

# A novel method to visualize local LD

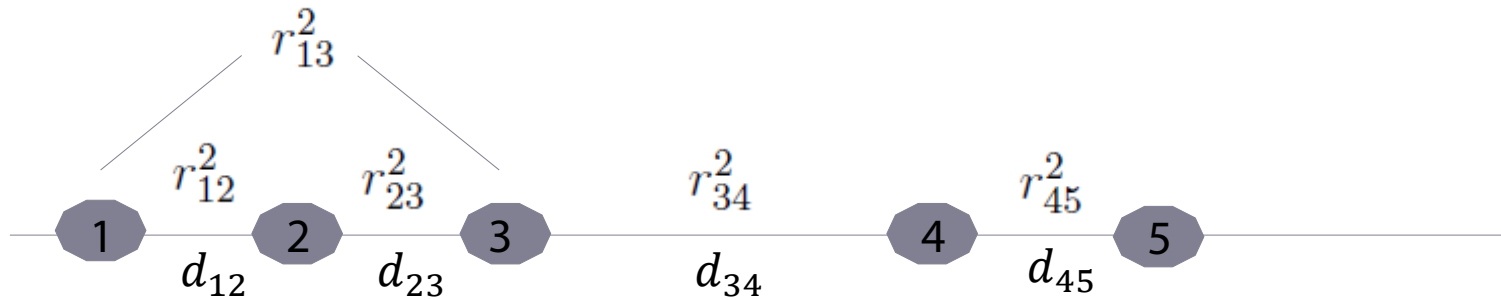
## -- a study on Danish pig breeds

Lei Wang

David Edwards, Peter Sørensen, Luc Janss and Tage Ostensen

# Linkage disequilibrium (LD)

Degree of non-random association of alleles at different loci (correlation between alleles)

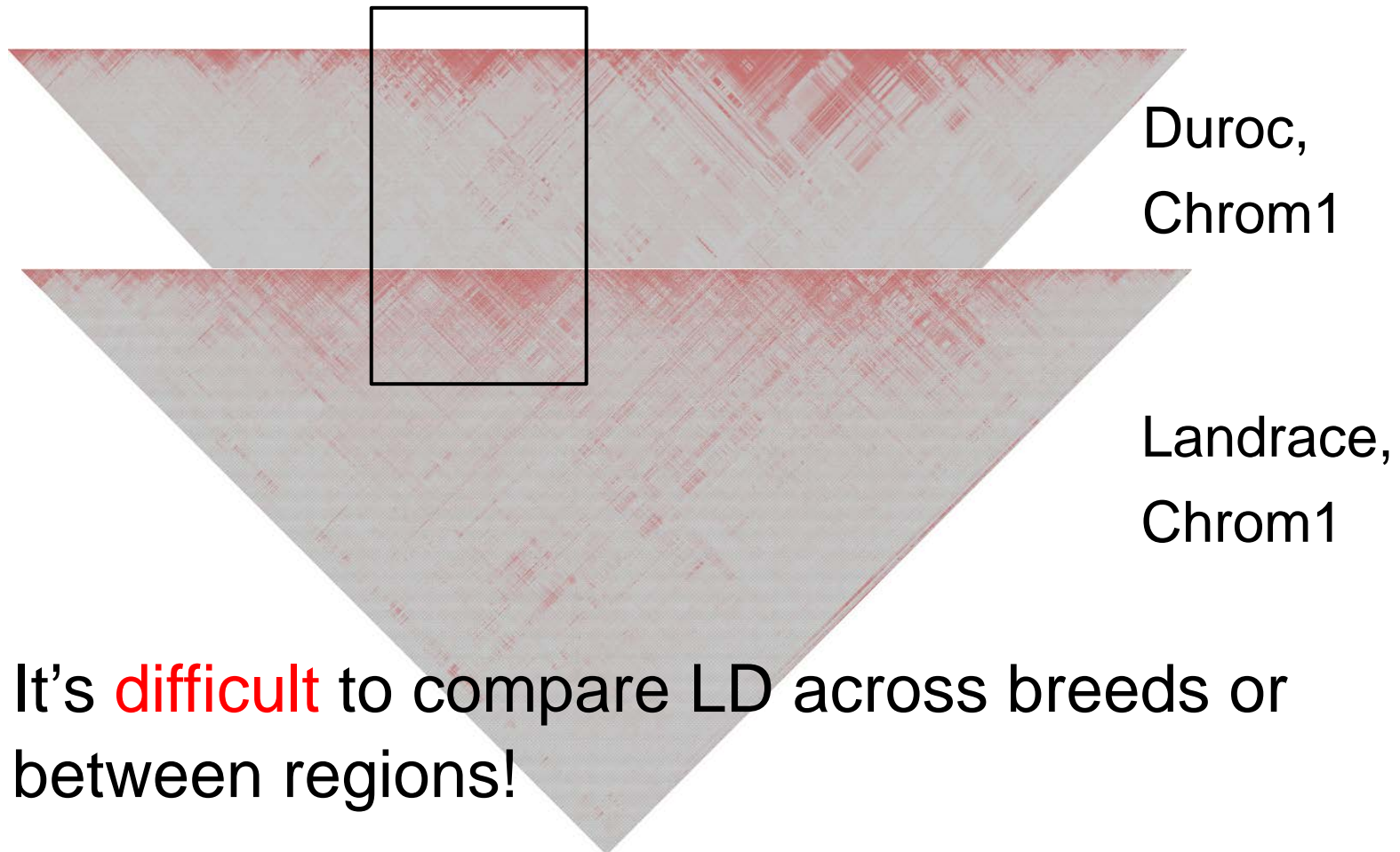


	SNP1	SNP2	SNP3		SNP1	SNP2	SNP3	...
P1	1	1	0	SNP1	1	0.13	0.42	
M1	0	1	1	SNP2	0.13	1	0.38	
P2	0	0	0	SNP3	0.42	0.38	1	
M2	1	0	1	...				
...								

LD matrix

# heatmap

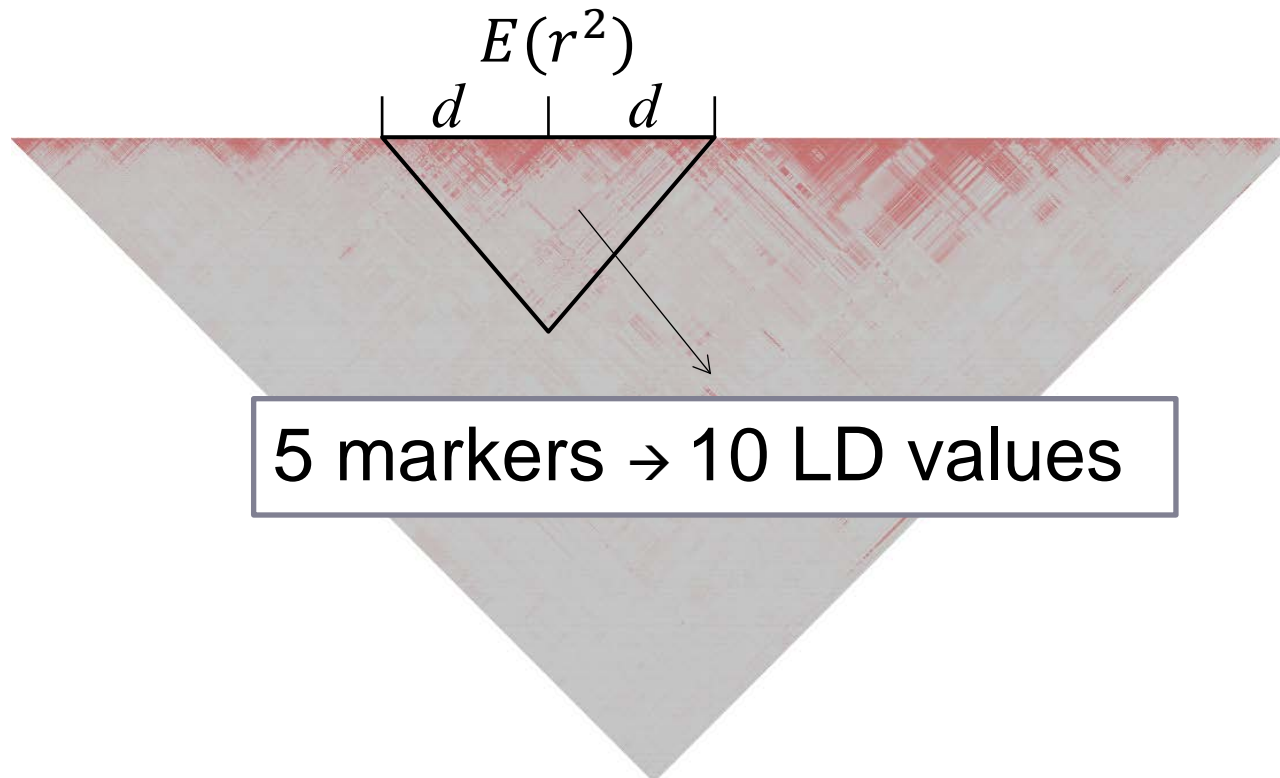
## Conventional method to visualize LD in color blocks



It's **difficult** to compare LD across breeds or between regions!

# Can we quantify LD block as a single value?

Example:



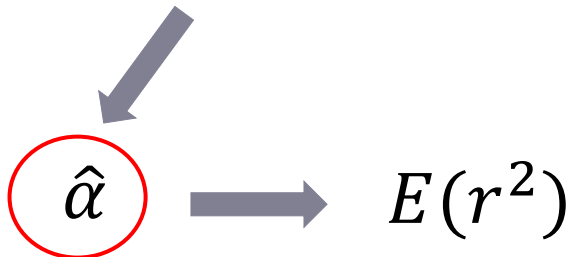
# Regress LD ( $r^2$ ) on distance ( $d$ )

## Sliding window

- window size: 50, 100, 200 markers
- 1 marker per step

## The model:

$$r_{ij}^2 = 1 / (1 + \alpha * d_{ij})$$

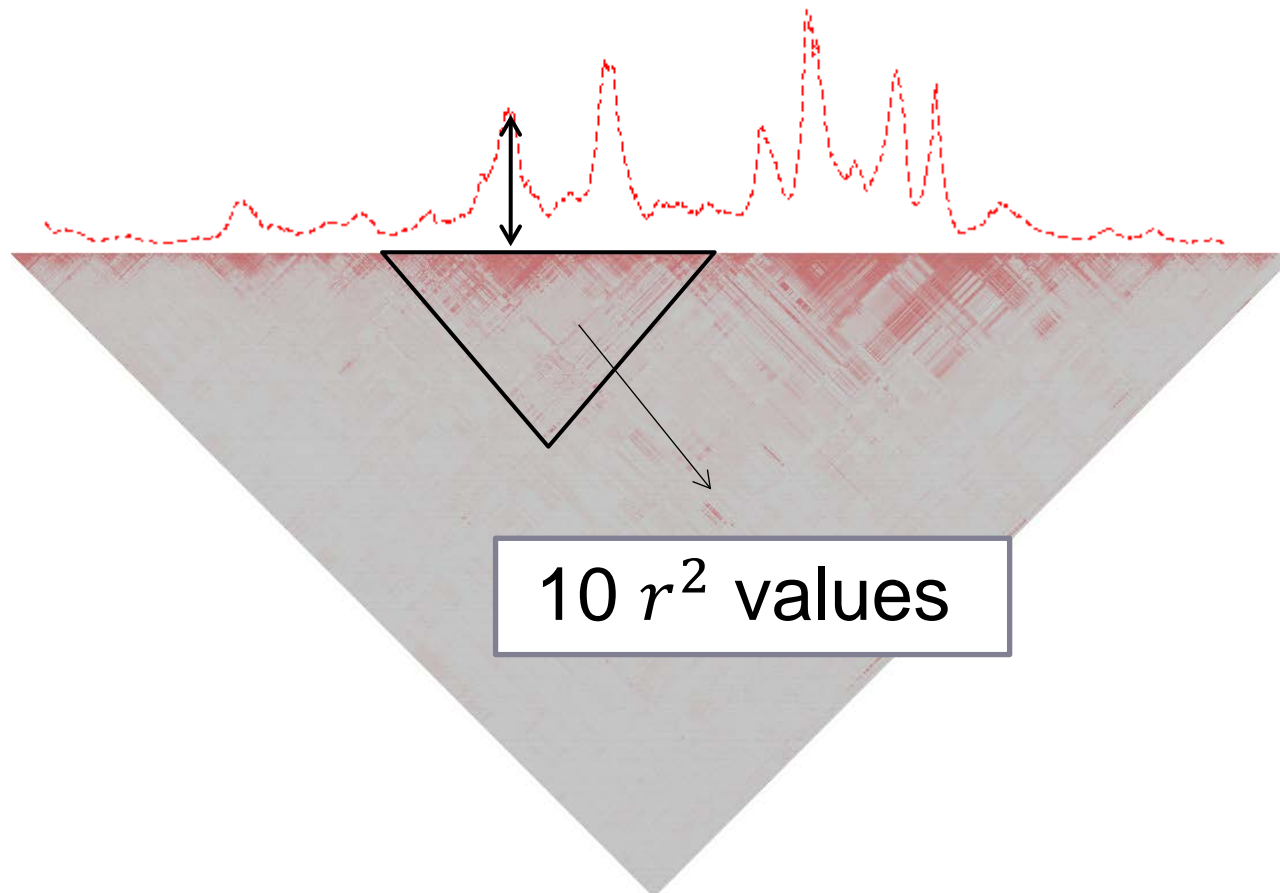


1. fit  $r^2$  and  $d$  into the model, estimate  $\hat{\alpha}$  for each window
2. substitute  $\hat{\alpha}$  and  $d$ , calculate  $E(r^2)$  for each window
3. plot  $E(r^2)$  against middle-SNP physical position

**”local” effective population size**

# single value represent LD at each marker locus

Example:



# Data

## 3 Danish pig breeds:

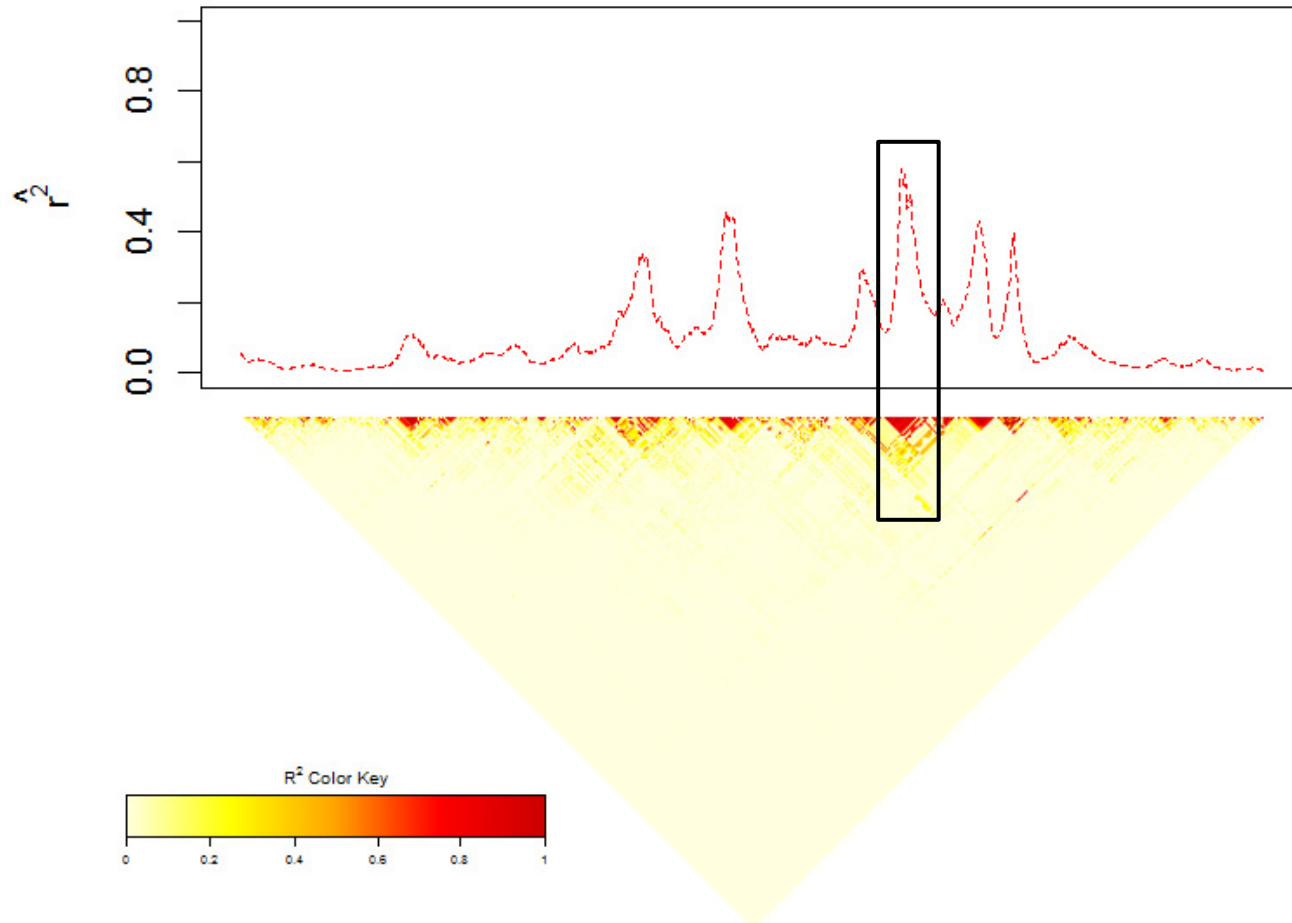
	No. of animals
Duroc	4249
Landrace	1979
Yorkshire	2123

## ILLUMINA PORCINE SNP60 BEADCHIP

	No. of markers
Duroc	29567
Landrace	33719
Yorkshire	33907

# local LD map **versus** heatmap

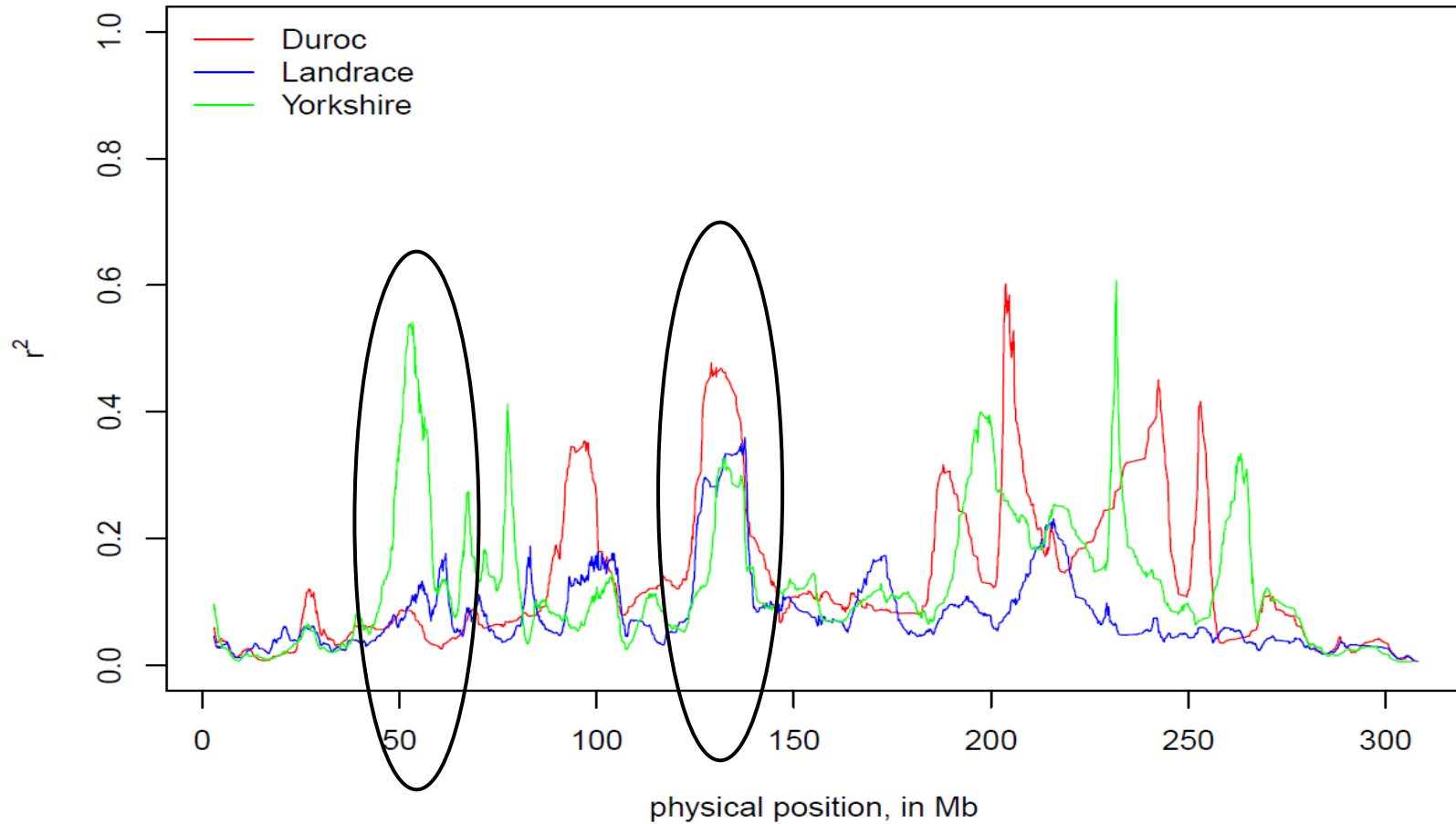
Duroc, Chromosome 1, N=100





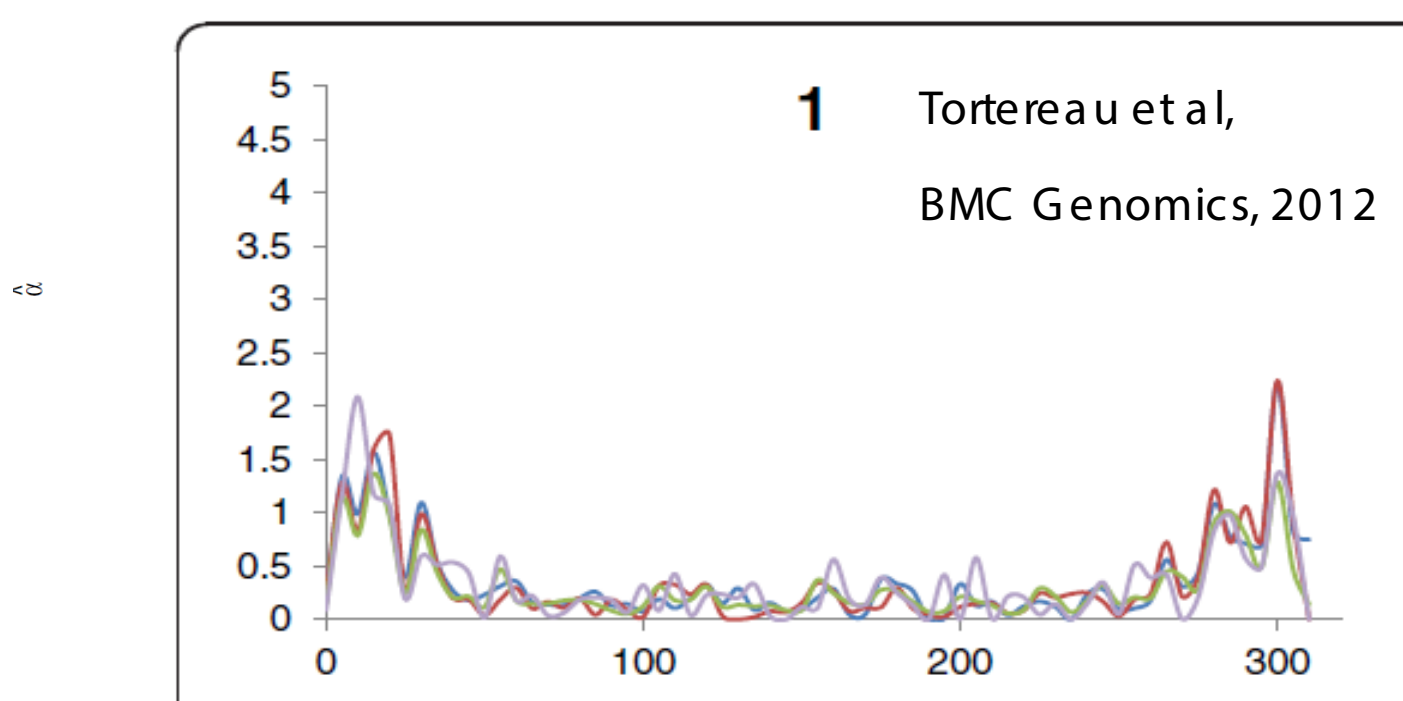
# Compare local LD across breeds

(b)



# Discussion

- **The robustness of the method is being tested**  
 $\alpha$  is population specific
- **Interpret  $\alpha$  as "local effective population size"**



# Conclusion

- **New method for visualizing local LD developed**
- **Method facilitates comparison of genetic variation across breeds**
- **Prospects of the method:**
  - Assist genomic prediction across breeds
  - Selection of markers for prediction

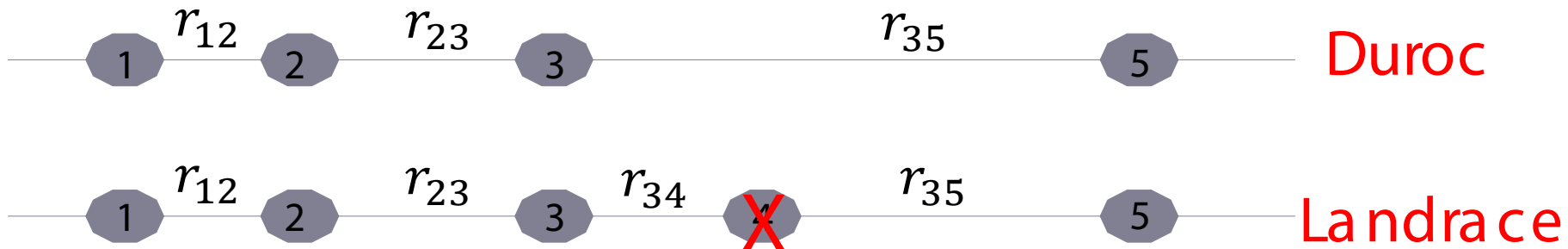
Thank you  
for your attention !!!



# Measure Persistence of phase

- correlation of  $r$  in two populations
- common markers between two populations

$$\text{cor}(r_{A,B}) = \frac{\sum_{(i,j) \in p} (r_{ij(A)} - \bar{r}_A)(r_{ij(B)} - \bar{r}_B)}{S_A S_B}$$



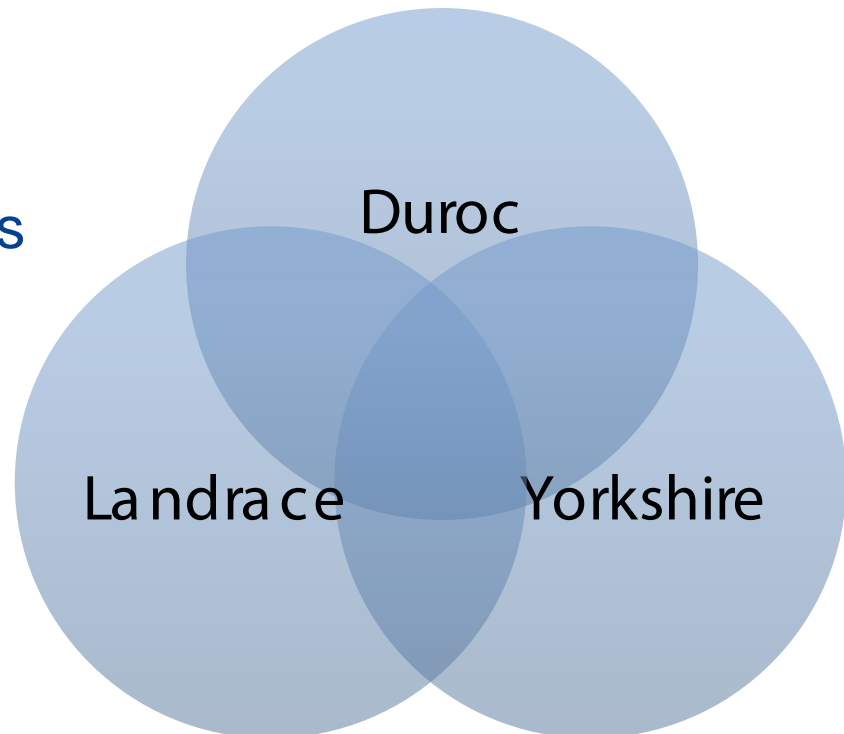
# Persistence of Phase over distance

- **Persistence of phase:**

- group  $r$  the same way as LD
- calculate correlation of  $A(r_{12}, r_{23}, r_{13}, \dots)$  with  $B(r_{12}, r_{23}, r_{13}, \dots)$

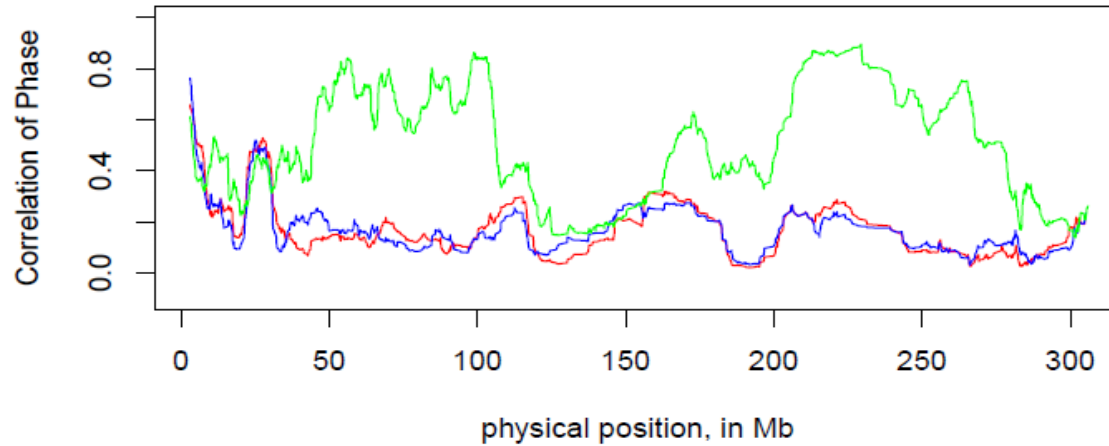
- **Two scenarios:**

- Scenario I: 17,918 all 3 breeds
- Scenario II:
  - Duroc-Landrace 22,142;
  - Duroc-Yorkshire 22,347;
  - Landrace-Yorkshire 26,505

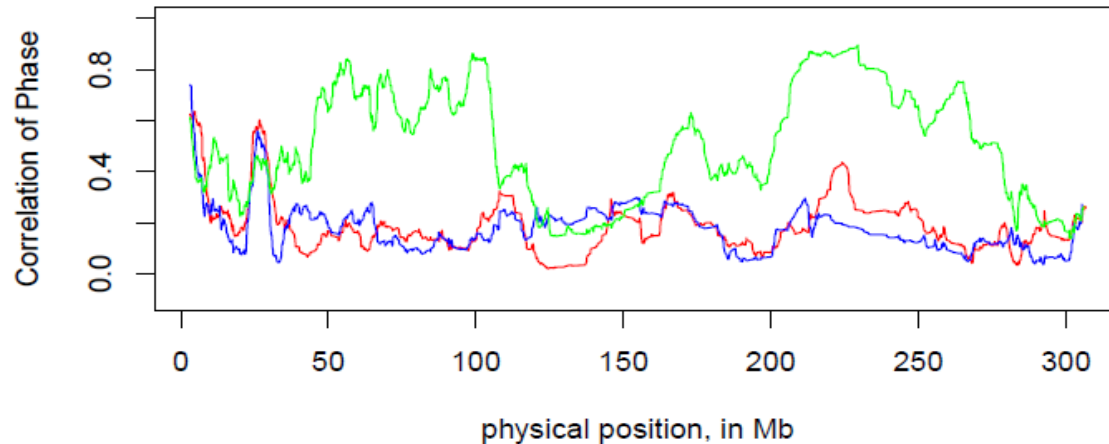


# local PoP in two scenarios

**(a) Scenario I**



**(b) Scenario II**



# Data

- **3 breeds**

- Duroc, Landrace, Yorkshire

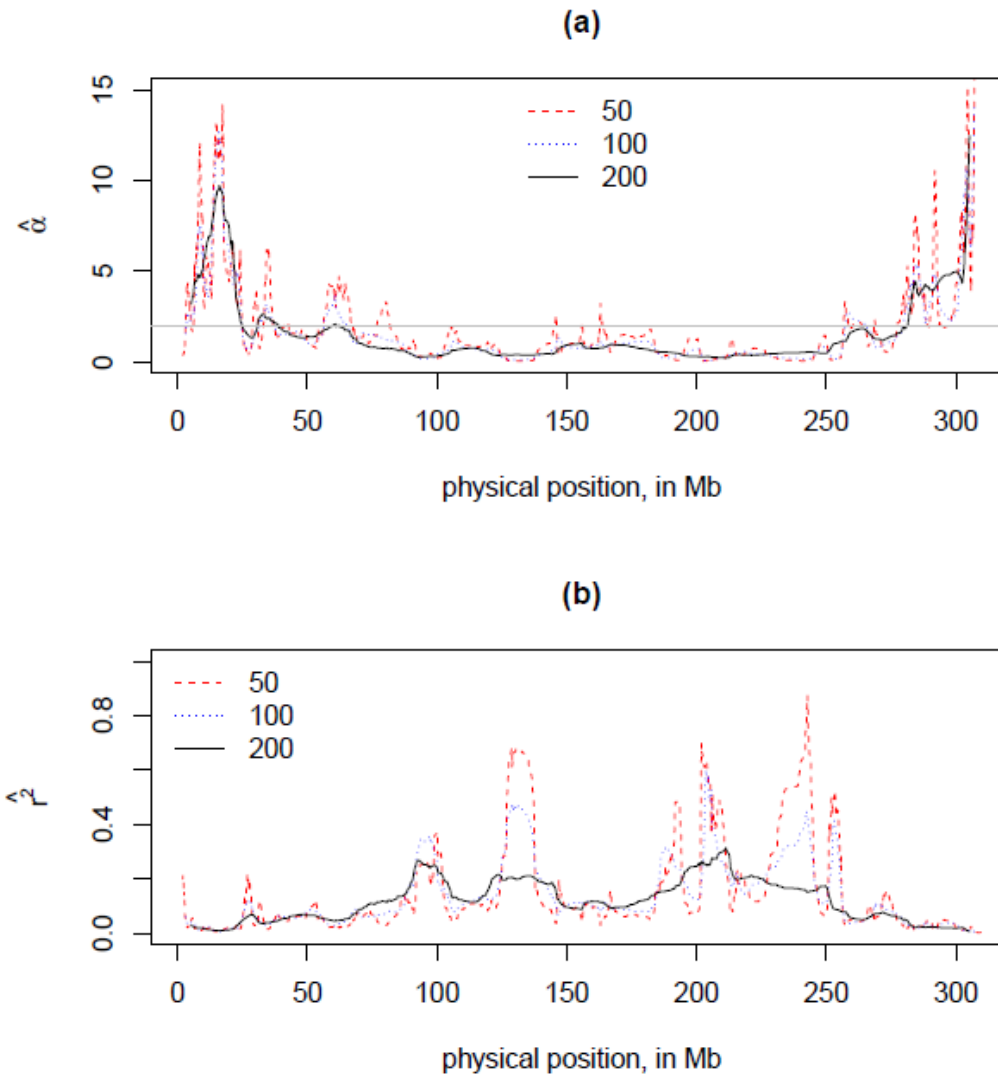
- **Illumina PorcineSNP60 BeadChip**

- Genotype missing rate < 10% (animals)
- CallRate > 0.9
- MAF < 0.05
- Hardy Weinberg Equilibrium ( $p < 1 \times 10^{-7}$ )

	Duroc	Landrace	Yorkshire
Sample size	4249	1979	2123
SNP number	29567	33719	33907
Aver. Space	83 kb	72 kb	72 kb

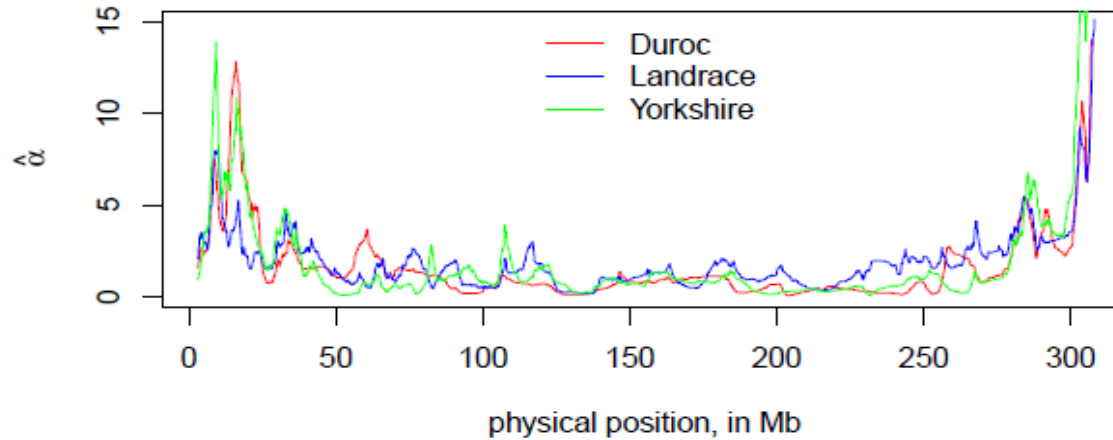


# sliding window size N



# compare local LD across breeds

(a)



(b)

