



SCIENCE AND  
EDUCATION **FOR**  
**SUSTAINABLE**  
**LIFE**

# Problematizing selection mapping from an animal science perspective

Martin Johnsson

Department of Animal Breeding and Genetics, Swedish University  
of Agricultural Sciences

# Acknowledgements

Dominic Wright

Per Jensen

Rie Henriksen

Eben Gering

Pamela Willis

Saioa Lopez

Lucy Van Dorp

Garrett Hellenthal

Amir Fallahshahroudi

Enrico Sorato

S J Kumari A Ubhayasekera

Jonas Bergquist

Jordi Altimiras

Anna Maria Johansson

Dolapo Adepoju

FORMAS 

Dnr. 2020-01637

Selection mapping

Population genetic limitations

Practical concerns

Slides & references:

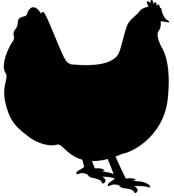
<https://github.com/mrtnj/asas2022>

GAATGTAACCAATTACTTGCCTCAGTGACAAAGGAATGAAACAATATGCCATCCTCTAGTTGAGAGCCTGCAGCA  
GAATTACTATGGATCACTAATAGGAAGTAGCTTATTAGTAAAGGTAAAGATAACATAAGGATATCTTGTACACAAAATGA  
GTAAGTATGATTGGGTTATGCAAACCTACTGAAATAACTCAAATCACGGGTCTACTTATCTGTCCACCAAAAAAAACA  
ACAACAAAAAAACACTAATATTCTTGCTTAGTCCACAAAAGCAGAACATAAAAGTAACACTGGACTCAGAATAAAA  
GCAAGGGAGAGTTGCAAGAGAGCACCCACAAGAGGTTACAAACATCACTCAATCAGTGTACCTTGAAGAGGGAGTA  
CAAGTCCCCAGTTCTTAGTCTCACCAAGGTCTTGTGAGAACATAGACTCTTGTAGTGCTCTTAGTTGT  
ATTGGATCATCAGGAAAGTTAAGAGAGGACAGAGATTAAGCTGGAAGTGGATTAAACACTATGGCTCAAGATTGACAAG  
CTGCCGCTGCAGTGTCAAGCAATGTTATGTATAAAACTATGGCTTCAACTATGGCTCTACTGCCTAGAAATCAA  
GTGTAGAATAGACTCATGGCAGCTAACAGACATTGATCCAGAGAACAGAGATTCCCTCTGGTATCTATTCAATTGCCA  
GATGATTTATTTCTCTGTGTCAAGCCTGTCAATCATTATTCGCCAAGTCAAGAGGAACGTGAGGAACTGAGGAACTGAGATT  
TCATGTCCTGTTGCTAACAGAACATGCGGTGTCCCTGCATTCTGGGACAAGCTGGCTATTTAGCAGTAAGAACACATGGAATGTTA  
ATCAGCAGAGAACAGTTGATACTCTAGAAAAATCCTGAGGGCAGCTGAAACATGCTGAATATTACAAACACATGGAATGTTA

Cast your mind back to 2010, the “(post-)genomic era” ...

Reference genome

2004

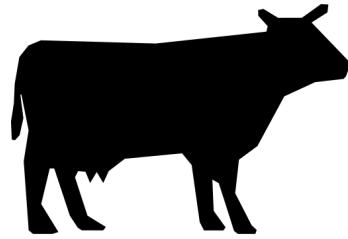


SNP chip

2009

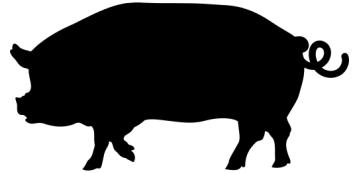
2009

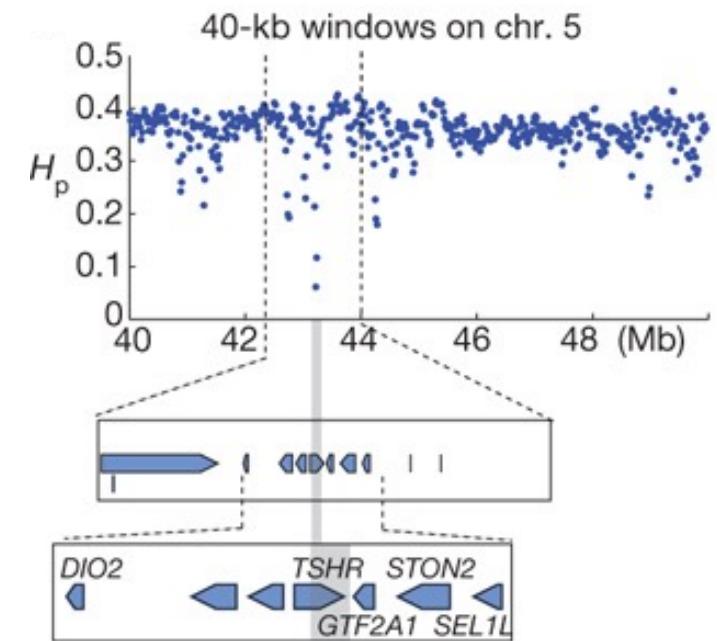
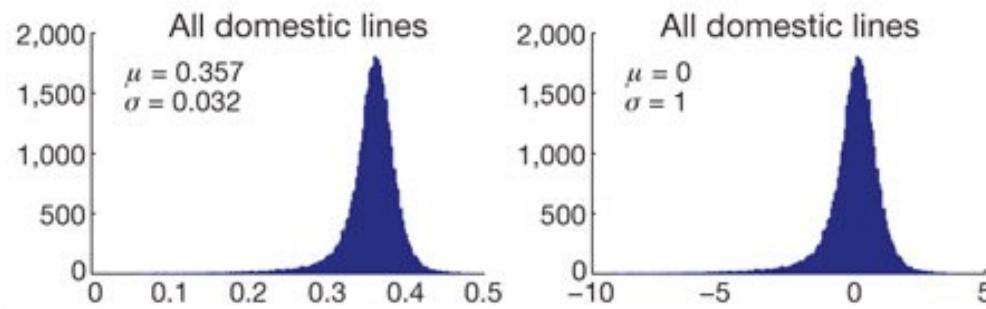
2009



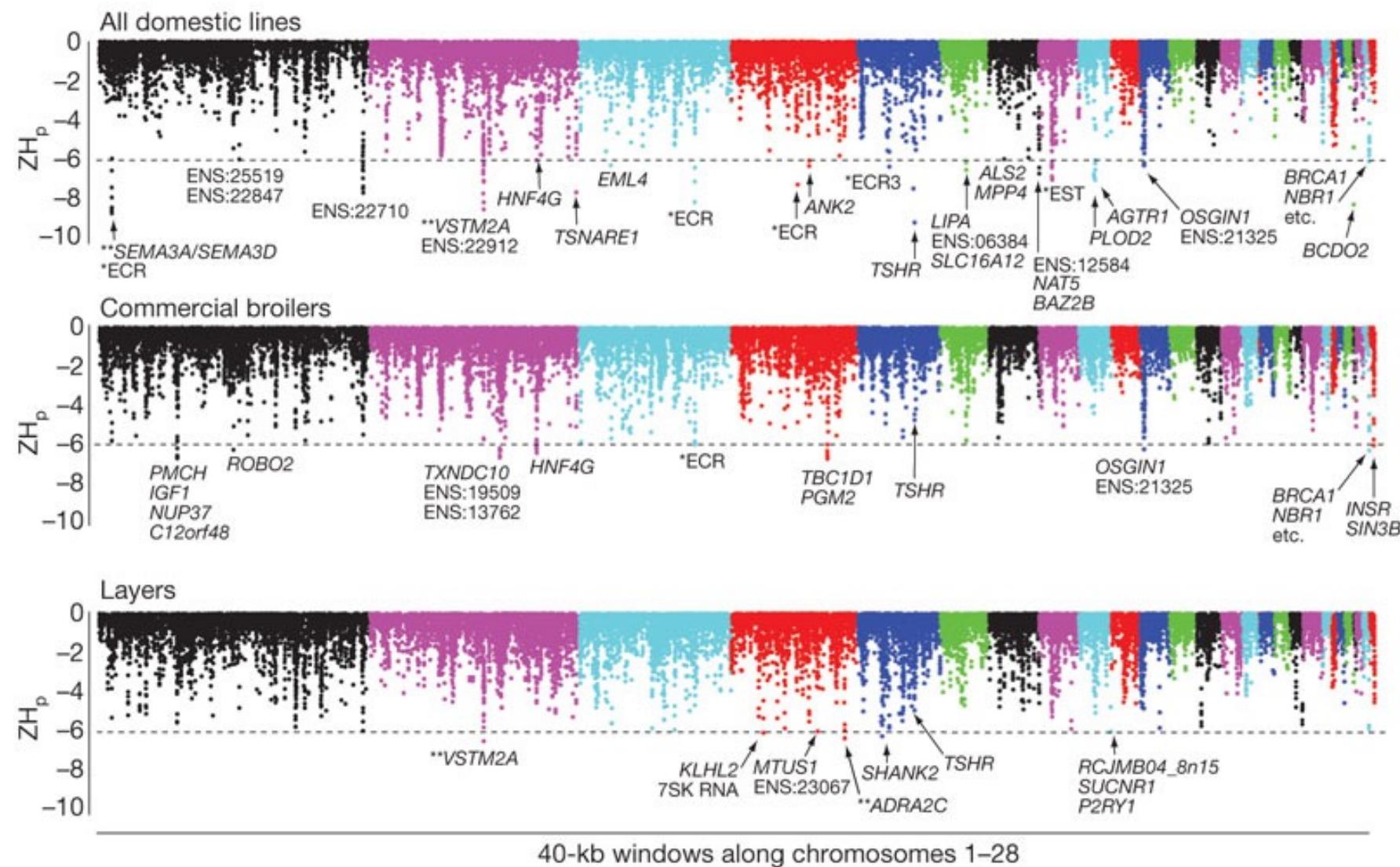
2012

2011





(Rubin et al. 2010)



(Rubin et al. 2010)



Walmart parking lot, Lihue, Kauai

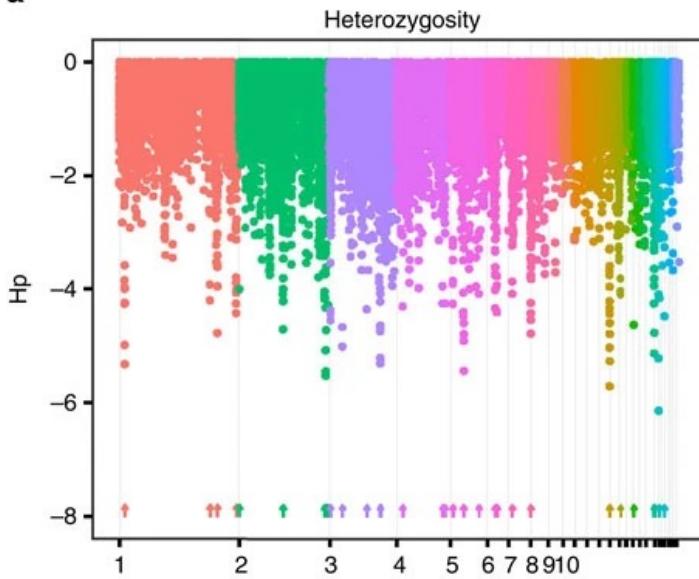
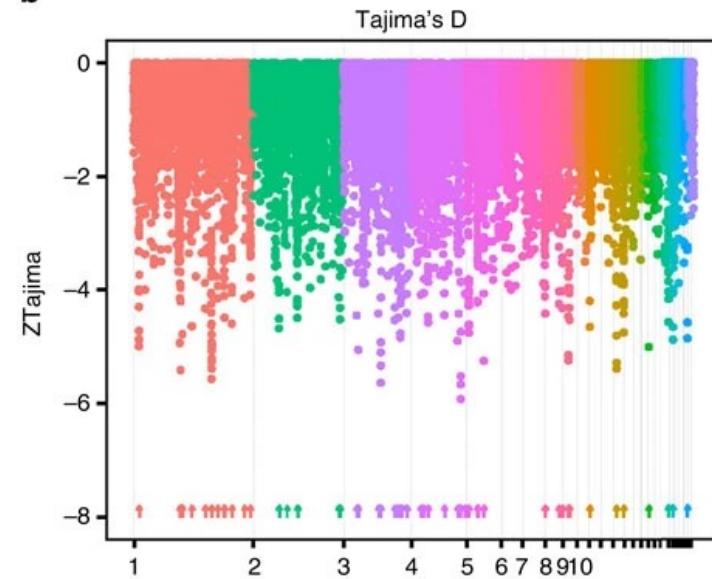
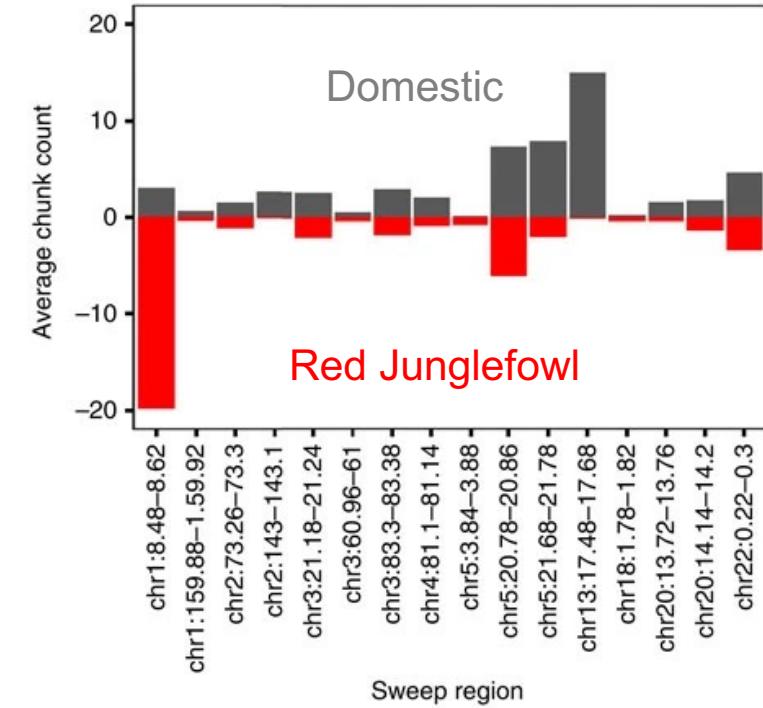




23 low-coverage SOLiD sequences

Selection mapping (heterozygosity, Tajima's D)

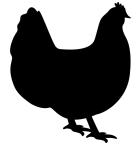
Chromosome painting to compare haplotypes to  
domestic pools

**a****b****c**

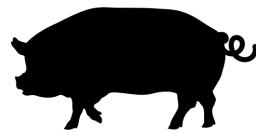
Different sweeps than in domestic chickens

Some more similar to domestic, some more similar to  
Red Junglefowl

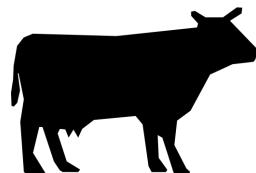
Chromosome 13 overlap with comb size locus



*BCDO2* yellow legs (Rubin et al. 2010; Qanbari et al. 2019)



*MC1R* major growth locus, *NR6A1* vertebrae number, *PLAG1*,  
*LCORL* possibly growth (Rubin et al. 2012)



*KIT*, *PMELO*, *MC1R*, *MITF* pigmentation, *MSTN* double-muscling  
(Druet et al. 2013; Qanbari et al. 2014)

And too many others to cite

Wilkins, Wrangham & Fitch (2014): neural crest cell hypothesis of domestication

Selection on tameness changed migration or proliferation of neural crest cells

“By comparing village dogs and wolves, we identified 246 candidate domestication regions in the dog genome. Analysis of gene function in these regions suggests that perturbation of crucial neural crest signaling pathways could result in the broad phenotypes associated with the domestication syndrome.”

Pendleton et al. (2018)

"Intriguingly, several of the genes that manifested significant allele frequency changes in our tame foxes may play a role in neural crest cell fate (45). Wnt signaling plays a key role in initial neural crest cell differentiation, and both *Wnt3* and *Wnt4* in the fox have more than one SNP with significant allele frequency changes. Protocadherins are also important in neural crest cell function. Direct assessment of whether these genes play roles in neural crest cell function in the fox presents an interesting experimental challenge."

Wang et al. (2018)

## Features of selection mapping:

1. Works best for recent selection
2. Works best for large effects, simple architectures
3. Dependent on genome features (population history, mutation, recombination, standing variation)

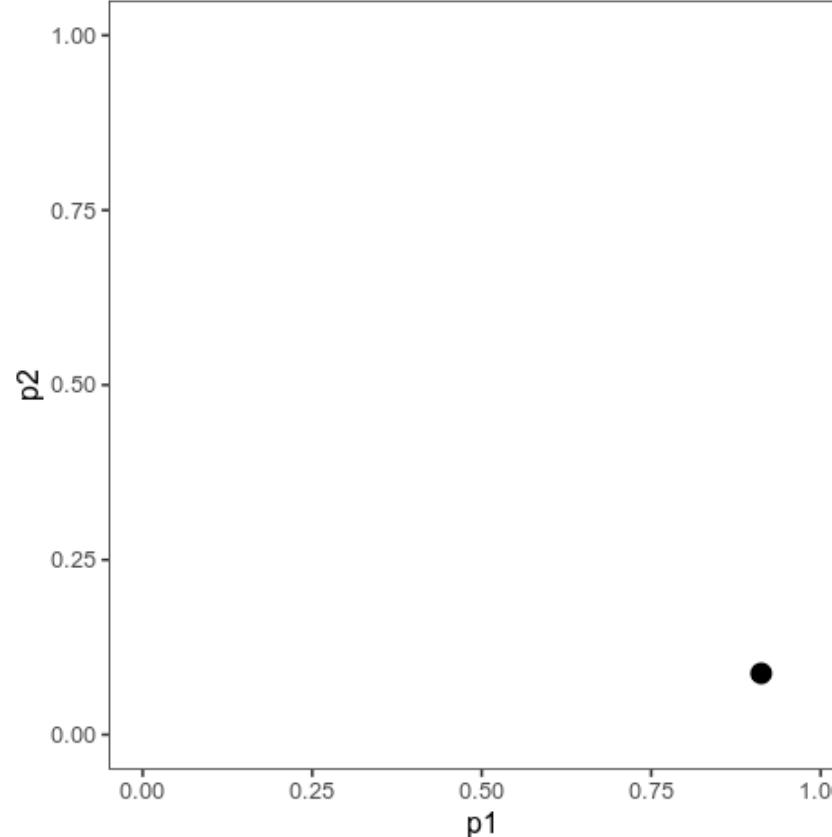
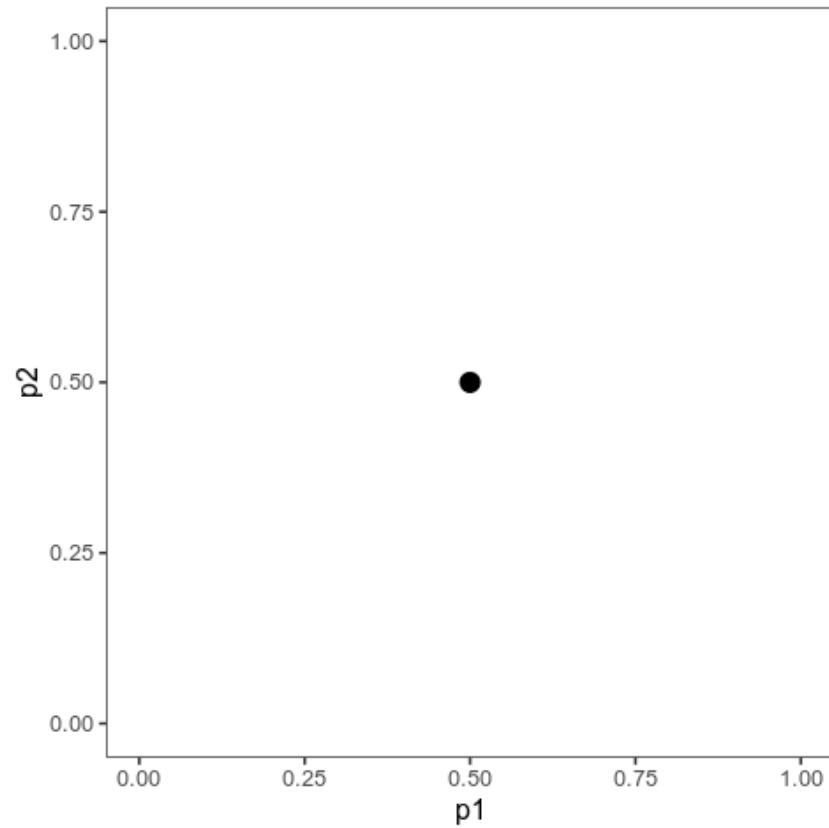
Average fixation time of neutral allele is  $4N_e$ , selected alleles fix faster than that

Selection signature frequently erased much earlier ( $0.1N_e$ ;  
Przeworski 2002)

$N_e$	$4N_e$	$0.1N_e$
100	400	10
500	2000	50
10,000	40,000	1000

Example: *TSHR* selected last 500-1000 years, not early in domestication (Girdland Flink et al. 2014; Loog et al. 2017)

Adaptation by small shifts of small-effect alleles or large shifts of large-effect alleles (e.g., model by Jain & Stephan 2017)



Example: major loci, pigmentation loci etc.

## Mutation

Signature of selection  
or mutation coldspot?

?

## Recombination

Reduces linkage  
disequilibrium and  
selective interference

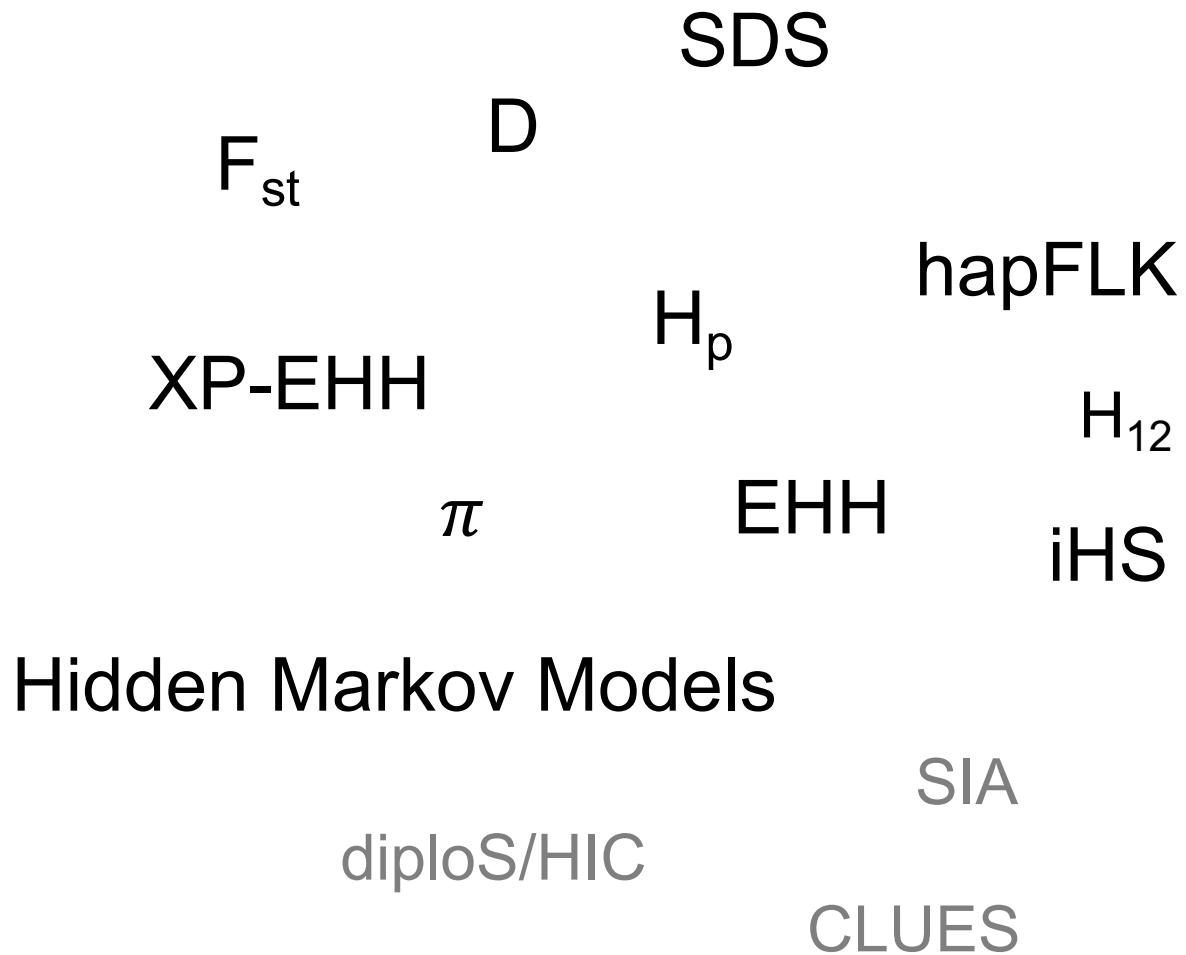


## Population history

Bottlenecks reduce  
diversity, erase  
historical record

Methods concerns:

1. Researcher degrees of freedom
2. Hard to check design properties
3. Hard to interpret



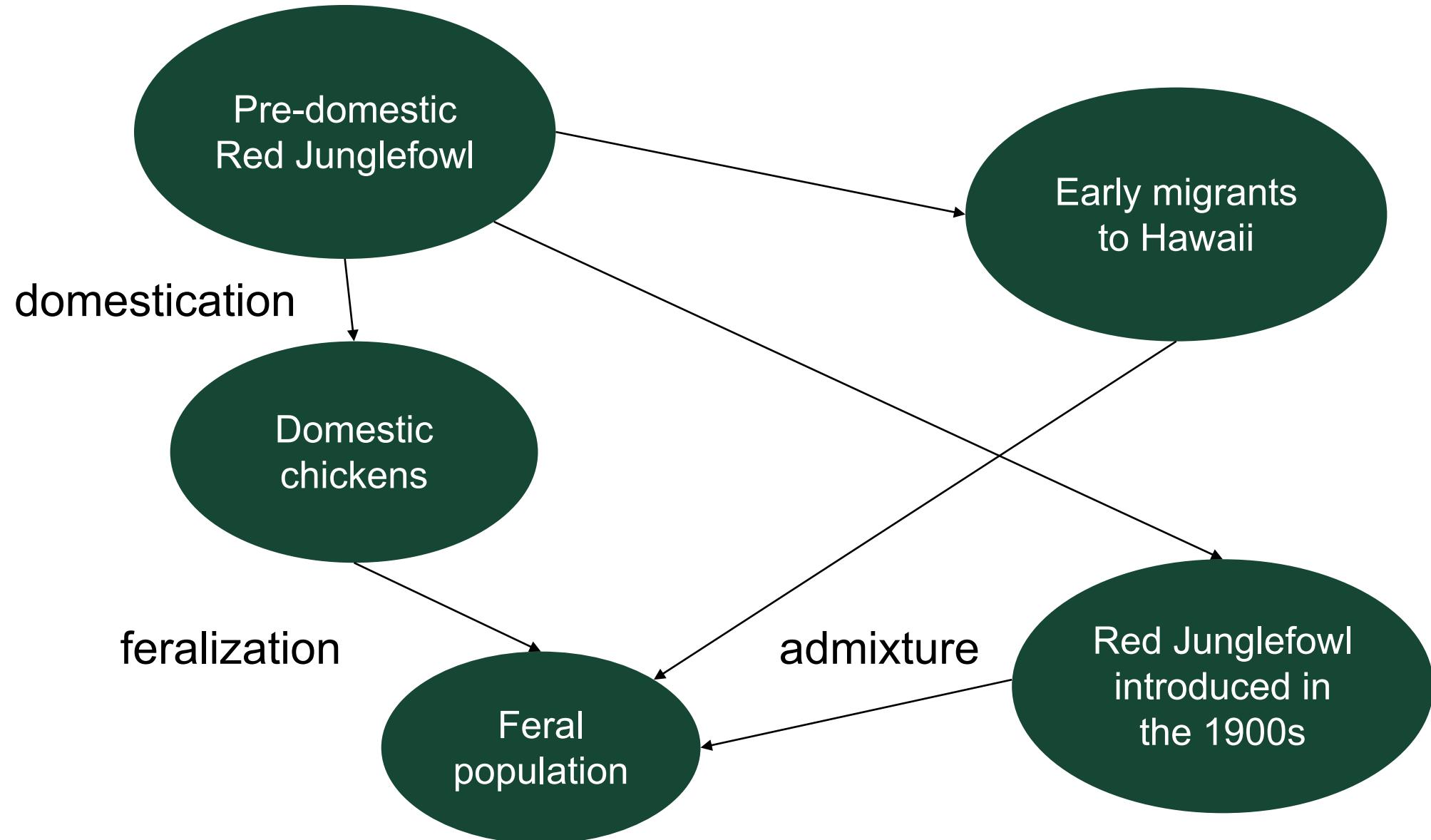
Statistical phasing

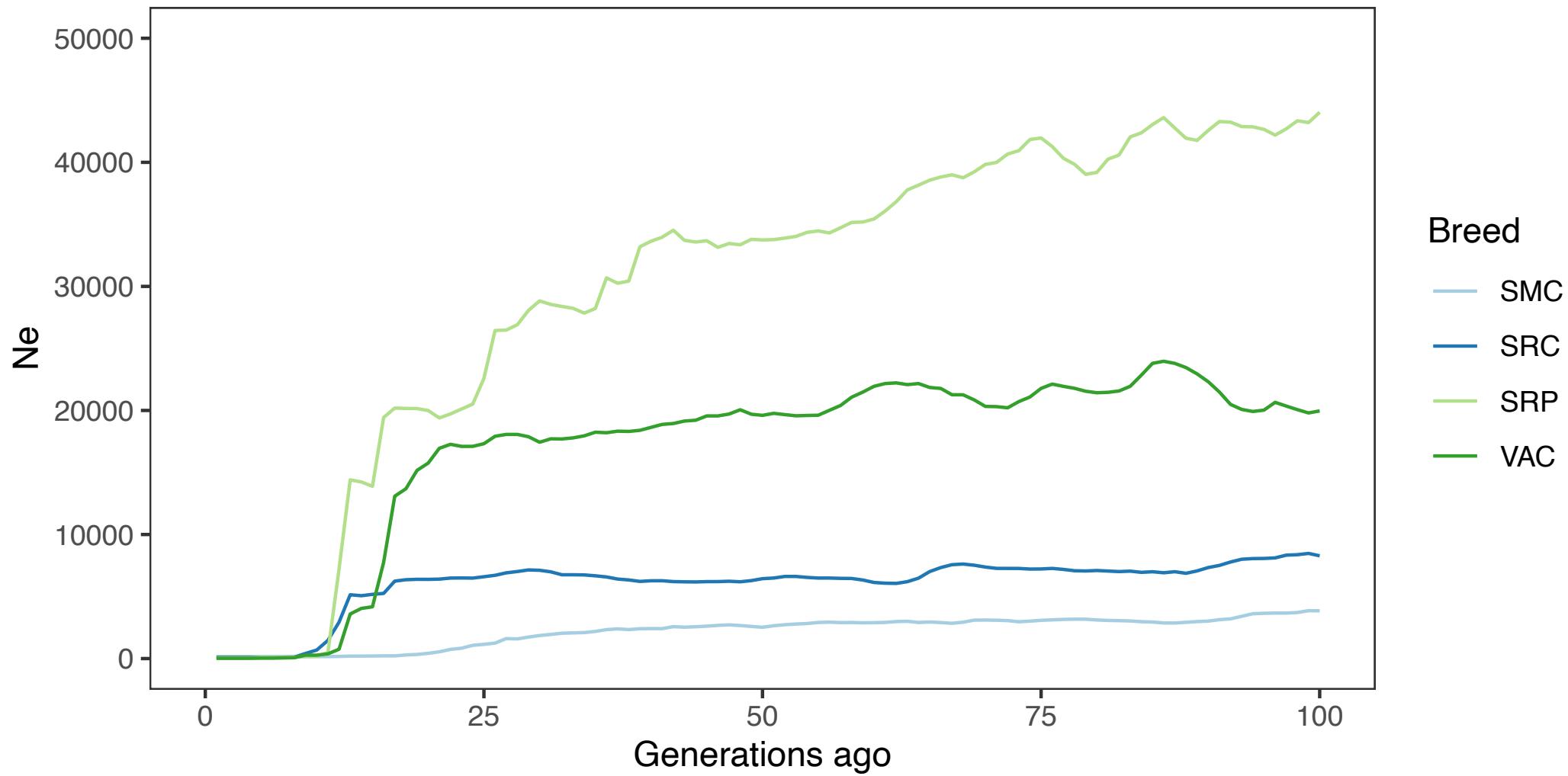
Ancestral alleles

Windows along the genome

Thresholds

Union, intersection,  
meta-analysis





Population history estimates for Swedish cattle breeds  
(Dolapo Adepoju)

## Annotation term enrichment

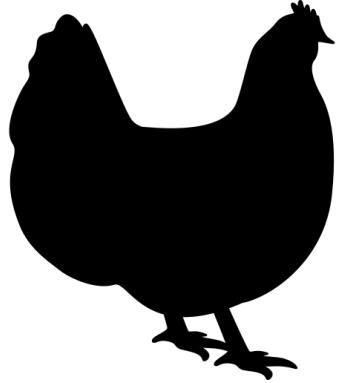
annotation biases

researcher degrees of freedom

context-dependent pathway recycling

Example: neural crest cell hypothesis (Johnsson et al. 2021)

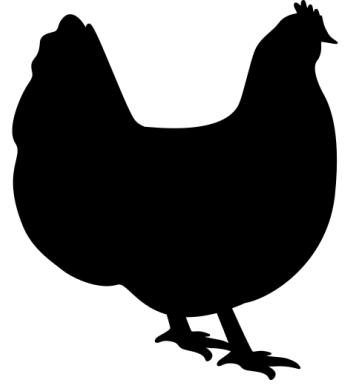
Readily creates stories out of noise (Pavlidis et al. 2012)



## Function

“We did not find compelling genotype effects in open field, tonic immobility, aerial predator, associative learning or fecundity. Therefore we conclude that *ADRA2C* is probably not involved in the domestication of the stress response in chicken ...”

Elfwing et al. (2014)



“Functionally, the individuals with the domestic *TSHR* variant had a lower expression of the *TSHR* in the hypothalamus and marginally higher in the thyroid gland than wild-type *TSHR* individuals. Expression of *TSHB* and *DIO2*, two regulators of sexual maturity and reproduction in birds, was higher in the pituitary gland of the domestic-variant chickens.”

Fallahshahroudi et al. (2021)

$$\dot{p}_i = -s \gamma_i p_i q_i (c_1 - z') - \frac{sy^2}{2} p_i q_i (q_i - p_i) + \mu (q_i - p_i)$$

Jain & Stephan (2017)

“The late, greatly loved, James Crow used to say that he had no objection to graduate students in his department not taking his course on population genetics, but that he would like them to sign a statement that they would not make any pronouncements about evolution. There are still many papers published with confused ideas about evolution, suggesting that we need a ‘Crow’s Law’, requiring authors who discuss evolution to have acquired a knowledge of basic population genetics.”

Brian Charlesworth (2019)

## Upsides:

1. Natural history of the genome
2. Bring genomics to some interesting population
3. Maybe even find a monogenic adaptation

We need:

Analysis workflows that make modern selection mapping as easy as a Hardy-Weinberg scan

```
nextflow domestic_selection_mapping.nf --species sheep --variants sheep.vcf.gz
```

Simulation workflows that let us check performance, accounting for genome features

To think of selection mapping like an exploratory analysis

? ! @

Martin.Johnsson@slu.se



@mrtnj

<https://github.com/mrtnj/asas2022>