

What do we mean by runs of homozygosity? Assessing the effect of parameters involved in their detection



S. T. Rodríguez-Ramilo and J. Fernández

Introduction

ROH: long, uninterrupted stretches of homozygous genotypes

ROH can arise by different mechanisms...
...but the first cause seems to be **inbreeding (F)**

Long ROH: recent inbreeding

Short ROH: ancient inbreeding

Introduction

The **theoretical** definition of a ROH is straight forward...

BUT...

The use of **different parameters** for ROH detection may lead to different estimates of F

Objective

To investigate the effect of the different parameters involved in the detection of ROH...

...using animals from the Spanish Holstein population genotyped with the BovineSNP50 BeadChip



Genomic and pedigree data

- **36,693** autosomal SNPs from the Illumina BovineSNP50 BeadChip
- **10,569** genotyped individuals
- **31,203** individuals in genealogy (5 generations of ancestors of genotyped individuals)

Inbreeding estimates

- Pedigree-based inbreeding (F_{ped})
- SNP-by-SNP-based inbreeding (F_{snp}):

Proportion of homozygous SNPs

- ROH-based inbreeding (F_{roh}):

$$F_{ROH_i} = \frac{\sum_{k=1}^{n_{ROH_i}} l_{ROH_{ik}}}{l_g}$$

where n_{ROH_i} is the total number of ROHs in individual i , $l_{ROH_{ik}}$ is the length of the k^{th} ROH in individual i in base pairs and l_g is the length of the genome in base pairs

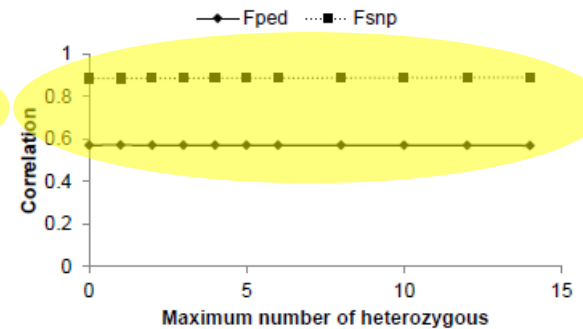
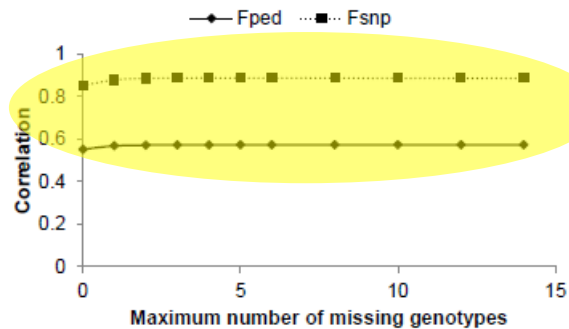
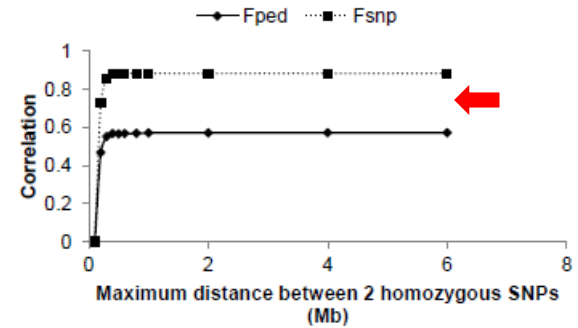
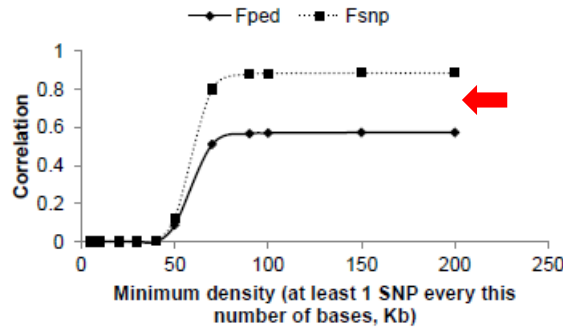
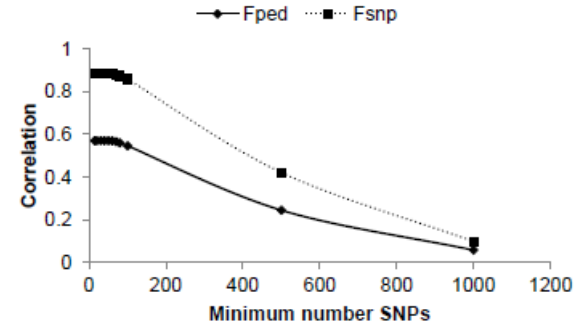
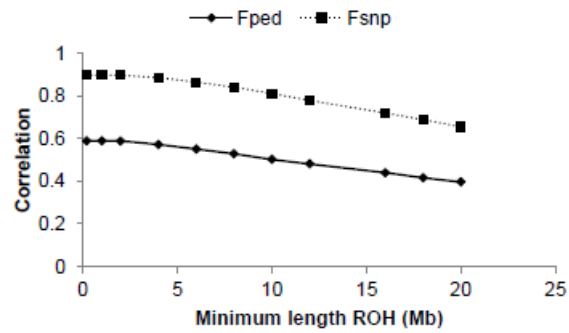
Correlation between F_{ped} (or F_{snp}) and F_{roh}

Parameters involved in the detection of ROH

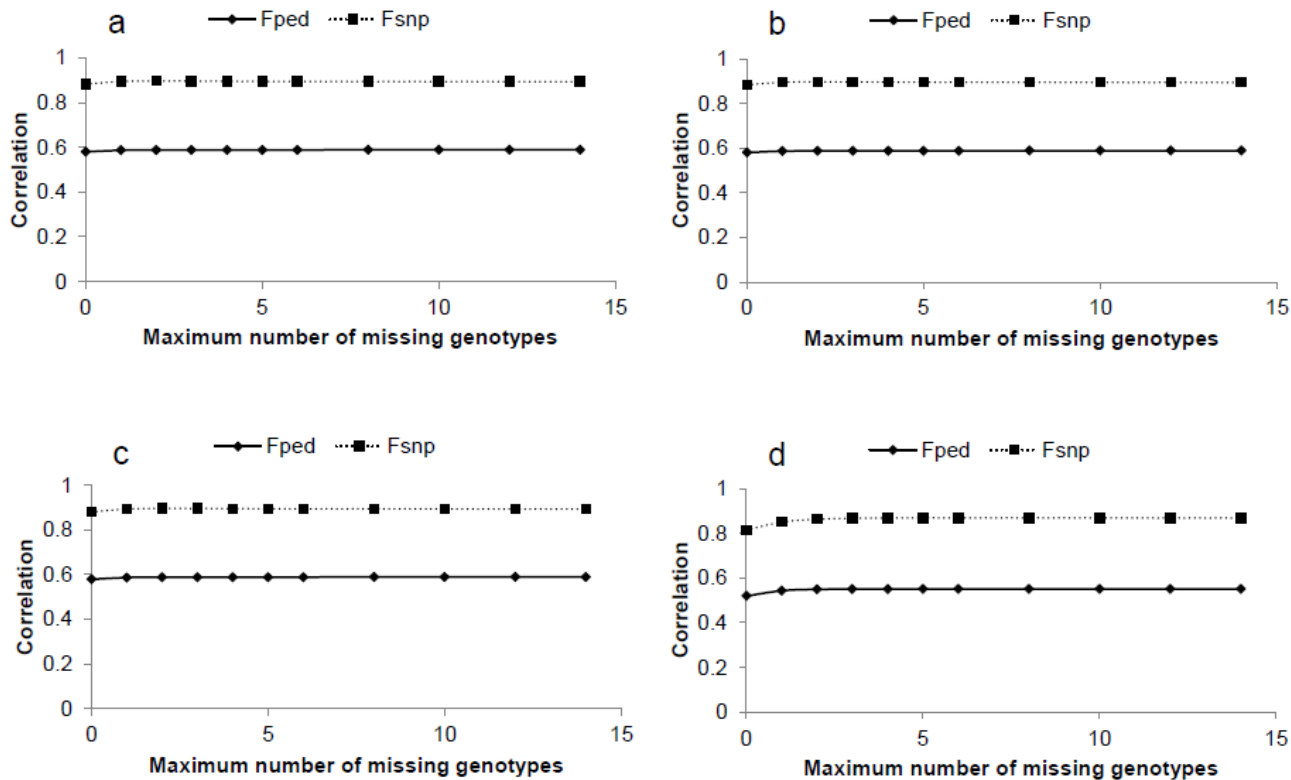
- The minimum length of a ROH
- The minimum number of SNPs within a ROH
- The minimum marker density of a ROH
- The maximum distance between two adjacent SNPs
- The maximum number of missing genotypes
- The maximum number of heterozygous SNPs admitted within a ROH

Results

Correlation
between
Fped (or Fsnp)
and Froh

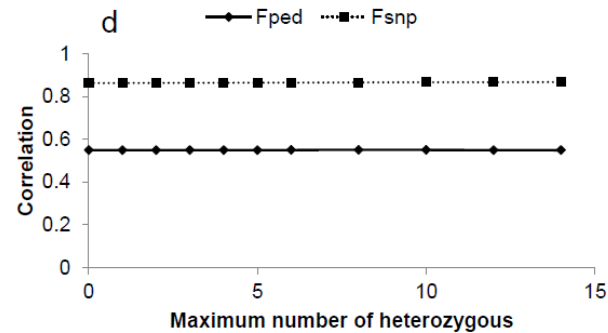
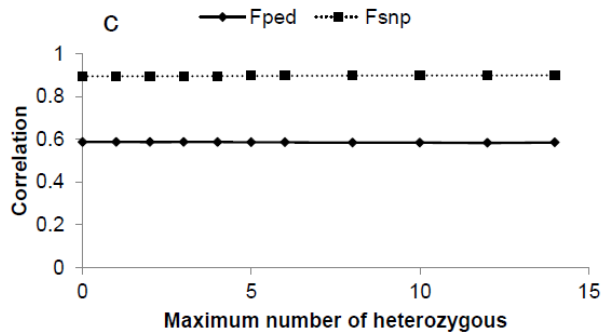
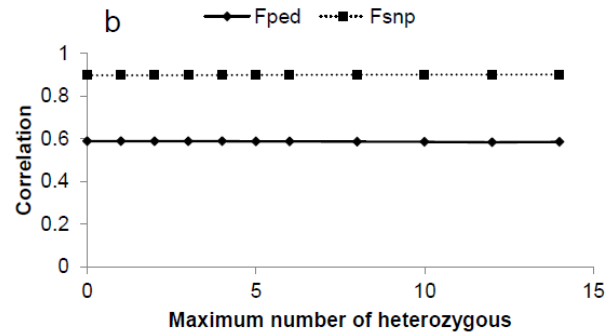
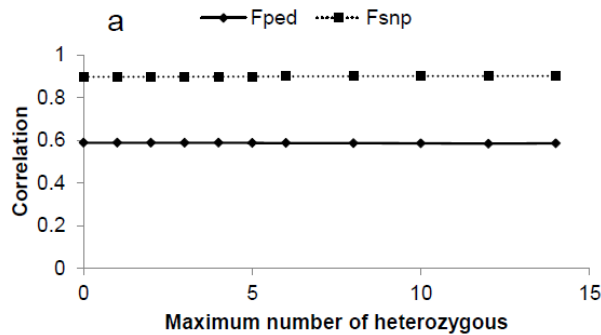


Results



Correlation between **Fped** (or **Fsnp**) and **Froh** when evaluating the maximum number of missing genotypes for different length of a ROH: (a) **0.2 Mb**, (b) **1 Mb**, (c) **2 Mb** and (d) **6 Mb**

Results



Correlation between **Fped** (or **Fsnp**) and **Froh** when evaluating the maximum number of heterozygous for different length of a ROH: (a) **0.2 Mb**, (b) **1 Mb**, (c) **2 Mb** and (d) **6 Mb**

Take home message

- The maximum number of missing genotypes and the maximum number of heterozygous have **no effect** in the identification of ROH
- The other four factors seem to have **effect** on ROH-based inbreeding
- The **optimal combination** of parameters could depend on the population characteristics and the type of SNP-chip available

Thanks for your attention