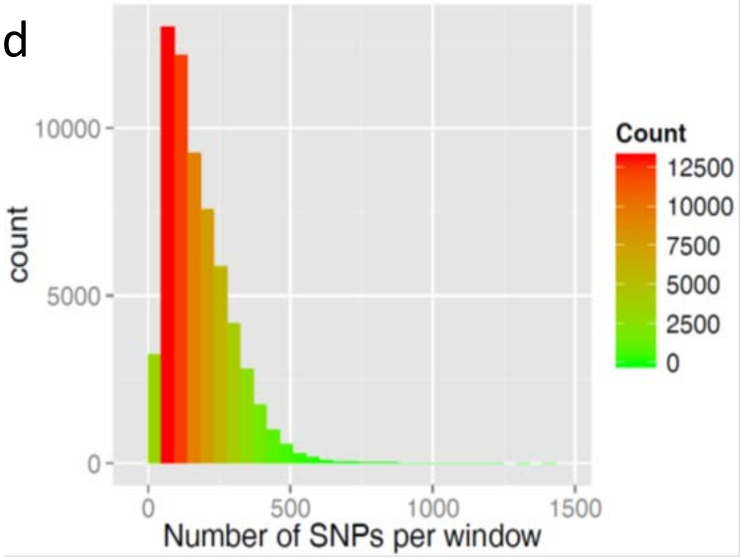
Allele Frequency Spectrum



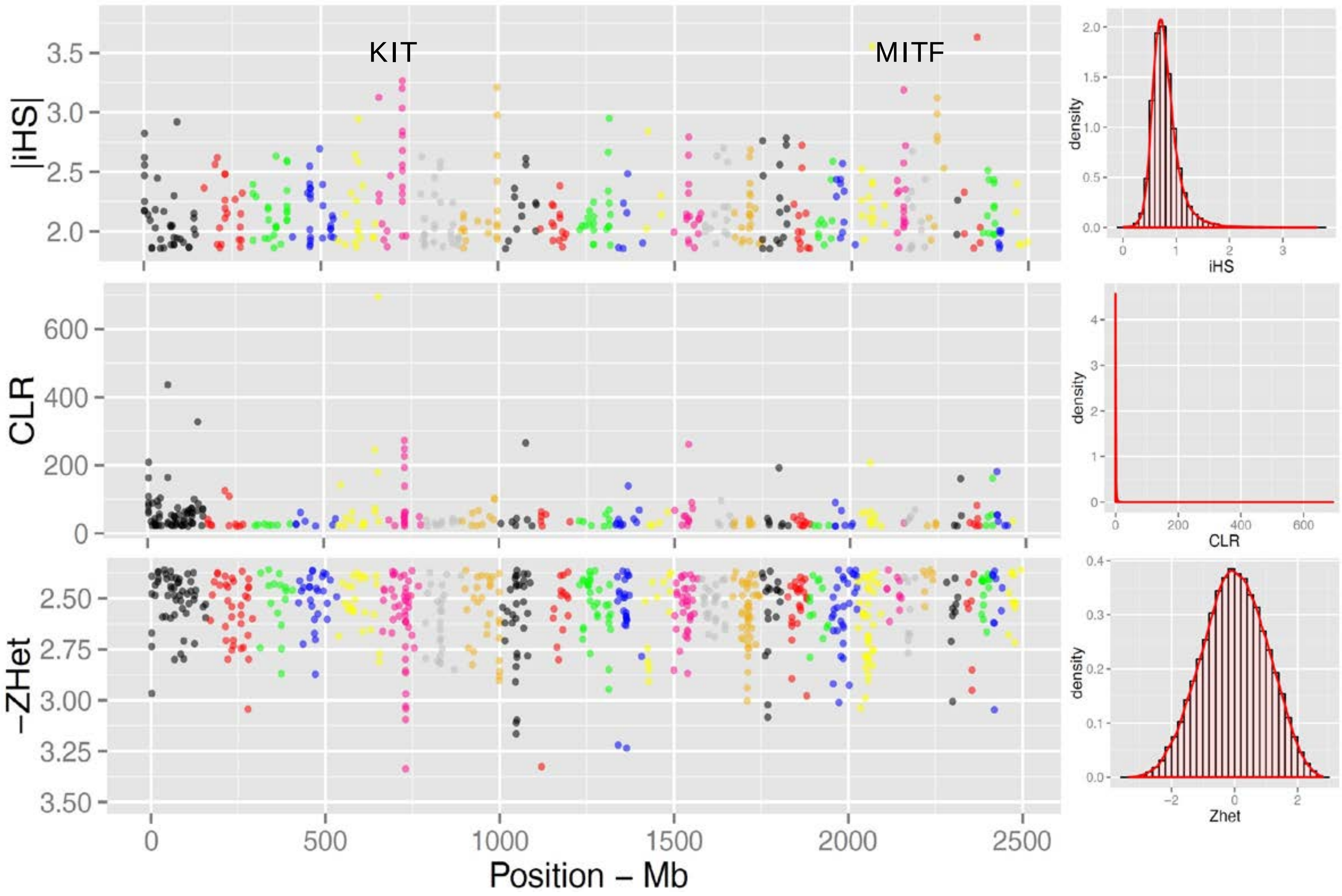
(Nielsen, 2005)

Mapping strategy

- Genome was scanned by |iHS|, CLR metrics (ongoing + fixed selection)
- Heterozygosity, Tajima D, Fay & Wu H and |nSL| were also measured for further inspection
- [except CLR] metrics were accumulated in windows of 40-Kb
- In all 62’196 windows/values were estimated (in average 240 SNPs per window)
- Only top 1% (genome-wide percentile) signals considered as selection candidates
- 106 candidate sweeps were detected



Genomic distribution of candidate regions

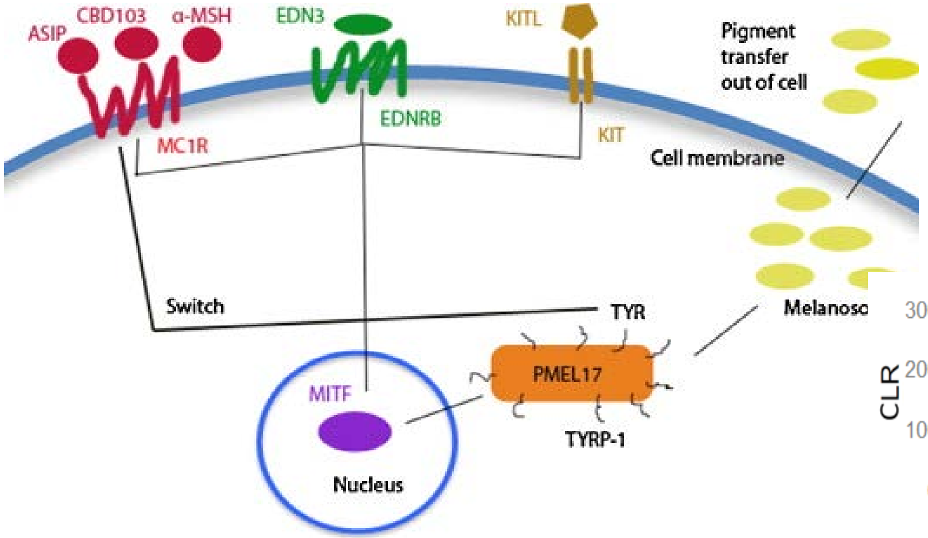


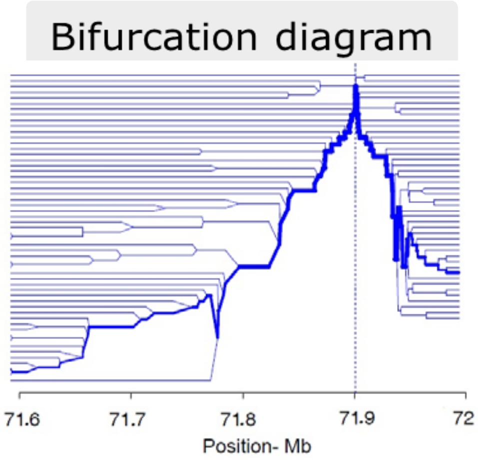
Candidates genes under selection …

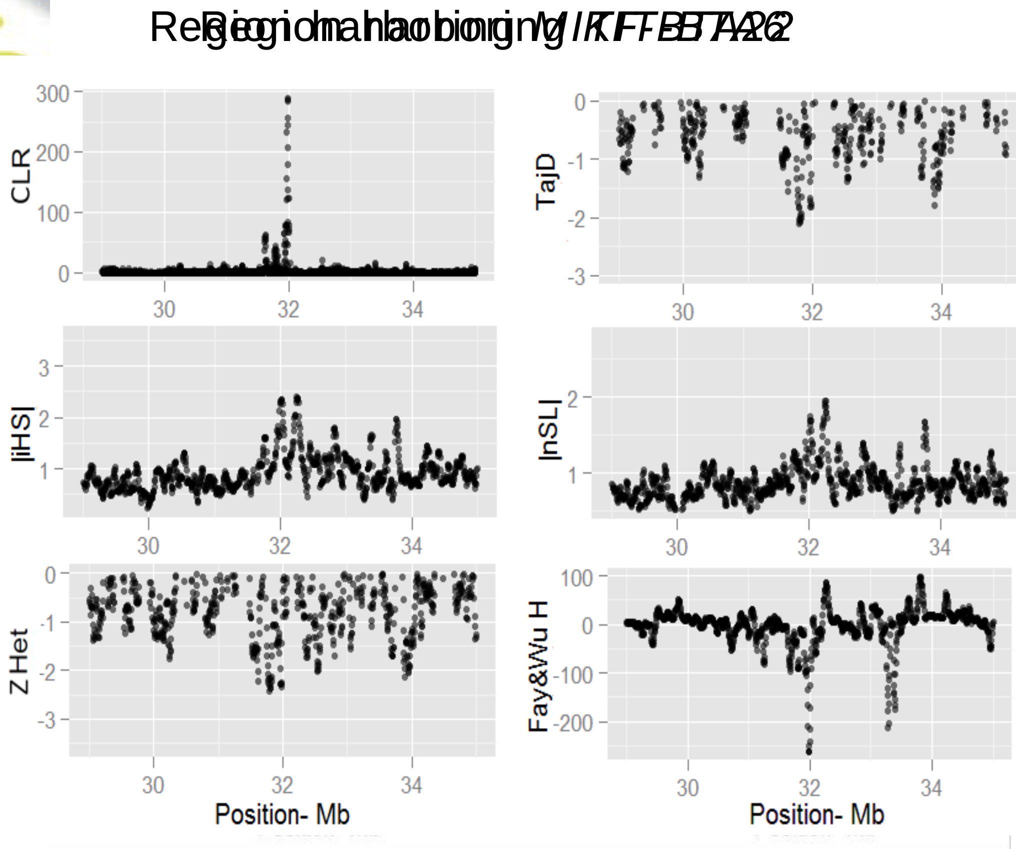
Chr	Gene	Position (bp)	$P_{\|iHS\|}$	P_{CLR}	Function / association
6	KIT/KDR	72’087’087	0.00029	0.00011	Pigmentation
22	MITF	32’017’564	0.00172	0.00709	pigmentation
18	MC1R	14’785’816	0.00785	0.00156	pigmentation
21	NRG4	32’012’049	0.00368	0.00221	Pigmentation
2	Erbb4	99’904’366	0.00203	0.00855	Pigmentation
17	ULBP3	40’044’391	0.00691	0.00917	Pigmentation, sudden whitening of the hair
15	OR51A7	50’749’853	0.87230	0.00048	Olfaction perception
8	OR13C8	96’134’318	0.57280	0.00043	Olfaction perception
7	OR2	15’246’693	0.00502	0.00194	Olfaction perception
7	OR	43’810’382	0.00113	0.00905	Olfaction perception
4	TAS2R16	88’279’974	0.00143	0.00887	bitter taste receptor, type 2
17	TMEM132D	48’769’405	0.00087	0.00021	Neurobehavioral functioning (Anxiety)
5	CACNA1C	109’068’036	0.48366	0.00013	Bipolar disorder and schizophrenia
3	GRIK3	109’293’546	0.00761	0.00761	Reward-related learning

Candidates genes under selection …

I. Pigmentation …



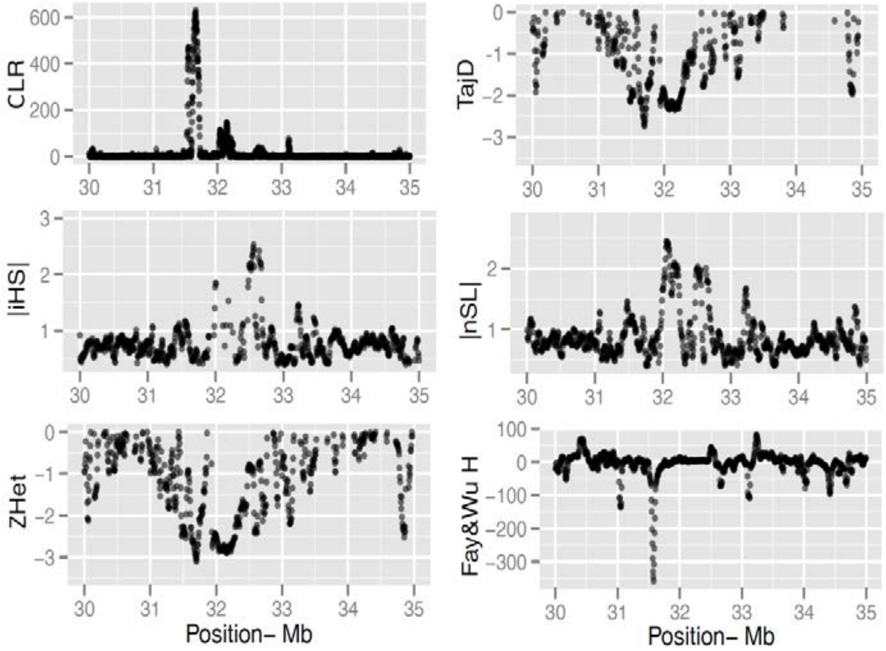




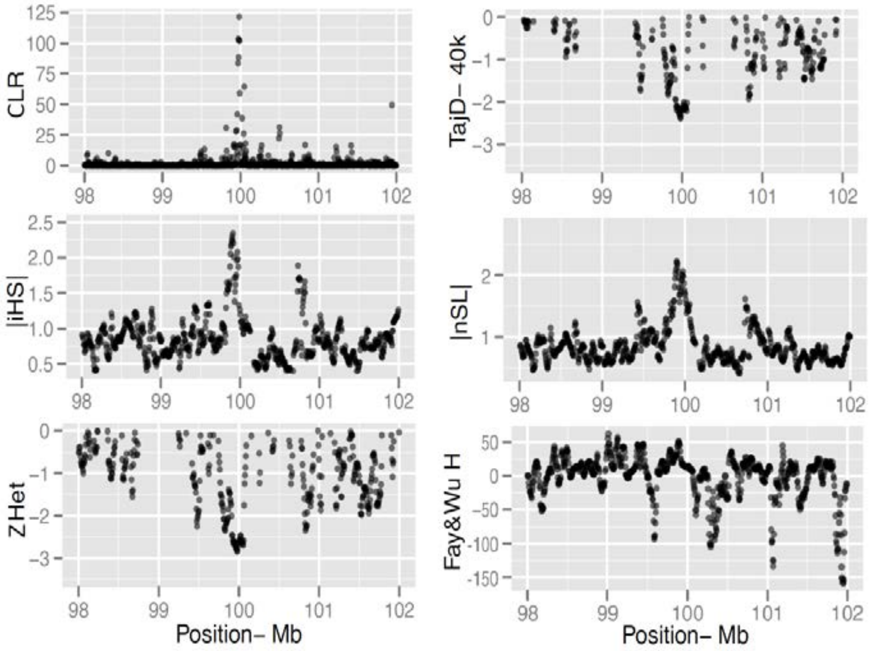
I. Pigmentation - *NRG–ErbB4* pathway

- Pigmentation (targeted by positive selection in humans Pickrell et al. 2009)
- ✚ QTLs for psychiatric disorders

NRG4-BTA21



ErbB4-BTA2

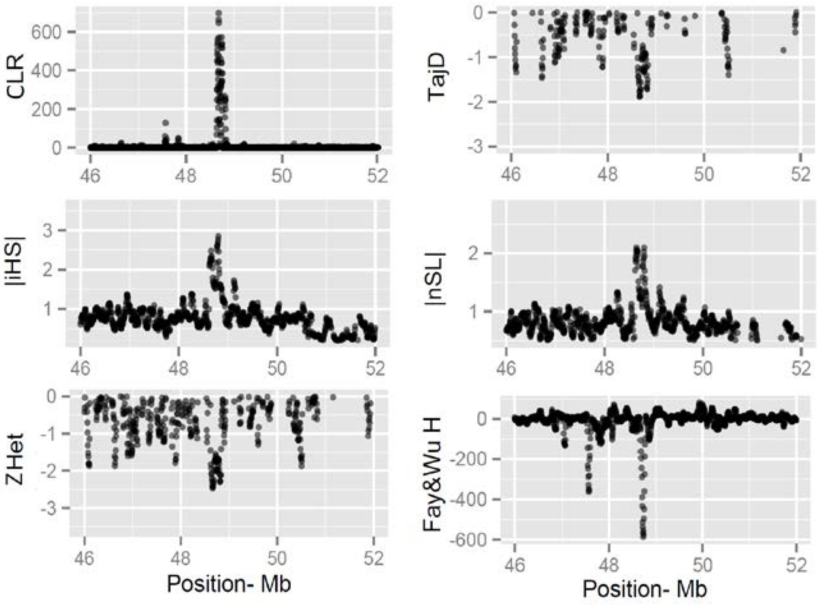


More examples:

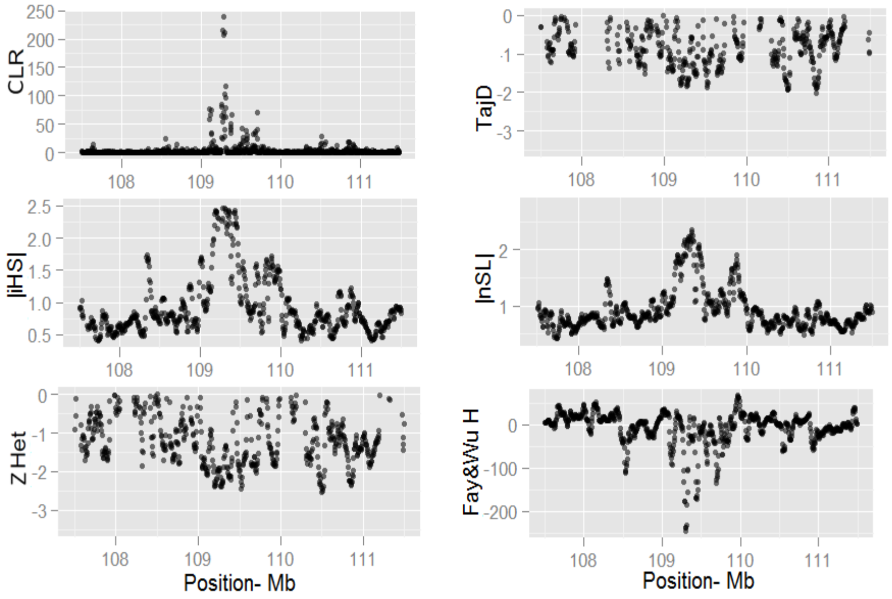
II. Neurobehavioral functioning

Anxiety-related behavior gene (*TMEM132D*) and reward-related learning gene (*GRIK3*)

*TMEM132*D-BTA17



GRIK3-BTA3



GWAS as additional evidence

Sequenced-based imputation

A two-step imputation:

50K SNPs
n=2’309

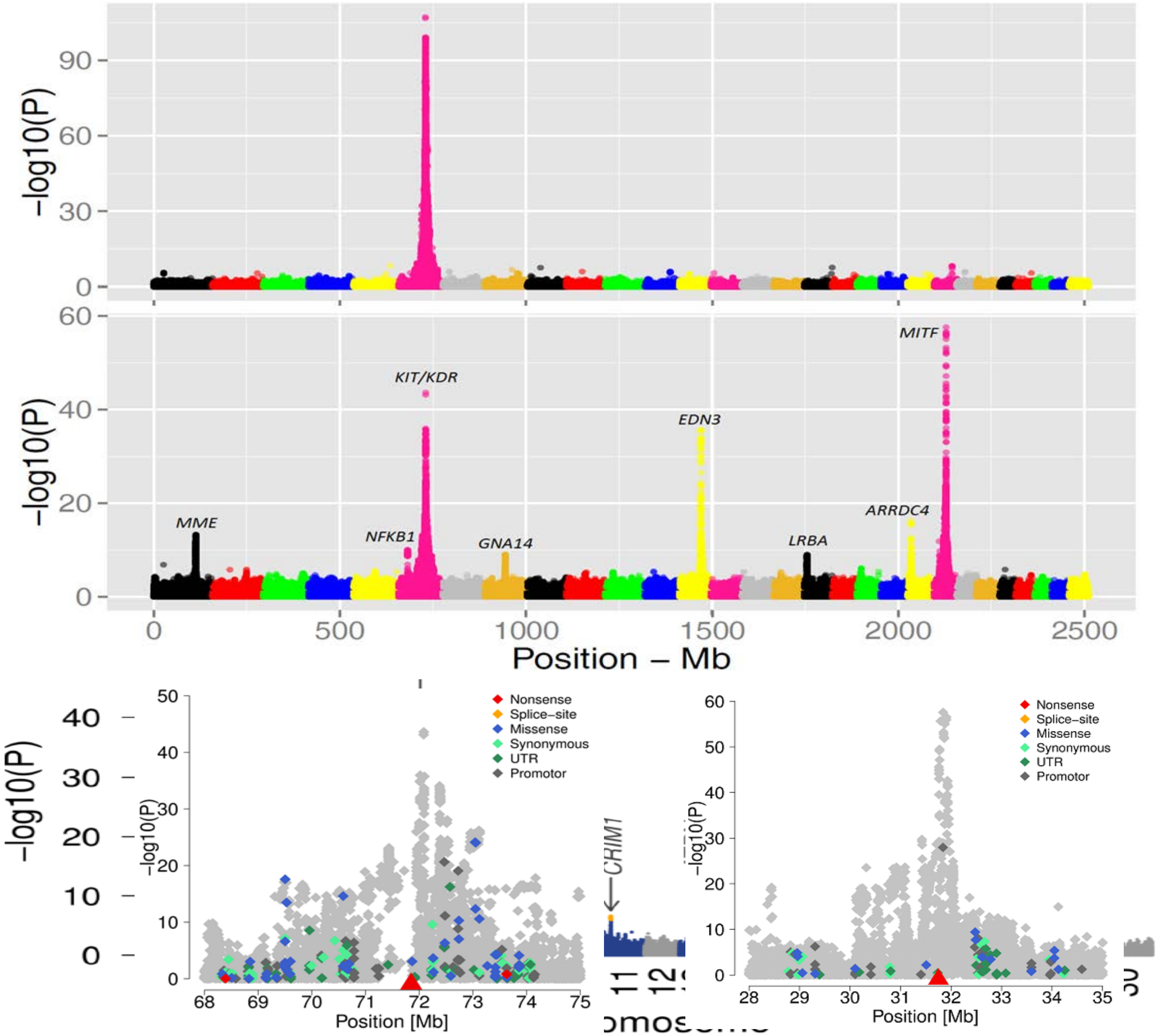
700K SNPs
n=1’293

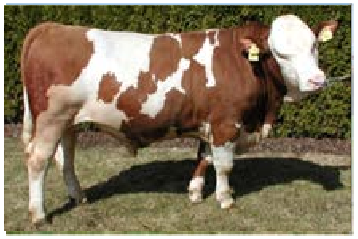
15M SNPs
n=43

GWA studies (15M SNPs and 3062 individuals) for 4 traits:

- The “proportion of daughters without spotting”
- The “proportion of daughters with red head”
- “Somatic cell count”
- and “body size”

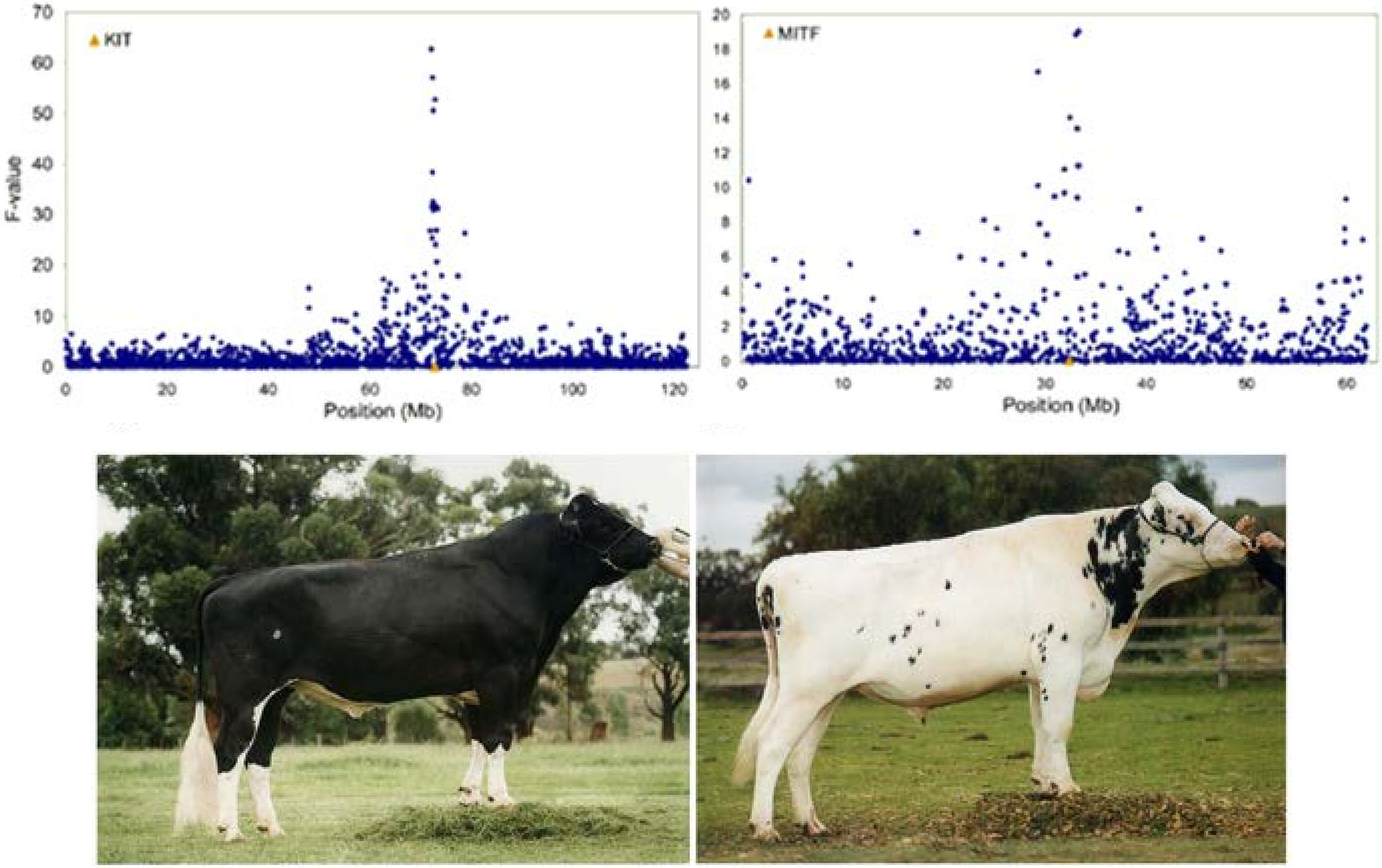
GWAS: Coat coloring …





Pausch et al. (2011)

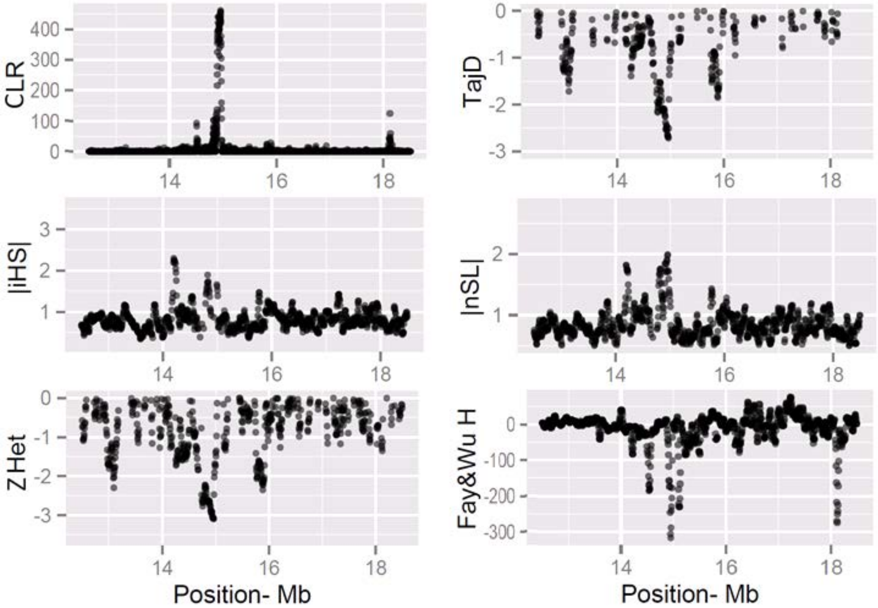
GWAS in Holstein: Coat coloring

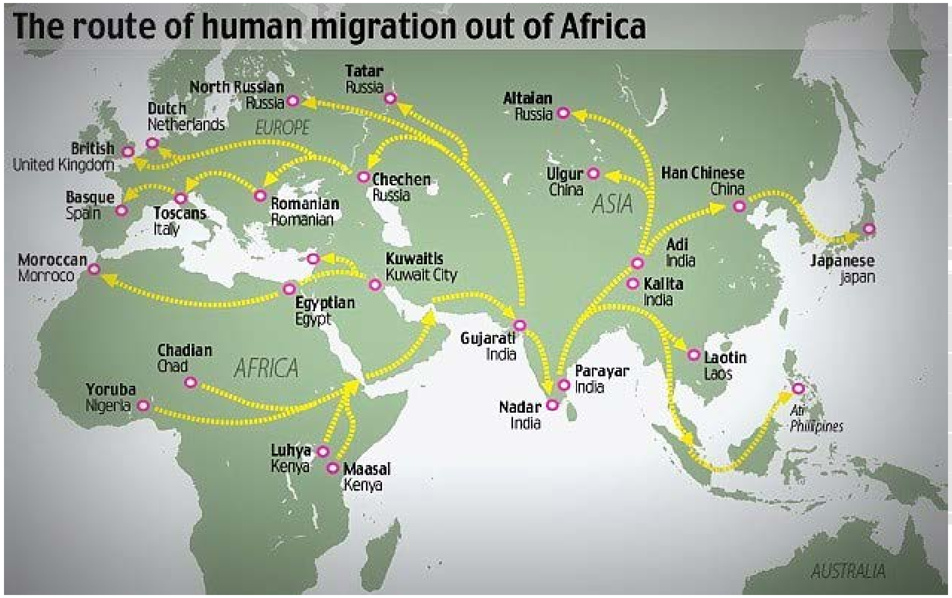


A GWAS for proportion of black phenotype. (Hayes et al. PLoS Genetics 2010)

MC1R gene: a closer look

Region harboring *MC1R*-BTA18





The route of human migration out of Africa

Conclusion

- Genome-wide, 106 classic sweeps were identified
- GWAS shows, coat colorings are co-localized with major QTLs
- Candidate genes are related to domestication phenotypes such as coat coloring, sensory perceptions and neurobehavioral functioning.
- As demonstrated for *MC1R*, selection can be detected if desired allele is fixed , and GWAS fails.
- Comparing GWAS results, shows complexity of coat coloring genetics in cattle.
- LD in cattle decays at a much faster rate than previously thought .

Thank you for attention !