Instituto Nacional de Investigación
y Tecnología Agraria y Alimentaria

Genomic Inbreeding Coefficients Based on the Length of Runs of Homozygosity

L. Gomez-Raya, C. Rodríguez, C. Barragán,
W.M. Rauw and L. Silió

E-mail: gomez.luis@inia.es

INTRODUCTION

Genomic Inbreeding Coefficients

-Molecular Inbreeding Coefficient (proportion of homoz. sites)

-Total ROH content (proportion)

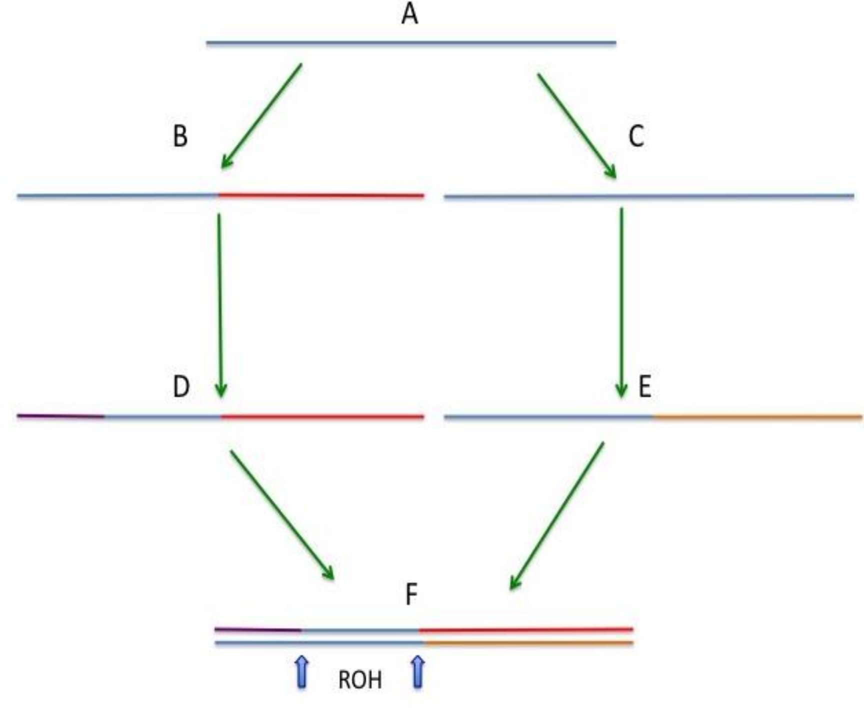
↓

NOT ACCOUNT FOR BY ROH LENGTH

↓

recent or old inbreeding

Scheme representing the generation of a ROH



Factors

1. the number of steps in the paths from “A” to “F” (opportunities for recombination)

2. the recombination rate in a chromosome

3. the length of the chromosome

Objectives

1- Propose novel genomic inbreeding coefficients based on ROH length.

2.- Compare genomic inbreeding coefficients in Iberian pigs (Torbiscal strain).

METHODS

ROH length as a random variable

- number of contiguous SNPs in homozygosis

- length covered by homozygous SNPs measured in units of physical distance in Mb

METHODS

Genomic Inbreeding Coefficients based on ROH length

NON PARAMETRIC

a) Kolmogorov-Smirnov Inbreeding Coefficients

b) Quantile Inbreeding Coefficients

PARAMETRIC

c) Exponential Inbreeding Coefficients

Kolmogorov-Smirnov Inbreeding Coeff.

The Kolmogorov–Smirnov statistic (*D*) quantifies the distance between a cumulative distribution (T) and the cumulative distribution of a reference distribution (S) and has value:

$$\boldsymbol{D = max_x \left| F_T(x) - F_S(x) \right|}$$

The new inbreeding coefficient based on the KS test is:

$$\boldsymbol{F_{ROH-KS} = max_x \left(F_{I-ROH}(x) - F_{A-ROH}(x)\right)}$$

Quantile Inbreedig Coefficients

The *ROH-Q* inbreeding coefficient is defined as the sum of the distances between the quantiles of the CDF of a given individual ($q_{I\text{-}ROH}$) with respect to the reference population ($q_{A\text{-}ROH}$):

$$F_{ROH-Q} = \sum_{i=1}^{nq} \frac{q_{I-ROH} - q_{A-ROH}}{\sqrt{2}}$$

where *nq* is the number of quantiles $q_{I\text{-}ROH}$ and $q_{A\text{-}ROH}$ are the quantiles of the CDF of ROH-length from one individual or any reference distribution.

LEAD TO A GRAPHICAL REPRESENTATION

Exponential Inbreeding Coefficients

a) mean of the exponential density:

$$F_{ROH-E^m} = \frac{1}{\lambda_{I-ROH}} - \frac{1}{\lambda_{A-ROH}}$$

b) variance of the exponential density:

$$F_{ROH-E^v} = \frac{1}{\lambda^2_{I-ROH}} - \frac{1}{\lambda^2_{A-ROH}}$$

DEPENDS ON THE MEAN LENGTH OF ROH

Exponential-p Inbreeding

$$F_{ROH-E^{p}} = \int_{T}^{\infty} \lambda_{I-ROH}\, e^{-(\lambda_{I-ROH})\,x}\, d(x)$$

If T = 1 Mb → prob that one ROH drawn at random be larger than 1Mb

it is forced to be between 0 and 1 (being a probability).

Animal Material

217 Torbiscal sows

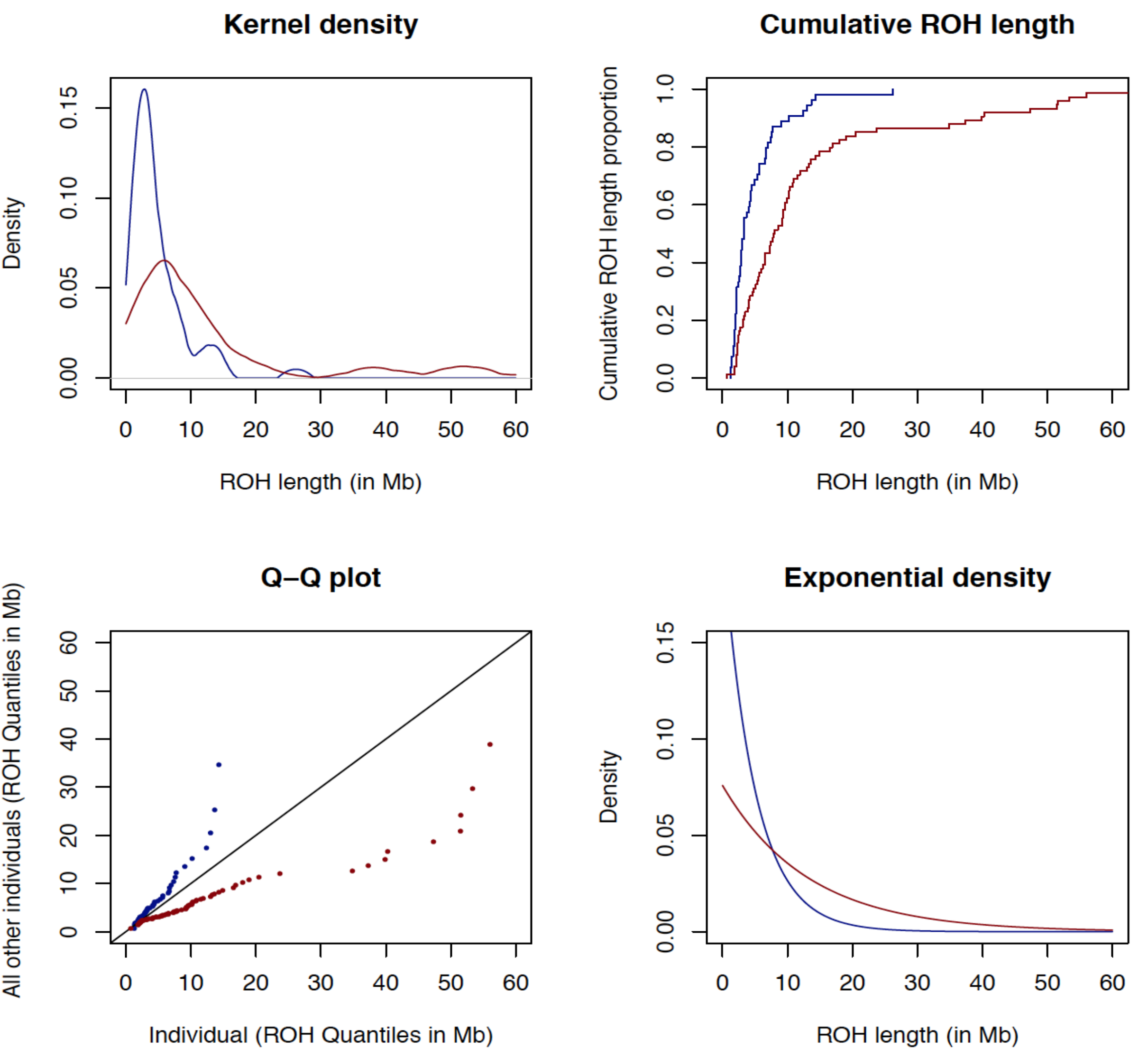
Full genealogical pedigree recording

new (from 1980) and old inbreeding (until 1980)

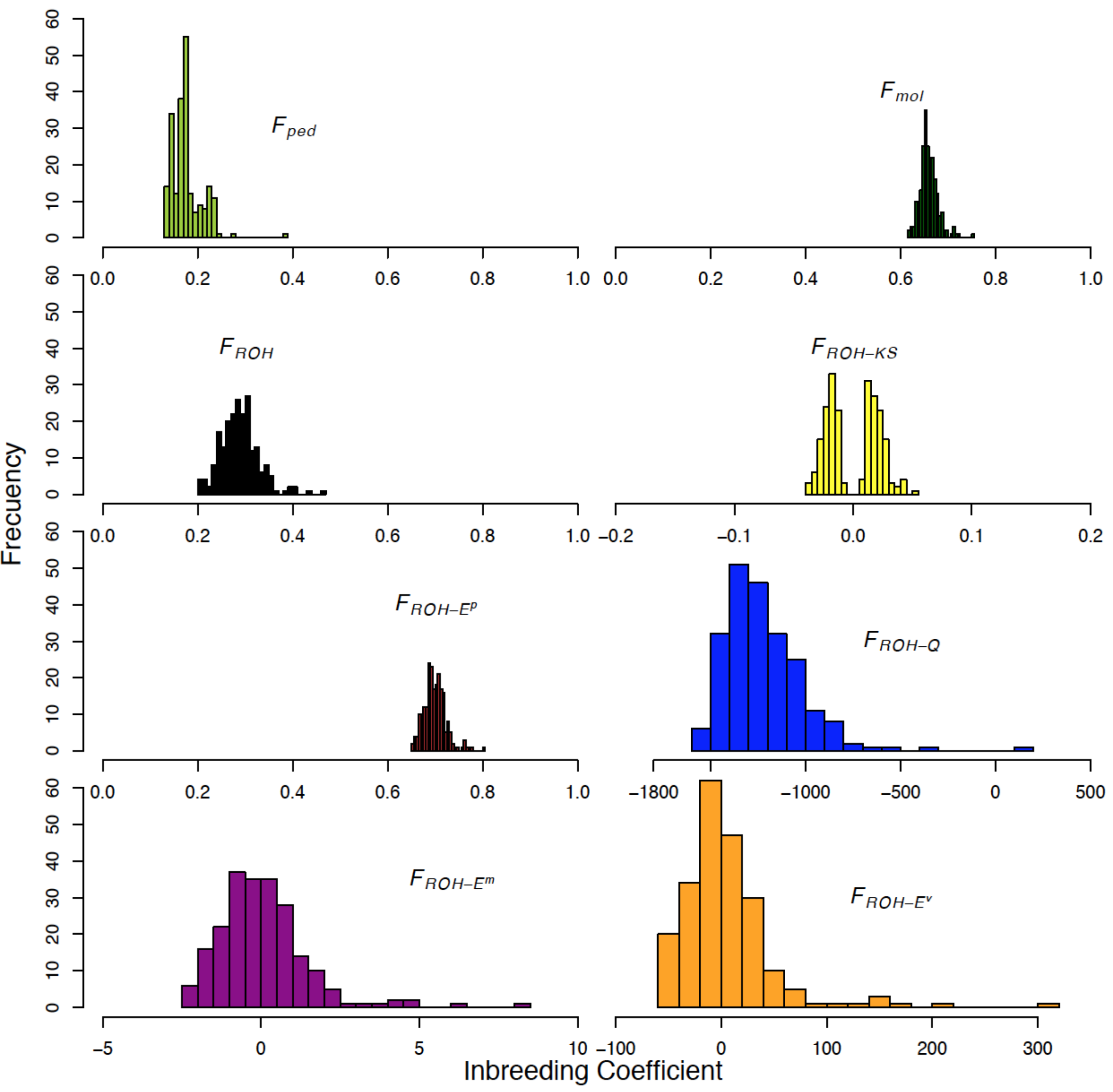
Genotyped with Illumina array (62,163 SNP)

RESULTS

Distributional properties of the distribution of ROH length. Kernel density, Q-Q plots, cumulative ROH length, and adjusted exponential density for the distributions of the length of ROH of two individuals with extreme pedigree inbreeding coefficient: sow 18705308 (highest inbreeding; dark red) and sow 13304804 (lowest inbreeding; dark blue). SNPs >35, ROH-length in Mb**.**



Histograms of frequencies of pedigree and genomic inbreeding coefficients

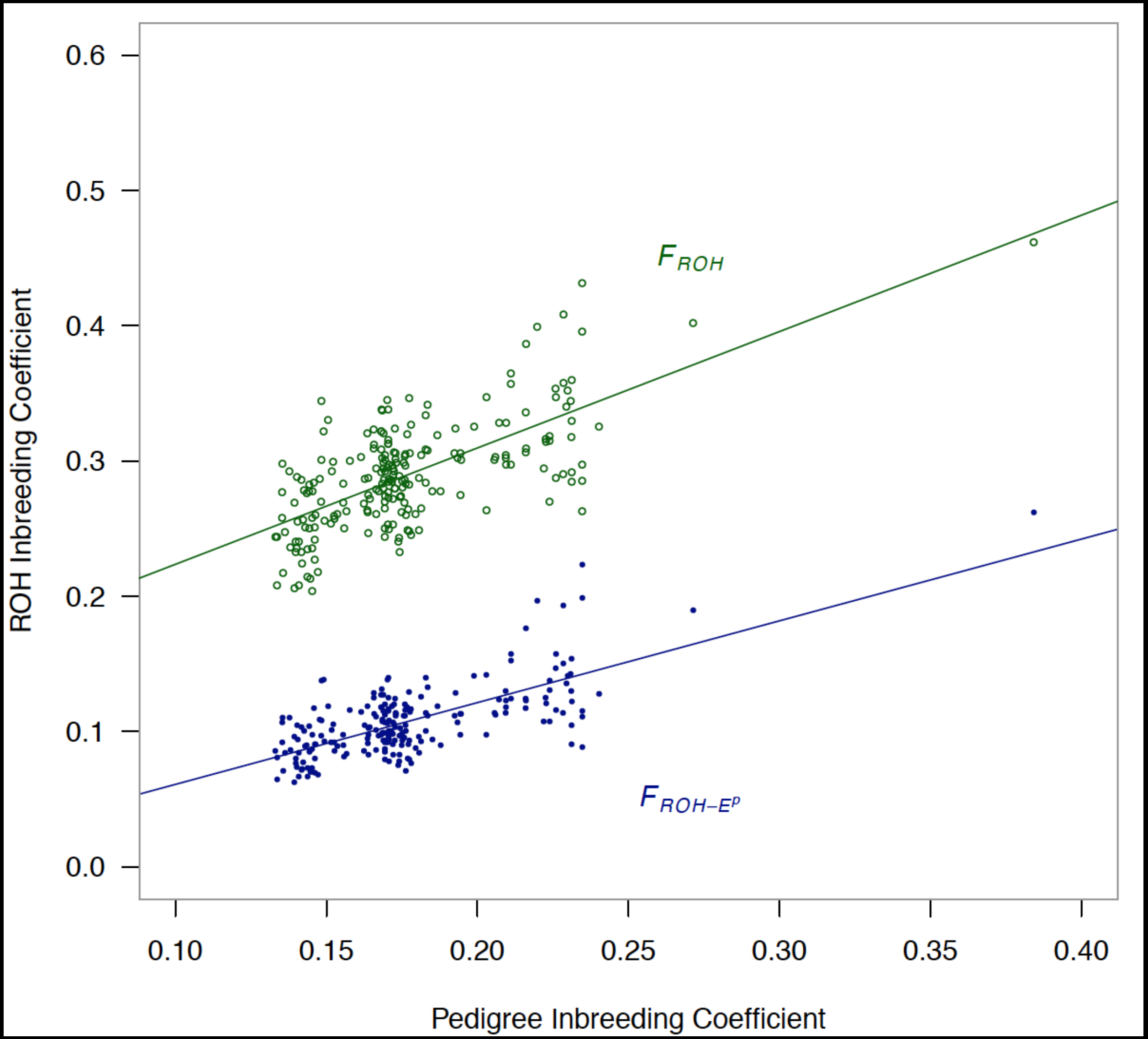


Correlations between current and novel genomic inbreeding coefficients based on ROH length in Mb. Standard errors ranged between 0.015 and 0.068.

Number of SNP > 5

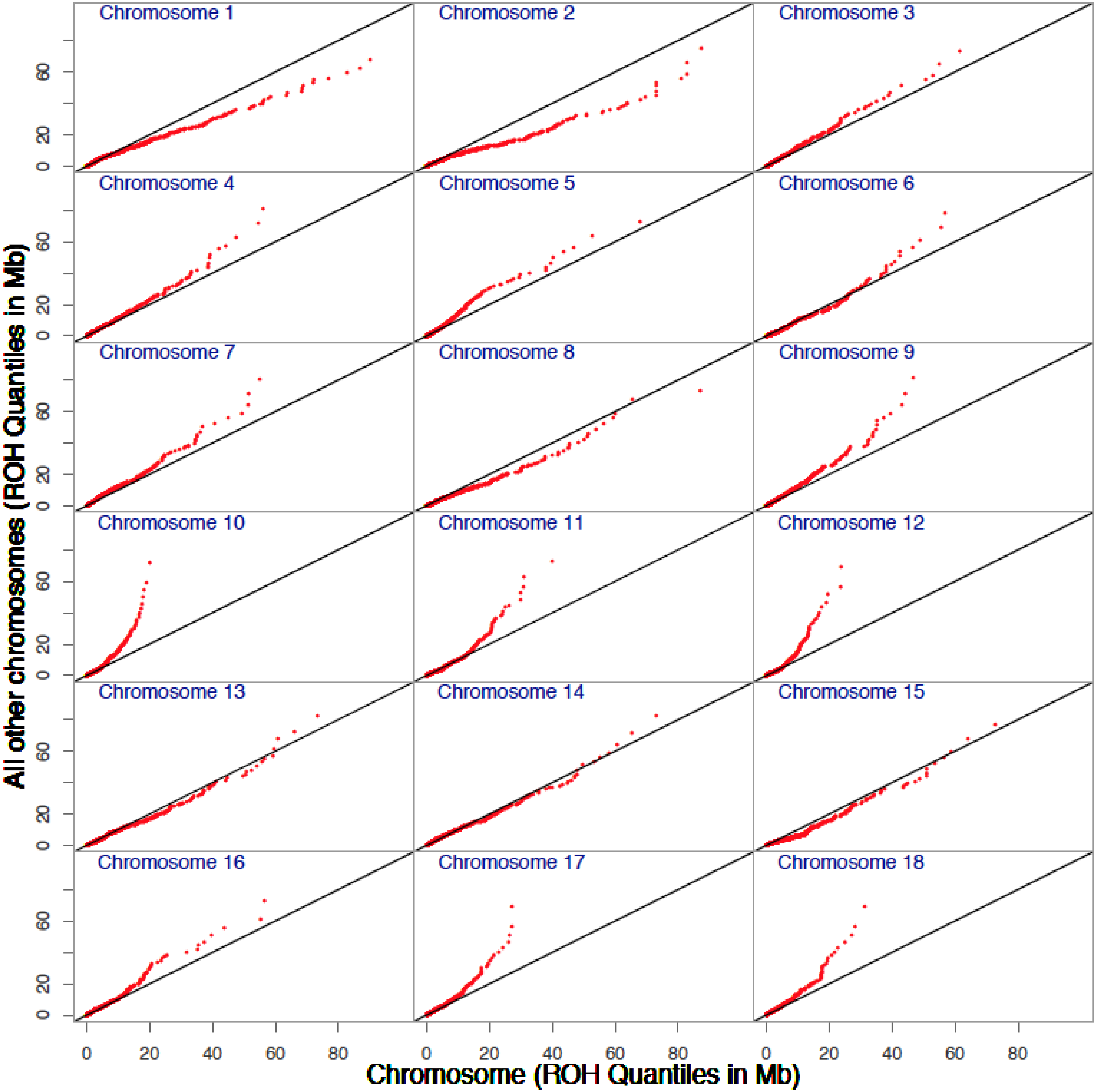
	$F_{ROH\text{-}KS}$	$F_{ROH\text{-}Q}$	$F_{ROH\text{-}E}^{m}$	$F_{ROH\text{-}E}^{v}$	$F_{ROH\text{-}E}^{p}$
F_{ped}	0.166	0.510	0.701	0.713	0.704
$F_{ped\text{-}new}$	0.169	0.510	0.707	0.714	0.705
$F_{ped\text{-}old}$	-0.081	0.150	0.158	0.151	0.161
F_{Mol}	0.317	0.696	0.960	0.939	0.960
F_{ROH}	0.295	0.717	0.956	0.931	0.956

Regressions of $F_{ROH\text{-}E^p}$ and F_{ROH} on pedigree inbreeding coefficients (F_{ped}).



Genomic autozygosity in chromosomes

SSC	# of SNPs (fixed)	# ROH
1	3,898 (1,038)	32,650
2	2,447 (1,197)	20,672
3	1,785 (387)	14,939
4	2,313 (647)	18,918
5	1,539 (433)	13,689
6	2,138 (796)	17,758
7	2,269 (688)	19,240
8	2,010 (741)	16,714
9	2,296 (703)	20,427
10	1,207 (404)	10,727
11	1,407 (506)	12,629
12	992 (344)	9,145
13	2,754 (982)	24,311
14	2,508 (977)	20,948
15	2,066 (748)	16,858
16	1,322 (493)	12,041
17	1,033 (337)	9,052
18	1,039 (281)	9,032
Total	35,023 (11,702)	299,750


Chromosome 1
Chromosome 2
Chromosome 3
Chromosome 4
Chromosome 5
Chromosome 6
Chromosome 7
Chromosome 8
Chromosome 9
Chromosome 10
Chromosome 11
Chromosome 12
Chromosome 13
Chromosome 14
Chromosome 15
Chromosome 16
Chromosome 17
Chromosome 18
All other chromosomes (ROH Quantiles in Mb)
0
20
60
Chromosome (ROH Quantiles in Mb)
0
20
40
60
80

Pearson correlations (upper diagonal) between pairs of chromosomal genomic inbreeding coefficients and length (number of SNPs per chromosome). Lower diagonal are approximate standard errors.

	Length	F_{Mol}	F_{ROH}	$F_{ROH\text{-}KS}$	$F_{ROH\text{-}Q}$	$F_{ROH\text{-}E}^{m}$	$F_{ROH\text{-}E}^{p}$
Length		0.04	0.16	0.57	0.84	0.57	0.57
F_{Mol}	0.25		0.45	0.01	-0.07	0.44	0.45
F_{ROH}	0.25	0.25		0.32	0.13	0.51	0.52
$F_{ROH\text{-}KS}$	0.21	0.25	0.24		0.60	0.83	0.82
$F_{ROH\text{-}Q}$	0.14	0.25	0.25	0.20		0.59	0.58
$F_{ROH\text{-}E}^{m}$	0.20	0.22	0.21	0.14	0.20		0.99
$F_{ROH\text{-}E}^{p}$	0.21	0.22	0.21	0.14	0.20	0.00	

CONCLUSSIONS

a) The newly proposed inbreeding coefficients add to existing methods to estimate inbreeding by accounting for the length of ROH, which incorporates information on recent inbreeding.

b) Among the proposed methods, quantile inbreeding coefficients might be the most sensitive for identifying individuals with longer ROH fragments.

c) Exponential-p inbreeding coefficients are less sensitive for detecting long ROH fragments but are defined as a probability (they range between 0 and 1) and are suitable for comparison of individuals across populations.

AKNOWLEDGEMENTS

1. Financial support was provided by RTA2011-00113 and RZ2012-00006 INIA grants.

2. We acknowledge the effort of Jaime Rodrigáñez and all the staff of the Iberian pig farm ‘Dehesón del Encinar’ for maintaining with strict pedigree and data recording the Torbiscal pigs and their ancestors since 1944 until the recent closure of farm facilities.

THANKS!

INIA
Instituto Nacional de Investigación
y Tecnología Agraria y Alimentaria

Genomic Inbreeding Coefficients Based on the Length of Runs of Homozygosity

L. Gomez-Raya, C. Rodríguez, C. Barragán,
W.M. Rauw and L. Silió

E-mail: gomez.luis@inia.es