

# Using dense SNP markers in runs of homozygosity to measure inbreeding

**Borghild Hillestad**

EAAP 2013, Nantes

Harald Grove (UMB)  
Dag Inge Våge (UMB)  
Theo Meuwissen (UMB)  
John Arthur Woolliams (UoE)  
Gunnar Klemetsdal (UMB)



## Aknowledgements

- Trygve Roger Solberg and Morten Svendsen from
- Solomon Antwi Boison and Johann Sölkner from

The logo for 'geno' is written in a bold, lowercase, teal-colored sans-serif font. A thin teal arc is positioned below the letters 'o' and 'n'.

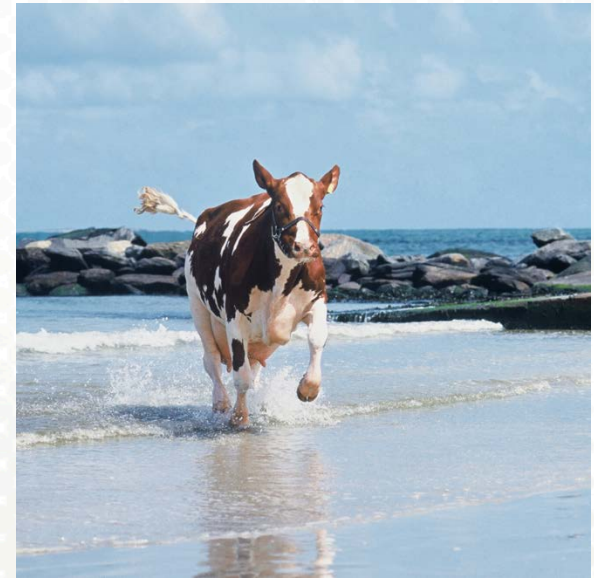
## Introduction

- SNP-density and genotyping quality controls
- Runs of homozygosity (ROH)
  - Long runs = recent inbreeding
  - Short runs = ancient inbreeding
- ROH-criteria and definition



## Objective

- Investigate how the different SNP-densities affect the correlation between  $F_{ROH}$  and  $F_{Ped}$  in Norwegian Red cattle
- Observe how different genotyping quality controls affect the discovery of ROH,  $F_{ROH}$  and its correlation to  $F_{Ped}$



## Genotyping and density sets

- 384 Norwegian Reds, Illumina HD chip, 777 K
- 3 289 Norwegian Reds, Illumina Bead chip, 54 K
- 4 different quality controls
  - 10 main density sets
  - 4 control density sets



## Genotyping quality controls

### Main sets:

- GC score > 0.7
- Autosomal SNPs only
- Individual call rate > 95 %
- SNP call rate > 90 %
- HWE  $p < 10^{-6}$

### Control sets:

- 610K(MAF>0.01)
  - as main sets
  - MAF > 0.01
- 597K(MAF>0.02)
  - as main sets
  - MAF > 0.02
- 48K(GS) and 539K(GS)
  - MAF > 0.05
  - Mendelian error rate < 2.5 %
  - Autosomal SNPs only



## SNP density sets

Name	Exact no of SNPS	SNP's coverage of bp	No. of animals
707K	707 609	2 509 326 386	381
530K	530 706	- 297 975	381
400K	398 029	- 888 366	381
300K	298 521	- 1 567 586	381
220K	223 890	- 1 839 909	381
167K	167 917	-2 038 250	381
125K	125 937	- 3 486 335	381
95K	94 452	- 4 076 257	381
70K	70 839	- 4 551 967	381
53K	53 129	- 6 074 180	381
Control sets			
610K(MAF>0.01)	610 885	- 1 167 558	381
597K(MAF>0.02)	597 454	- 1 167 558	381
539K(GS)	539 665	-2 736 273	384
48K(GS)	48 249	37 026 606	3 289

## ROH criteria

**Minimum length of a ROH**

Density	Min. no. of Kb	Min. no. of SNPs
48-95 K	2 000	15
125-300 K	1 000	25
400-707 K	500	50

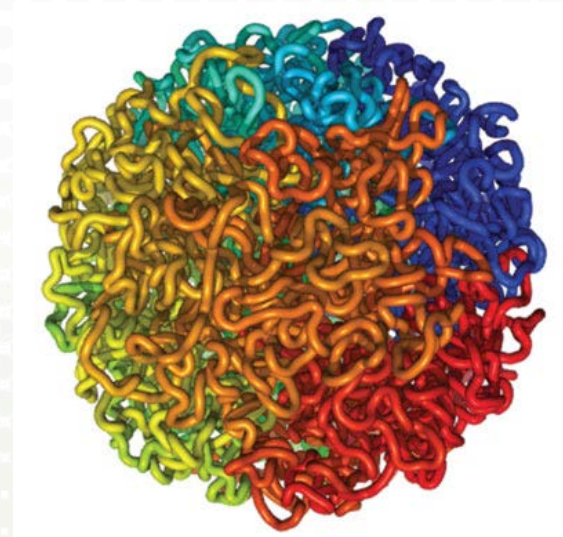


## F-value based on ROH

$$F_{ROH} = \frac{\sum L_{ROH}}{L_{AUTO}}$$

- $L_{ROH}$  = total length of ROH
- $L_{AUTO}$  = total length of genome covered by SNPs

([McQuillan, R. et al. 2008](#))



# Pedigree sets

Density sets	Pedigree depth	Number of animals	Genotyped animals	Birth years	$F_{Ped}$
53-707K, 597K(MAF>0.02) & 610K(MAF>0.01)	Deep	8 646	381	1875-2004	0.020
	5 generations	2 588	381	1935-2004	0.017
	3 generations	1 279	381	1947-2004	0.009
48K(GS)	Deep	30 005	3 289	1875-2009	0.022
	5 generations	16 129	3 289	1930-2009	0.019
	3 generations	9 714	3 289	1947-2009	0.012
539K(GS)	Deep	8 627	384	1875-2004	0.020
	5 generations	2 591	384	1935-2004	0.017
	3 generations	1 286	384	1947-2004	0.010

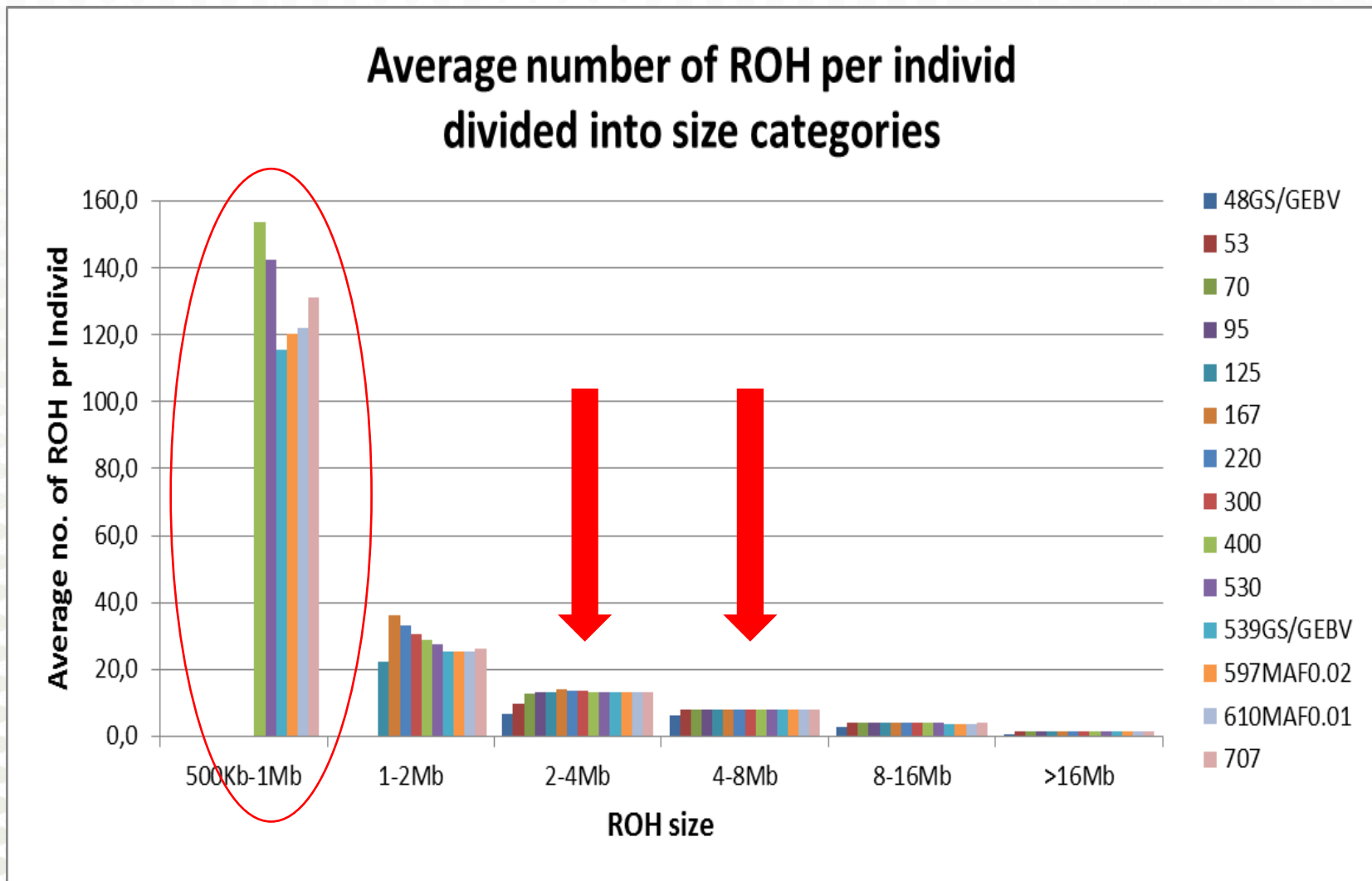


## Results

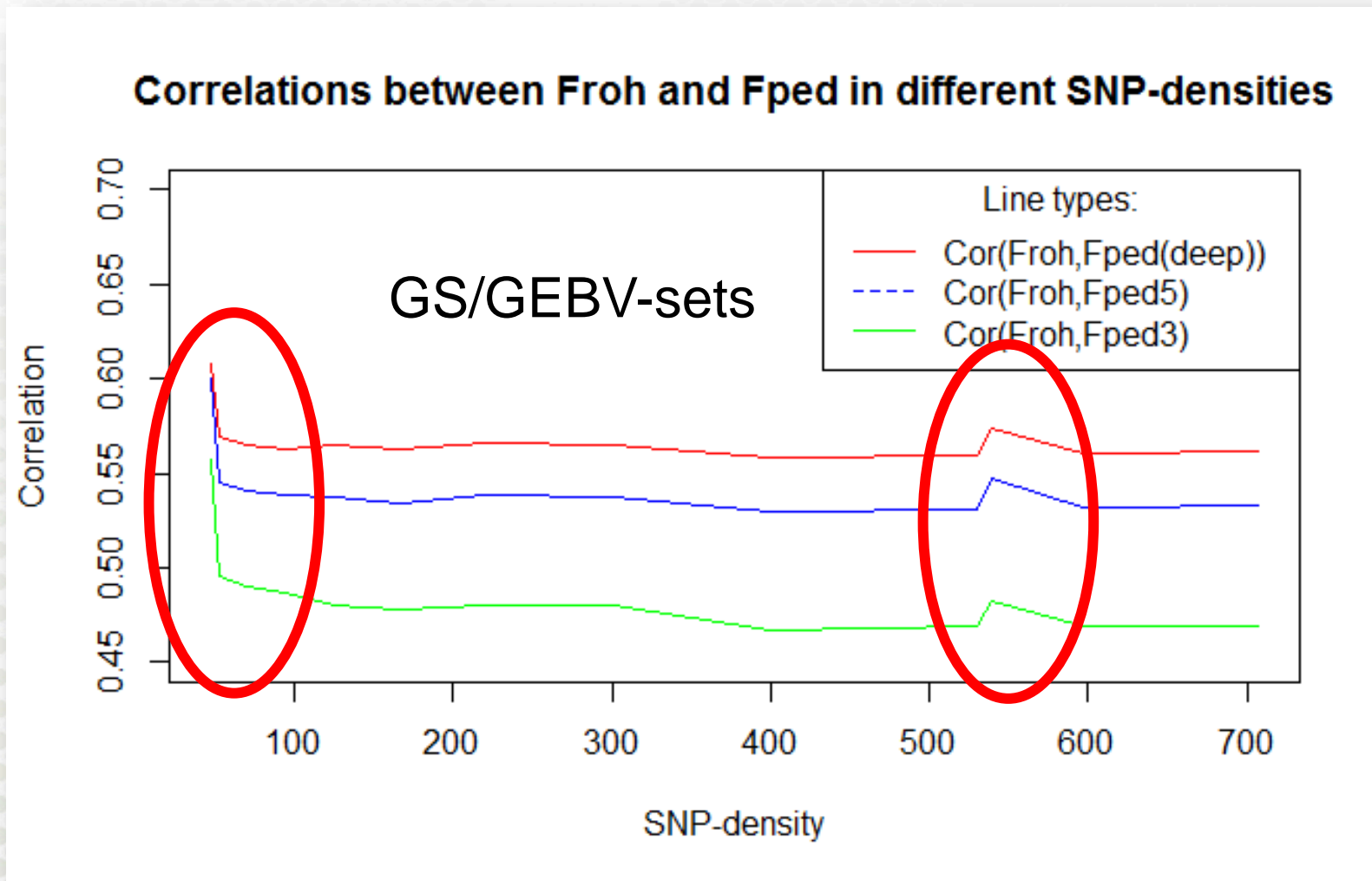
Name	No. of ROH pr animal	F-values
707K	184,1	0.111
530K	196,4	0.115
400K	209,1	0.119
300K	57,3	0.080
220K	60,3	0.082
167K	63,6	0.085
125K	48,4	0.075
95K	26,3	0.063
70K	26,2	0.063
53K	23,4	0.061
Control sets		
610K(MAF < 0.01)	173,5	0.107
597K(MAF < 0.02)	171,8	0.107
539K(GS)	166,7	0.105
48K(GS)	16,6	0.041



# Results



## Results



## Concluding marks

- SNP density does not affect detection of ROH > 4 Mb
- High SNP densities discovers more ROH < 1 Mb
- $F_{ROH}$  increase with SNP density
- $(F_{ROH}, F_{ped})$ -correlations are affected by pedigree depth, not SNP density
- Different genotyping quality controls might affect the results

*"It is that range of biodiversity that we must care for – the whole thing – rather than just one or two stars."*

David Attenborough, [BBC Interview](#)



Thank you for  
your attention!