



Inbreeding and effective population size in French dairy sheep: Comparison between genomic and pedigree estimates

S. T. Rodríguez-Ramilo,* J. M. Elsen, and A. Legarra
INRA, UMR 1388 GenPhySE, 31326 Castanet Tolosan, France

ABSTRACT

Before availability of dense SNP data, genetic diversity was characterized and managed with pedigree-based information. Besides this classical approach, 2 methodologies have been proposed in recent years to characterize and manage diversity from dense SNP data: the SNP-by-SNP approach and the alternative based on runs of homozygosity (ROH). The establishment of criteria to identify ROH is a current constraint in the literature dealing with ROH. The objective of this study was, using a medium-density SNP chip, to quantify by 3 methods (pedigree, SNP-by-SNP, and ROH) the genetic diversity on 5 selected French dairy sheep subpopulations and breeds and to assess the effect of the definition of ROH on these estimates. The data set available included individuals from the breeds Basco-Béarnaise, Manech Tête Noire, Manech Tête Rousse, and 2 subpopulations of Lacaune: Lacaune Confederation and Lacaune Ovitest. Animals were genotyped with the Illumina OvineSNP50 BeadChip (Illumina Inc., San Diego, CA). After filtering, the genomic data included 38,287 autosomal SNP and 8,700 individuals, which comprised 72,803 animals in the pedigree. The results indicated that no significant differences were observed in effective population size estimates obtained from pedigree or genomic (SNP-by-SNP or ROH) information. In general, estimates of effective population size were above 200 in Lacaune Confederation and Lacaune Ovitest subpopulations and below 200 in Basco-Béarnaise, Manech Tête Noire, and Manech Tête Rousse breeds. The minimum length that constituted a ROH, the minimum number of SNP that constituted a ROH, as well as the minimum density and the maximum distance allowed between 2 homozygous SNP are ROH-defining factors with important implications in the estimation of the rate of inbreeding. The ROH-based rates of inbreeding in concordance

with those obtained from pedigree information require a specific set of values. This particular set of values is different from that identified to obtain ROH-based rates of inbreeding similar to those obtained on a SNP-by-SNP basis. Factors to define ROH do not change the results much unless extreme values are considered, although further research on ROH-based inbreeding is still required.

Key words: runs of homozygosity, rate of inbreeding, effective population size

INTRODUCTION

The preservation of genetic diversity is usually achieved by maximizing the effective population size or, equivalently, by minimizing the rate of inbreeding in a population. Genomic selection can reduce the rate of inbreeding per generation but can lead to a higher rate of inbreeding per year (Daetwyler et al., 2007). Such an increase in the rate of inbreeding can result in lower genetic diversity, lower response to selection, and a higher exposure of deleterious alleles (Curik et al., 2014). For this reason, controlling the rate of inbreeding is an important goal in breeding programs.

Before the widespread availability of dense SNP data, genetic diversity was characterized and managed with pedigree-based inbreeding estimates. This approximation refers to the proportion of the genome that is expected to be identical by descent (**IBD**). However, this pedigree-based approach has several constraints. First, pedigree completeness and quality are essential (Oliehoek and Bijma, 2009). Second, pedigree information does not take into account either Mendelian sampling variation (Hill and Weir, 2011) or linkage disequilibrium caused by selection (e.g., Smith and Haigh, 1974).

The realized proportion of the genome that 2 individuals share can be more accurately estimated from genome-based information than from pedigree-based information (Howard et al., 2017). Two alternatives have been proposed to characterize and manage diversity from dense SNP data. The first is the SNP-by-SNP approach (e.g., VanRaden, 2008), which involves the es-

Received July 18, 2018.

Accepted December 23, 2018.

*Corresponding author: silvia.rodriguez-ramilo@inra.fr

timization of the alleles that are identical by state (**IBS**) for each individual SNP. Consequently, it captures relationships caused by common ancestors going back to a base population where all alleles are supposed to be unique. The second is the segment-based approach (e.g., McQuillan et al., 2008), which considers IBS segments, rather than individual SNP. The segments of homozygous genotypes are called runs of homozygosity (**ROH**). The ROH-based inbreeding estimates could capture the relatedness between nominally unrelated founders of the pedigree (Peripolli et al., 2017), consider the stochastic mechanism of recombination (Keller et al., 2011), and also consider the linkage disequilibrium between loci caused by selection (Curik et al., 2002). The ROH could arise due to autozygosity, where the same chromosomal segment has been passed to an offspring from parents who inherited it from a common ancestor (Broman and Weber, 1999). Accordingly, ROH with ancient origin are expected to be shorter, because recombination from repeated meiosis breaks up IBD segments. On the other hand, ROH due to recent inbreeding are expected to be longer because the probability of breaking up IBD segments from recombination is reduced.

There is a need to establish consistent and reproducible criteria for identifying and quantifying ROH. The comparison between studies is challenging because inconsistency is high among the criteria used to identify ROH both within and also between species (Ku et al., 2011; Peripolli et al., 2017). This lack of standard criteria for identifying ROH enhances the probability of introducing bias on ROH-based inbreeding estimates. Accordingly, some parameters and thresholds imposed during the sequence analysis can affect ROH identification (Howrigan et al., 2011). In addition, pruning SNP that show low minor allele frequency, that deviate from Hardy-Weinberg equilibrium, or show high linkage disequilibrium can influence ROH characterization (Albrechtsen et al., 2010). Likewise, the number of heterozygous genotypes and the SNP chip density also can influence ROH identification in cattle (Purfield et al., 2012; Ferenčaković et al., 2013; Mastrangelo et al., 2016). Several studies have evaluated ROH across all major livestock species [see Peripolli et al. (2017) for a review]. However, further exploration is desirable to better understand the criteria for estimating ROH-based inbreeding.

Selection of French dairy sheep has been implemented for each local breed separately. The largest breed is the Lacaune, which consists of 2 subpopulations: Lacaune Confederation (**LACCon**) and Lacaune Ovitest (**LACovi**). The population in the western Pyrenees Mountains consists of 3 breeds: Manech Tête Rouse

(**MTR**), Manech Tête Noire (**MTN**), and Basco-Béarnaise (**BB**). Breeding programs, started in the sixties, are now clearly profitable (Larroque et al., 2014) and genomic selection is being established (Legarra et al., 2014). Accordingly, genomic selection should include genomic management of inbreeding (Sonesson et al., 2012); however, it is still unclear how to implement genomic management of diversity within these breeds.

The objectives of our study were the following. First, to quantify the genetic diversity in 5 selected French sheep subpopulations and breeds with several pedigree- and marker-based methods and to compare the existing results in terms of rate of inbreeding and on effective population size. Second, to establish the robustness of ROH-based measures of inbreeding to the criteria used to define a ROH. We used a medium-density SNP chip to complete these objectives.

MATERIALS AND METHODS

Data

Rams were genotyped with the OvineSNP50 Bead-Chip (Illumina Inc., San Diego, CA). Supplemental Figure S1 (<https://doi.org/10.3168/jds.2018-15405>) shows the distribution of progeny-tested rams genotyped per year of birth. The genotyped animals in the BB breed were born between 1999 and 2008; in the MTN breed between 1996 and 2007; in MTR breed between 1999 and 2009; in the LACCon subpopulation between 1996 and 2012; and in the LACovi subpopulation between 1999 and 2012.

The SNP quality control included the absence of parent-offspring Mendelian segregation incompatibilities (<3%), SNP call rate >97%, and large deviations from Hardy-Weinberg equilibrium [SNP with $P < 10^{-6}$ were discarded; see Baloche et al. (2014) and Legarra et al. (2014) for more details]. The final data set included 38,287 autosomal SNP and 8,700 genotyped animals. The pedigree data set was constructed with all known ancestors of the genotyped individuals and comprised 72,803 animals, with all breeds merged. Table 1 shows per subpopulation and breed the number of genotyped animals, the individuals included in the pedigree, and the equivalent number of complete generations.

Pedigree-Based Inbreeding Estimates

The pedigree-based inbreeding estimates (F_{PED}) of the genotyped individuals were obtained using the software PEDIG (Boichard, 2002) with the option that implements the algorithm of Meuwissen and Luo (1992). The generation interval was also calculated us-

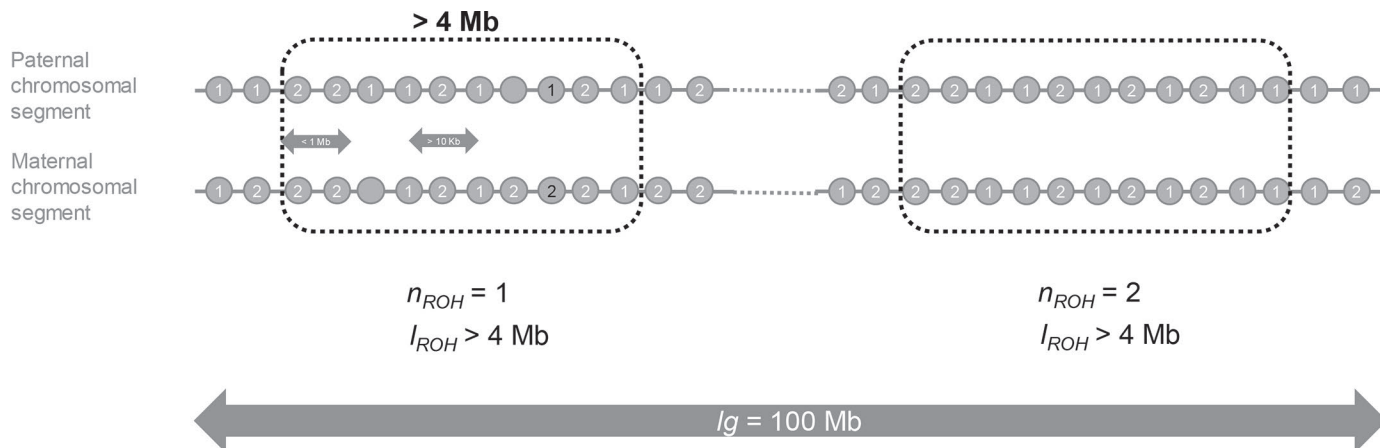


Figure 1. Example of run of homozygosity (ROH) within the dotted squares. The first allele is indicated by 1 and the second allele is indicated by 2 within the circles. The minimum length that constituted a ROH was 4 Mb, the minimum number of SNP was 10, the minimum density was 1 SNP per 10 kb, the maximum distance allowed between 2 consecutive homozygous SNP in a run was 1 Mb, a maximum of 2 missing genotypes were allowed (empty circles,) and 1 heterozygous genotype within a particular ROH was permitted (alleles indicated in black). n_{ROH} refers to the number of ROH, l_{ROH} indicates the length of the ROH, and l_g is the length of the genome.

ing PEDIG and was 4.95, 4.51, 4.20, 3.61, and 3.55 for BB, MTN, MTR, LACCon, and LACovi, respectively, in accordance with previous estimates (Larroque et al., 2014).

SNP-by-SNP Inbreeding Estimates

Following the logic of Malécot (1948), IBS relationships are twice the probability that 2 alleles taken at random, 1 per individual, are identical. For 1 individual with itself, the sampling is with replacement, and the relationship reduces to 1 + the probability that 1 allele is identical to the other. Thus, the inbreeding coefficient based on individual SNP (F_{SNP}) of individual i (F_{SNP_i}) is simply the probability that the 2 alleles are identical, measured across the genome, and was equal to the proportion of typed loci at which this individual is homozygous (Silió et al., 2013).

ROH-Based Inbreeding Estimates

Runs of homozygosity are long, uninterrupted stretches of homozygous genotypes (McQuillan et al., 2008). More specifically, the inbreeding estimator (F_{ROH}), is the proportion of the genome that is in ROH. For individual i , F_{ROH_i} was calculated as

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

where n_{ROH_i} is the total number of ROH 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 (which was 2,438,309,318 bp in the data set analyzed in the present study). An example of ROH is shown in Figure 1.

Table 1. Number of genotyped animals, individuals included in the pedigree, and equivalent number of complete generations per subpopulation and breed

Breed ¹	Genotyped individuals	Individuals in pedigree	Equivalent number of complete generations
BB	321	1,861	6.10
MTN	329	1,616	5.29
MTR	1,906	11,574	7.61
LACCon	3,030	29,255	10.84
LACovi	3,114	28,497	11.69

¹BB = Basco-Béarnaise; MTN = Manech Tête Noire; MTR = Manech Tête Rousse; LACCon = Lacaune Confederation; LACovi = Lacaune Ovitest.

There are 6 criteria for identifying a ROH, which consider the information provided by the SNP itself (i.e., homozygous, heterozygous, or missing genotype) and its map position in the chromosome. These 6 criteria are (1) the minimum length, in megabases, that constituted a ROH; (2) the minimum number of SNP that constituted a ROH; (3) the minimum density, at least 1 SNP in each direction, in kilobases; (4) the maximum distance allowed between 2 consecutive homozygous SNP, in megabases; (5) the maximum number of missing genotypes allowed; and (6) the maximum number of heterozygous genotypes permitted within a particular ROH. Table 2 shows the different values tested for each ROH defining factor and their default values (based on Peripolli et al., 2017).

Rate of Inbreeding and Effective Population Size

Rates of inbreeding (ΔF) per year ($\Delta F_{PED_{year}}$, $\Delta F_{SNP_{year}}$, and $\Delta F_{ROH_{year}}$) were computed from the regression coefficient of the inbreeding coefficient for each individual on the year of birth, taking into account the 8,700 animals. Rates of inbreeding per generation (ΔF_{PED} , ΔF_{SNP} , and ΔF_{ROH}) were calculated by multiplying the rates of inbreeding per year by the generation interval (L). Finally, estimates of effective population size (N_e) were obtained from the rate of inbreeding per generation as $N_{ePED} = 1/2\Delta F_{PED}$, $N_{eSNP} = 1/2\Delta F_{SNP}$, and $N_{eROH} = 1/2\Delta F_{ROH}$ (Falconer and Mackay, 1996). Confidence intervals for effective population sizes were obtained from the standard error of the rates of inbreeding per year; for example, the estimator of N_{eSNP} is $1/2L\Delta F_{SNP_{year}}$ and its confidence interval is

$$\frac{1}{2L \left[\Delta F_{SNP_{year}} + 1.96se \left(\Delta F_{SNP_{year}} \right) \right]}$$

RESULTS

Comparison of Pedigree- and Genome-Based Inbreeding Estimators

In this section, the default values used for identifying a ROH were (1) 4 Mb as the minimum length that constituted a ROH, (2) 30 as the minimum number of SNP, (3) 1 SNP/100 kb as the minimum density, (4) 1 Mb as the maximum distance allowed between 2 consecutive homozygous SNP in a run, (5) 2 missing genotypes maximum, and (6) 1 heterozygous genotype within a particular ROH. These default values are highlighted in bold in Table 2.

Table 2. Values tested for each run of homozygosity (ROH)-defining factor

Factor	Tested values ¹											
Minimum length ROH (Mb)	0.2	1	2	4	6	8	10	12	16	18	20	
Minimum number SNP	15	20	30	40	50	60	70	80	100	500	1,000	
Minimum density (at least 1 SNP every this number of bases, kb)	5	10	20	30	40	50	70	90	100	150	200	
Maximum distance between 2 homozygous SNP (Mb)	0.1	0.2	0.3	0.4	0.5	0.6	0.8	1	2	4	6	
Maximum number of missing genotypes	0	1	2	3	4	5	6	8	10	12	14	
Maximum number of heterozygous	0	1	2	3	4	5	6	8	10	12	14	

¹Default values in bold.

Table 3. Mean ± SE of the inbreeding coefficient, rate of inbreeding per generation (ΔF) ± SE, and effective population size (N_e ; 95% CI) for each evaluated subpopulation and breed

Item ¹	Breed or subpopulation ²				
	BB	MTN	MTR	LACCon	LACovi
Mean ± SE					
F_{PED}	0.0296 ± 0.0009	0.0298 ± 0.0010	0.0239 ± 0.0003	0.0234 ± 0.0002	0.0311 ± 0.0002
F_{SNP}	0.6294 ± 0.0005	0.6300 ± 0.0006	0.6226 ± 0.0002	0.6202 ± 0.0001	0.6230 ± 0.0001
F_{ROH}^3	0.0379 ± 0.0012	0.0440 ± 0.0013	0.0368 ± 0.0004	0.0346 ± 0.0003	0.0407 ± 0.0003
ΔF ± SE					
ΔF_{PED}	0.0099 ± 0.0017	0.0094 ± 0.0012	0.0045 ± 0.0004	0.0019 ± 0.0002	0.0022 ± 0.0002
ΔF_{SNP}	0.0044 ± 0.0010	0.0028 ± 0.0007	0.0025 ± 0.0002	0.0016 ± 0.0001	0.0014 ± 0.0001
ΔF_{ROH}^3	0.0085 ± 0.0024	0.0062 ± 0.0017	0.0046 ± 0.0006	0.0022 ± 0.0003	0.0014 ± 0.0003
N_e (95% CI)					
N_{ePED}	51 (38–76)	53 (43–71)	111 (95–135)	263 (218–332)	227 (193–277)
N_{eSNP}	114 (79–205)	179 (120–350)	200 (173–237)	313 (278–356)	357 (313–415)
N_{eROH}^3	59 (38–132)	81 (52–174)	109 (87–146)	227 (179–310)	357 (252–616)

¹ F_{PED} = inbreeding calculated from pedigree-based information; F_{SNP} = inbreeding calculated on a SNP-by-SNP basis; F_{ROH} = ROH-based inbreeding.

²BB = Basco-Béarnaise; MTN = Manech Tête Noire; MTR = Manech Tête Rousse; LACCon = Lacaune Confederation; LACovi = Lacaune Ovitest.

³ROH identifying values were: (1) the minimum length that constituted a ROH was 4 Mb; (2) the minimum number of SNP was 30; (3) the minimum density was 1 SNP per 100 kb; (4) the maximum distance allowed between 2 consecutive homozygous SNP in a run was 1 Mb; (5) a maximum of 2 missing genotypes; and (6) 1 heterozygous genotype within a particular ROH were permitted.

Table 3 shows the average inbreeding values for each evaluated subpopulation and breed. The ROH-based inbreeding estimates were slightly higher than those obtained from pedigree-based information. Inbreeding estimates obtained on a SNP-by-SNP basis (IBS) result in higher values than other estimators because they cannot distinguish between IBD and IBS. The difference is roughly a constant [e.g., for a single locus IBS = IBD + (1 - IBD)($p^2 + q^2$), where p and q are the allele frequencies] that vanishes when computing the rates of inbreeding (Toro et al., 2011). The lowest values of inbreeding were observed in LACCon subpopulation. In addition, the highest values of inbreeding were obtained for LACovi subpopulation (F_{PED}) and MTN breed (F_{SNP} and F_{ROH}).

The correlation coefficient between the inbreeding estimates is shown in Table 4. For all the evaluated

subpopulations and breeds the correlation coefficient between pedigree- and genomic-based inbreeding estimates is moderate (ranged between 0.38 and 0.58). The highest correlation was observed between both genomic inbreeding estimates (ranged between 0.78 and 0.88). Moreover, the lowest correlation was observed between F_{PED} and F_{ROH} , with the exception of MTR breed, where the lowest correlation was observed between F_{PED} and F_{SNP} . Differences are, however, very small.

The rate of inbreeding per generation is also shown in Table 3 for each evaluated subpopulation and breed. Similar estimates of rate of inbreeding were obtained both with pedigree- and genomic-based information, in contrast to the average inbreeding estimates. The highest values were observed for ΔF_{PED} (in BB, MTN, and LACovi) and ΔF_{ROH} (in MTR and LACCon). The lowest values of rate of inbreeding were obtained with

Table 4. Correlation coefficient of the inbreeding estimates for each evaluated subpopulation and breed¹

Breed ²	F_{PED} vs. F_{SNP}	F_{PED} vs. F_{ROH}^3	F_{SNP} vs. F_{ROH}^3
BB	0.53	0.51	0.86
MTN	0.50	0.48	0.88
MTR	0.56	0.58	0.84
LACCon	0.38	0.38	0.78
LACovi	0.42	0.40	0.79

¹ F_{PED} = inbreeding calculated from pedigree-based information; F_{SNP} = inbreeding calculated on a SNP-by-SNP basis; F_{ROH} = ROH-based inbreeding.

²BB = Basco-Béarnaise; MTN = Manech Tête Noire; MTR = Manech Tête Rousse; LACCon = Lacaune Confederation; LACovi = Lacaune Ovitest.

³ROH identifying values were: (1) the minimum length that constituted a ROH was 4 Mb; (2) the minimum number of SNP was 30; (3) the minimum density was 1 SNP per 100 kb; (4) the maximum distance allowed between 2 consecutive homozygous SNP in a run was 1 Mb; (5) a maximum of 2 missing genotypes; and (6) 1 heterozygous genotype within a particular ROH were permitted.

ΔF_{SNP} , except for LACovi, where the lowest values were also observed with ΔF_{ROH} . For Lacaune, values of ΔF are consistent with the 3 methods; this was not the case for BB, MTN, and MTR, where ΔF_{SNP} were consistently lowest.

Table 3 also shows estimates of N_e with confidence intervals. The lowest values of effective population size were obtained from ΔF_{PED} (in BB, MTN and LACovi) and ΔF_{ROH} (in MTR and LACCon). The highest values of effective population size were observed with ΔF_{SNP} , except for LACovi, where the highest values were also observed with ΔF_{ROH} . The 3 rates of inbreeding agree, ranking subpopulations and breeds by effective population size. In order of decreasing N_e , the breeds and subpopulations were LACovi, LACCon, MTR, MTN, and BB. The only exception was observed with ΔF_{PED} , which inverted LACCon and LACovi. In general, N_e estimates were above 200 in LACCon and LACovi subpopulations and below 200 in BB, MTN, and MTR breeds. However, no significant differences were observed in the estimates of N_e obtained from the 3 evaluated inbreeding estimates. This highlights the agreement between pedigree-based and SNP-by-SNP estimates of inbreeding with those obtained from ROH with the default values.

Effect on Estimates of Factors Defining ROH

In the evaluation of each ROH-defining factor, the other 5 ROH-defining factors remained the default value highlighted in bold in Table 2.

Figures 2 and 3 and Supplemental Figure S2 (<https://doi.org/10.3168/jds.2018-15405>) show the evolution of the rate of inbreeding in all evaluated breeds and subpopulations for each ROH-defining factor. The 6 evaluated factors showed, in general, the same behavior independently of the tested subpopulation or breed. Four ROH-defining factors have implications on the rate of inbreeding. These factors are (1) the minimum length that constituted a ROH, (2) the minimum number of SNP that constituted a ROH, (3) the minimum density (at least 1 SNP every some specific number of bases), and (4) the maximum distance allowed between 2 consecutive homozygous SNP. Figure 2 shows that, if the minimum length that constituted a ROH is equal or lower than 4 Mb (which could probably reflect ancient inbreeding), ΔF_{ROH} are, in general, in concordance with ΔF_{PED} . However, if the minimum length that constituted a ROH is higher than 4 Mb (which could probably reflect more recent inbreeding), ΔF_{ROH} are close to ΔF_{SNP} . In addition, reduced ROH lengths provided higher rates of inbreeding than greater ROH lengths. Accordingly, ΔF_{ROH} obtained with ROH of ≤ 80 SNP are consistent with ΔF_{PED} , and ΔF_{ROH}

estimated from ROH of more than 80 SNP are, overall, similar to ΔF_{SNP} . The ΔF_{ROH} assessed with a minimum density equal or lower than 1 SNP every 90 kb is close to ΔF_{SNP} , and the ΔF_{ROH} estimated with a minimum density higher than 1 SNP every 90 kb is close to ΔF_{PED} . Finally, a maximum distance between 2 homozygous SNP higher than 0.3 Mb makes ΔF_{ROH} agree with ΔF_{PED} , and equal or lower to 0.3 Mb makes ΔF_{ROH} match ΔF_{SNP} (Figure 3).

Rates of inbreeding higher than those obtained with ROH default values have implications on the estimation of diversity because a ΔF of 0.1% means that 0.1% of heterozygosity is lost, and the actual constraints on management of inbreeding (e.g., using optimal contributions) are typically based on ΔF (Meuwissen, 1997; Sonesson et al., 2012). Thus, different factors for ROH will lead to dissimilar ΔF_{ROH} , which will lead to diverse selection decisions with more or less emphasis on genetic diversity.

Two ROH-defining factors generally have no consequences on the estimation of the rate of inbreeding across the evaluated subpopulations and breeds. These 2 factors are (1) the maximum number of missing genotypes allowed and (2) the maximum number of heterozygous permitted (Supplemental Figure S2; <https://doi.org/10.3168/jds.2018-15405>).

The ranking of subpopulations and breeds regarding the rate of inbreeding is the same when evaluating the 6 ROH-defining factors, where BB has the highest ΔF followed by MTN, MTR, LACCon, and, finally, LACovi. This ranking is only disrupted when the minimum length that constituted a ROH is higher than 10 Mb and when the minimum number of SNP is higher than 300, probably due to the difficulty in finding these types of ROH; consequently, ROH-based inbreeding might be close to 0.

DISCUSSION

Inbreeding happens when an individual inherits chromosomal fragments that are IBD from both parents, leading to faster allele fixation, reduction of additive genetic variance (in principle), inbreeding depression, and a decline in the response to selection (Kristensen and Sorensen, 2005). Traditionally, inbreeding coefficients have been estimated using known pedigrees (Wright, 1922). Inspired by the pedigree-based inbreeding coefficients, Caballero and Toro (2002) proposed to estimate inbreeding on a SNP-by-SNP basis, where each SNP is considered individually. An alternative way to estimate inbreeding coefficients is to determine the IBD probability of chromosome segments (ROH) instead of each SNP individually (Howard et al., 2017).

Inbreeding coefficients evaluated in the present study are defined in relation to the base population. In F_{PED} , it is relative to the pedigree depth; for F_{ROH} , it is dependent on the ROH length; finally, in F_{SNP} , it is relative to Hardy-Weinberg equilibrium. Thus, some differences arise from the different definitions. The F_{PED} estimates were lower than F_{SNP} and F_{ROH} , suggesting that F_{PED} may possibly be underestimating inbreeding. This can be explained considering that F_{PED} is an expectation and variability exists around this expectation (Keller et al., 2011). More specifically, at one locus, the true inbreeding based on IBD has an expected value of F_{PED} and a variance of $F_{PED}(1 - F_{PED})$ (García-Cortés et al.,

2013). In addition, when calculating F_{PED} , it is assumed that founder animals are unrelated. This assumption may also lead to the underestimation of F_{PED} if the recorded pedigree is not deep enough (depth of pedigree is very heterogeneous in sheep breeds, as shown in Table 1), is incomplete, or has errors due to misidentification (8% in these sheep breeds; Tortereau et al., 2017) and incorrect recording. In addition, F_{PED} assumes that the loci are neutral, so it does not consider loci submitted to selection. As pedigree information has been used extensively to measure the proportion of the genome that is IBD, the relatively moderate correlation between F_{PED} and F_{ROH} supports that F_{ROH} can be used to predict

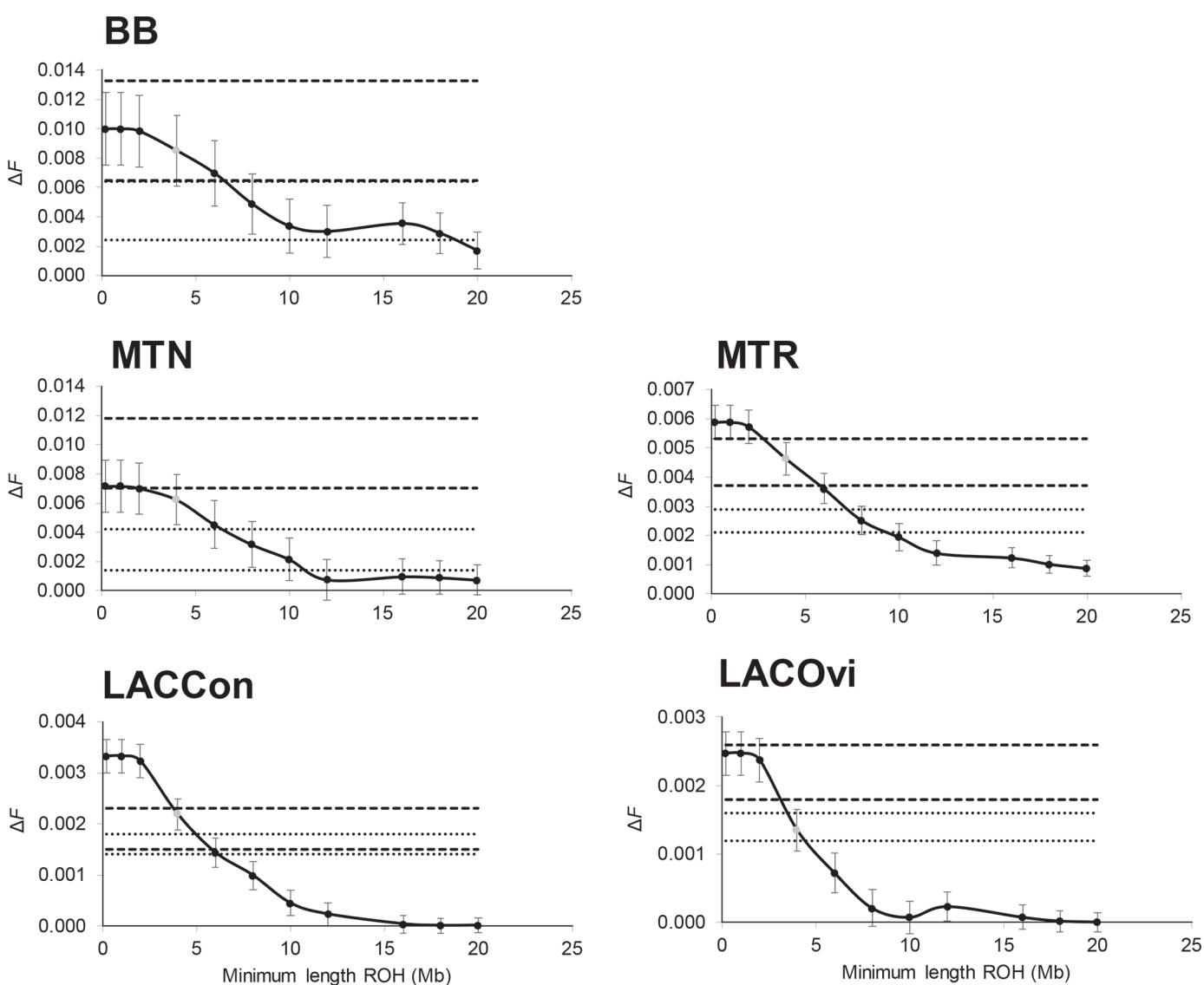


Figure 2. Evolution of the rate of inbreeding per generation (ΔF) in relation to the minimum length that constituted a run of homozygosity (ROH; Mb). BB = Basco-Béarnaise; MTN = Manech Tête Noire; MTR = Manech Tête Rousse; LACCon = Lacaune Confederation; LACovi = Lacaune Ovitest. Gray point indicates the default value. Bars indicate the SE. Dashed lines indicate the 95% CI of pedigree-based inbreeding. Dotted lines indicate the 95% CI of SNP-by-SNP-based inbreeding.

IBD in sheep populations (Purfield et al., 2017). Accordingly, similar correlations between F_{PED} and F_{ROH} were previously reported in cattle populations (Zhang et al., 2015). However, this correlation depends on the minimum length that constitute a ROH. For example, Silió et al. (2013) indicated a correlation between F_{PED}

and F_{ROH} longer than 1 Mb and F_{ROH} longer than 5 Mb of 0.78 and 0.82, respectively.

The absolute value of the inbreeding coefficient is not a valuable alternative to consider whether there is too much inbreeding or not because it is affected by several factors, such as (1) the number of generations consid-

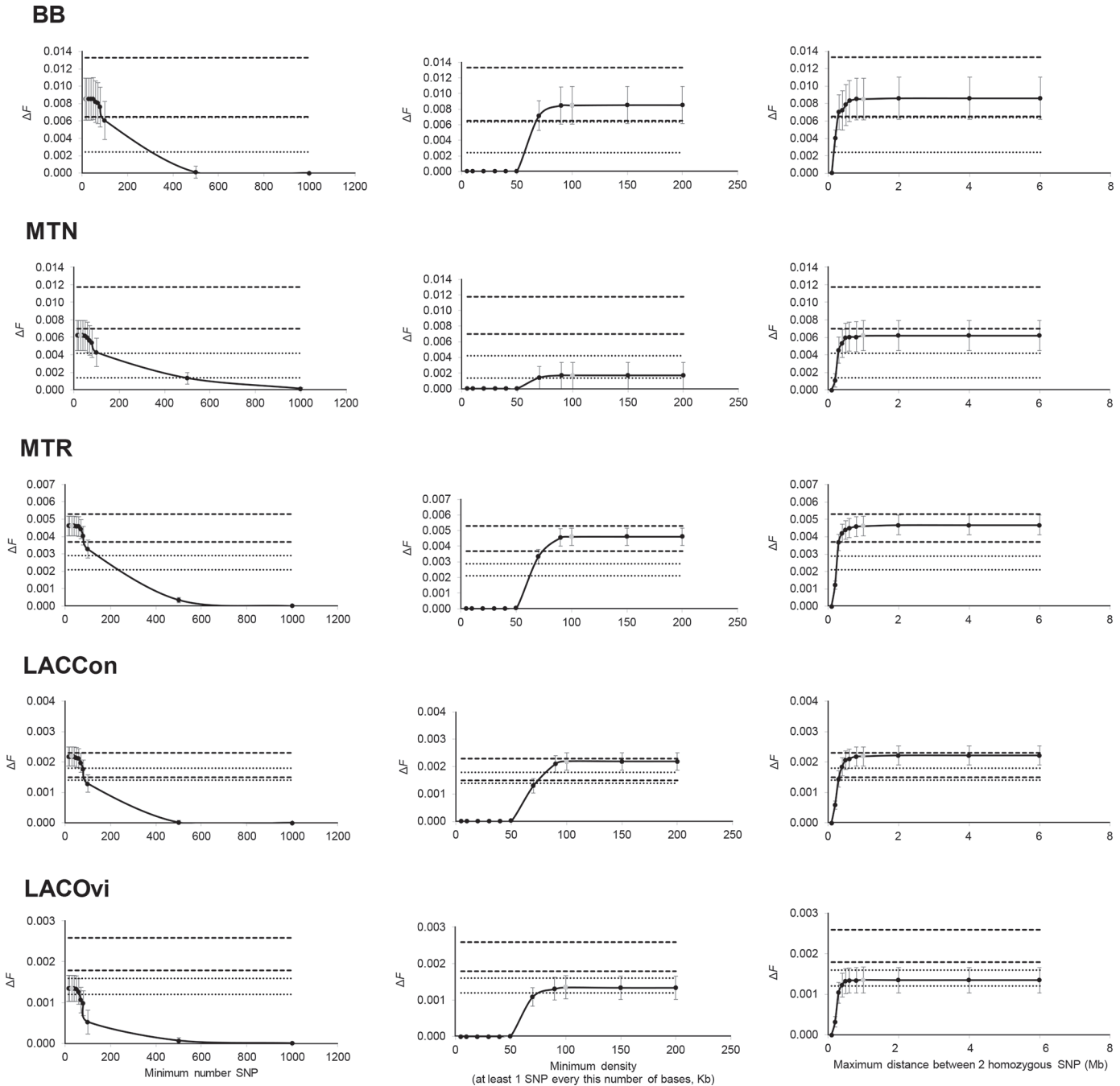


Figure 3. Evolution of the rate of inbreeding per generation (ΔF) in relation to the minimum number of SNP, the minimum density, and the maximum distance between 2 homozygous SNP. BB = Basco-Béarnaise; MTN = Manech Tête Noire; MTR = Manech Tête Rousse; LACCon = Lacaune Confederation; LACOvi = Lacaune Ovitest. Gray point indicates the default value. Bars indicate the SE. Dashed lines indicate the 95% CI of pedigree-based inbreeding. Dotted lines indicate the 95% CI of SNP-by-SNP-based inbreeding.

ered, (2) if markers are used, (3) the allele frequency spectrum of those markers, or (4) factors defining ROH. The rate of inbreeding is a better criterion to handle inbreeding management, as it translates to the effective size of a population (Meuwissen et al., 2018).

Sheep have kept a considerable level of genetic diversity. The estimates of effective population size shown in our study are similar to those previously reported in the same breeds (Larroque et al., 2014) and in other comparable sheep breeds (Li et al., 2009; García-Gómez et al., 2012; Kijas et al., 2012; Al-Mamun et al., 2015; Beynon et al., 2015; Chitneedi et al., 2017; Mastrangelo et al., 2017a,b; Purfield et al., 2017). An effective population size smaller than 100 could compromise the long-term viability of the population (Meuwissen, 2009). However, although estimated N_e are widely used, its estimation is remarkably complex (Wang, 2005) and assumptions are generally not accomplished (Hayes et al., 2003).

Runs of homozygosity have been studied in humans (e.g., Broman and Weber, 1999; Gibson et al., 2006; Lencz et al., 2007; Curtis et al., 2008; McQuillan et al., 2008), cattle (e.g., Ferenčaković et al., 2011; Purfield et al., 2012; Bjelland et al., 2013; Mastrangelo et al., 2016), and pigs (e.g., Bosse et al., 2012; Herrero-Medrano et al., 2013; Silió et al., 2013; Zhang et al., 2014), but less commonly in other livestock species, such as sheep (Al-Mamun et al., 2015; Beynon et al., 2015; Muchadeyi et al., 2015). Al-Mamun et al. (2015) indicated the Border Leicester breed has the largest total number of ROH, followed by Poll Dorset and, finally, Merino; those authors also identified ROH within autosomes. Beynon et al. (2015) used the haplotype homozygosity method to infer population history and structure of Welsh sheep breeds. This approach is based on the genome-wide distribution of ROH. Finally, Muchadeyi et al. (2015) related the ROH regions with subvital performance in Swakara breed. In addition, relatively few studies have assessed which set of parameters is optimal for identifying ROH to better understand their effects on detecting inbreeding. Accordingly, F_{ROH} does not accumulate over generations, as pedigree-based inbreeding estimates, because ROH break down due to recombination (Meuwissen et al., 2018). Finally, the probability of finding long ROH could be higher in regions surrounding the centromere, where the recombination rate is supposed to be reduced. Thus, efforts should be made to obtain correct ROH-based inbreeding estimates when defining ROH. It has been shown that 4 defining factors influence the identification of ROH-based inbreeding. The minimum length that should constitute a ROH has been previously studied by Ferenčaković et al. (2013), who considered a minimum length of 4 Mb to

obtain reliable estimates of inbreeding. In addition, ROH-based rates of inbreeding were higher with short ROH than with long ROH, in agreement with results obtained from pedigree-based information in the same breeds (Buisson et al., 2018). The minimum number of SNP that constituted a ROH is also an important ROH-defining factor. Accordingly, Forutan et al. (2018) indicated that a minimum ROH length of 20 to 50 SNP gave the most accurate and efficient estimates of ROH-based inbreeding estimates. The other 2 important ROH-defining factors identified in our study were the maximum distance between 2 homozygous SNP and the minimum density between 2 SNP. The ROH-based rates of inbreeding in concordance to those obtained from pedigree-based information require a specific set of values in the other ROH-defining factors. This particular set of values is different to that identified to obtain ROH-based rates of inbreeding similar to those obtained on a SNP-by-SNP basis.

Detecting ROH based on 50K chip data gave estimates similar to ROH from sequence data (Zhang et al., 2015). Consequently, in the absence of full sequence data, ROH based on 50K can be used to assess inbreeding. Notwithstanding, genotypes denser than 50K could be required to accurately detect short ROH.

In conclusion, our results show that rates of inbreeding and effective population sizes are empirically comparable across methods (pedigree, SNP-by-SNP, and ROH). Effective population sizes of these dairy sheep breeds are in the low hundreds. In addition, factors to define ROH do not change much unless extreme values are considered, although further research on ROH-based inbreeding is still required.

ACKNOWLEDGMENTS

We thank Carole Moreno (INRA, Castanet Tolosan, France) and 2 anonymous referees for useful suggestions on the manuscript. This project has received funding from the European Unions' Horizon 2020 Research & Innovation programme under grant agreement no. 772787 – SMARTER (Small Ruminants Breeding for Efficiency and Resilience, Castanet Tolosan, France). This work was also supported by ARDI (Research, Development and Innovation, Vitoria, Spain; grant agreement EFA 208/16) from POCTEFA funds, and GDivSelGen (Efficient Use of Genetic Diversity in Genomic Selection, Paris, France) action (INRA SelGen metaprogram). Authors are also grateful to the GenoToul bioinformatics platform Toulouse Midi-Pyrenees (Castanet Tolosan, France) for providing computing and storage resources.

REFERENCES

- Al-Mamun, H. A., S. A. Clark, P. Kwan, and C. Gondro. 2015. Genome-wide linkage disequilibrium and genetic diversity in five populations of Australian domestic sheep. *Genet. Sel. Evol.* 47:90.
- Albrechtsen, A., F. C. Nielsen, and R. Nielsen. 2010. Ascertainment biases in SNP chips affect measures of population divergence. *Mol. Biol. Evol.* 27:2534–2547.
- Baloche, G., A. Legarra, G. Sallé, H. Larroque, J. M. Astruc, C. Robert-Granié, and F. Barillet. 2014. Assessment of accuracy of genomic prediction for French Lacaune dairy sheep. *J. Dairy Sci.* 97:1107–1116.
- Beynon, S. E., G. T. Slavov, M. Farre, B. Sunduimijid, K. Waddams, B. Davies, W. Haresign, J. Kijas, I. M. MacLeod, C. J. Newbold, L. Davies, and D. M. Larkin. 2015. Population structure and history of the Welsh sheep breeds determined by whole genome genotyping. *BMC Genet.* 16:65.
- Bjelland, D. W., K. A. Weigel, N. Vukasinovic, and J. D. Nkrumah. 2013. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *J. Dairy Sci.* 96:4697–4706.
- Boichard, D. 2002. PEDIG: A Fortran package for pedigree analysis suited for large populations. In *Proceedings of the 7th World Congress of Genetics Applied to Livestock Production*, France. Am. Soc. Anim. Sci., Champaign, IL.
- Bosse, M., H.-J. Megens, O. Madsen, Y. Paudel, L. A. Frantz, L. B. Schook, R. P. Crooijmans, and M. A. Groenen. 2012. Regions of homozygosity in the porcine genome: consequence of demography and the recombination landscape. *PLoS Genet.* 8:e1003100.
- Broman, K. W., and J. L. Weber. 1999. Long homozygous chromosomal segments in reference families from the Centre d'Etude du Polymorphisme Humain. *Am. J. Hum. Genet.* 65:1493–1500.
- Buisson, D., J. M. Astruc, and F. Barillet. 2018. Bilan et perspectives de la gestion de la variabilité génétique des ovins laitiers en France. *INRA Prod. Anim.* 31:1–12.
- Caballero, A., and M. A. Toro. 2002. Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Genet.* 3:289–299.
- Chitneedi, P. K., J. J. Arranz, A. Suarez-Vega, E. García-Gámez, and B. Gutiérrez-Gil. 2017. Estimations of linkage disequilibrium, effective population size and ROH-based inbreeding coefficients in Spanish Churra sheep using imputed high-density SNP genotypes. *Anim. Genet.* 48:436–446.
- Curik, I., M. Ferenčaković, and J. Sölkner. 2014. Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livest. Sci.* 166:26–34.
- Curik, I., J. Sölkner, and N. Stipic. 2002. Effects of models with finite loci, selection, dominance, epistasis and linkage on inbreeding coefficients based on pedigree and genotypic information. *J. Anim. Breed. Genet.* 119:101–115.
- Curtis, D., A. E. Vine, and J. Knight. 2008. Study of regions of extended homozygosity provides a powerful method to explore haplotype structure of human populations. *Ann. Hum. Genet.* 72:261–278.
- Daetwyler, H. D., B. Villanueva, P. Bijma, and J. A. Woolliams. 2007. Inbreeding in genome-wide selection. *J. Anim. Breed. Genet.* 124:369–376.
- Falconer, D. S., and T. F. C. Mackay. 1996. *Introduction to Quantitative Genetics*. Longman Scientific & Technical, Harlow, United Kingdom.
- Ferenčaković, M., E. Hamzic, B. Gredler, I. Curik, and J. Sölkner. 2011. Runs of homozygosity reveal genome-wide autozygosity in the Austrian Fleckvieh cattle. *ACS Agric. Conspec. Sci.* 76:325–328.
- Ferenčaković, M., J. Sölkner, and I. Curik. 2013. Estimating autozygosity from high-throughput information: Effects of SNP density and genotyping errors. *Genet. Sel. Evol.* 45:42.
- Forutan, M., S. A. Mahyari, C. Baes, N. Melzer, F. S. Schenkel, and M. Sargolzaei. 2018. Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. *BMC Genomics* 19:98.
- García-Cortés, L. A., A. Legarra, C. Chevalet, and M. A. Toro. 2013. Variance and covariance of actual relationships between relatives at one locus. *PLoS One* 8:e57003.
- García-Gámez, E., G. Sahana, B. Gutiérrez-Gil, and J. J. Arranz. 2012. Linkage disequilibrium and inbreeding estimation in Spanish Churra sheep. *BMC Genet.* 13:43.
- Gibson, J., E. M. Newton, and A. Collins. 2006. Extended tracts of homozygosity in outbred human populations. *Hum. Mol. Genet.* 15:789–795.
- Hayes, B. J., P. M. Visscher, H. C. McPartlan, and M. E. Goddard. 2003. Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res.* 13:635–643.
- Herrero-Medrano, J. M., H.-J. Megens, M. A. M. Groenen, G. Ramis, M. Bosse, M. Pérez-Enciso, and R. P. M. A. Crooijmans. 2013. Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. *BMC Genet.* 14:106.
- Hill, W. G., and B. S. Weir. 2011. Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet. Res. (Camb.)* 93:47–64.
- Howard, J. T., J. E. Pryce, C. Baes, and C. Maltecca. 2017. Invited review: Inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of genomic variability. *J. Dairy Sci.* 100:6009–6024.
- Howrigan, D. P., M. A. Simonson, and M. C. Keller. 2011. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics* 12:460.
- Keller, M. C., P. M. Visscher, and M. E. Goddard. 2011. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics* 189:237–249.
- Kijas, J. W., J. A. Lenstra, B. Hayes, S. Boitard, L. R. Porto Neto, M. San Cristobal, B. Servin, R. McCulloch, V. Whan, K. Gietzen, S. Paiva, W. Barendse, E. Ciani, H. Raadsma, J. McEwan, S. Dalrymple, and International Sheep Genomics Consortium. 2012. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol.* 10:e1001258.
- Kristensen, T. N., and A. C. Sorensen. 2005. Inbreeding-Lessons from animal breeding, evolutionary biology and conservation genetics. *Anim. Sci.* 80:121–133.
- Ku, C. S., N. Naidoo, S. M. Teo, and Y. Pawitan. 2011. Regions of homozygosity and their impact on complex diseases and traits. *Hum. Genet.* 129:1–15.
- Larroque, H., F. Barillet, G. Baloche, J. M. Astruc, D. Buisson, F. Shumbusho, V. Clément, G. Lagriffoul, I. Palhiere, R. Rupp, C. Carillier, C. Robert-Granié, and A. Legarra. 2014. Toward genomic breeding programs in French dairy sheep and goats. Page 332 in *Proceedings of the 10th World Congress of Genetics Applied to Livestock Production*, Canada. Am. Soc. Anim. Sci., Champaign, IL.
- Legarra, A., G. Baloche, F. Barillet, J. M. Astruc, C. Soulas, X. Aguerre, F. Arrese, L. Mintegi, M. Lasarte, F. Maetzu, I. Beltrán de Heredia, and E. Ugarte. 2014. Within- and across-breed genomic predictions and genomic relationships for Western Pyrenees dairy sheep breeds Latxa, Manech, and Basco-Béarnaise. *J. Dairy Sci.* 97:3200–3212.
- Lencz, T., C. Lambert, P. DeRosse, K. E. Burdick, T. V. Morgan, J. M. Kane, R. Kucherlapati, and A. K. Malhotra. 2007. Runs of homozygosity reveal highly penetrant recessive loci in schizophrenia. *Proc. Natl. Acad. Sci. USA* 104:19942–19947.
- Li, M. H., I. Strandén, and J. Kantanen. 2009. Genetic diversity and pedigree analysis of the Finnsheep breed. *J. Anim. Sci.* 87:1598–1605.
- Malécot, G. 1948. *Les Mathématiques de l'Hérédité*. Masson & Cie, Paris, France.
- Mastrangelo, S., B. Portolano, R. Di Gerlando, R. Ciampolini, M. Tolone, M. T. Sardina, and International Sheep Genomics Consortium. 2017a. Genome-wide analysis in endangered populations: A case study in Barbaresca sheep. *Animal* 11:1107–1116.
- Mastrangelo, S., M. Tolone, R. D. Gerlando, L. Fontanesi, M. T. Sardina, and B. Portolano. 2016. Genomic inbreeding estimation in

- small populations: Evaluation of runs of homozygosity in three local dairy cattle breeds. *Animal* 10:746–754.
- Mastrangelo, S., M. Tolone, M. T. Sardina, G. Sottile, A. M. Sutura, R. Di Gerlando, and B. Portolano. 2017b. Genome-wide scan for runs of homozygosity identifies potential candidate genes associated with local adaptation in Valle del Belice sheep. *Genet. Sel. Evol.* 49:84.
- McQuillan, R., A. Leutenegger, R. Abdel-Rahman, C. S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A. K. Macleod, S. M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S. H. Wild, M. G. Dunlop, A. F. Wright, H. Campbell, and J. F. Wilson. 2008. Runs of homozygosity in European populations. *Am. J. Hum. Genet.* 83:359–372.
- Meuwissen, T. H. E. 1997. Maximizing the response of selection with a predefined rate of inbreeding. *J. Anim. Sci.* 75:934–940.
- Meuwissen, T. H. E. 2009. Genetic management of small populations: A review. *Acta Agric. Scand. A Anim. Sci.* 59:71–79.
- Meuwissen, T. H. E., and Z. Luo. 1992. Computing inbreeding coefficients in large populations. *Genet. Sel. Evol.* 24:305–313.
- Meuwissen, T. H. E., A. K. Sonesson, and J. A. Woolliams. 2018. Genomic management of inbreeding in breeding schemes. In *Proceedings of the 11th World Congress of Genetics Applied to Livestock Production*. ICAR and Interbull, New Zealand. Am. Soc., Anim. Sci., Champaign, IL.
- Muchadeyi, F. C., M. T. Malesa, P. Soma, and E. F. Dzomba. 2015. Runs of homozygosity in Swakara pelt producing sheep: Implications on sub-vital performance. Pages 310–313 in *Proceedings of the AAABG. Association for the advancement of Animal Breeding and Genetics, Australia and New Zealand*.
- Oliehoek, P. A., and P. Bijma. 2009. Effects of pedigree errors on the efficiency of conservation decisions. *Genet. Sel. Evol.* 41:9.
- Peripolli, E., D. P. Munari, M. V. G. B. Silva, A. L. F. Lima, R. Irgang, and F. Baldi. 2017. Runs of homozygosity: Current knowledge and applications in livestock. *Anim. Genet.* 48:255–271.
- Purfield, D. C., D. Berry, S. McParland, and D. G. Bradley. 2012. Runs of homozygosity and population history in cattle. *BMC Genet.* 13:70.
- Purfield, D. C., S. McParland, E. Wall, and D. P. Berry. 2017. The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. *PLoS One* 12:e0176780.
- Silió, L., M. C. Rodríguez, A. Fernández, C. Barragán, R. Benítez, C. Óvilo, and A. I. Fernández. 2013. Measuring inbreeding and inbreeding depression on pig growth from pedigree or SNP-derived metrics. *J. Anim. Breed. Genet.* 130:349–360.
- Smith, J. M., and J. Haigh. 1974. The hitch-hiking effect of a favourable gene. *Genet. Res.* 23:23–35.
- Sonesson, A. K., J. A. Woolliams, and T. H. E. Meuwissen. 2012. Genomic selection requires genomic control of inbreeding. *Genet. Sel. Evol.* 44:27.
- Toro, M. A., L. A. García-Cortés, and A. Legarra. 2011. A note on the rationale for estimating genealogical coancestry from molecular markers. *Genet. Sel. Evol.* 43:1–10.
- Tortereau, F., C. R. Moreno, G. Tosser-Klopp, B. Servin, and J. Raoul. 2017. Development of a SNP panel dedicated to parentage assignment in French sheep populations. *BMC Genet.* 18:50.
- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423.
- Wang, J. 2005. Estimation of effective population sizes from data on genetic markers. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360:1395–1409.
- Wright, S. 1922. Coefficients of inbreeding and relationship. *Am. Nat.* 56:330–338.
- Zhang, Q., B. Guldbbrandtsen, M. Bosse, X. Sun, A. Wolc, and J. C. M. Dekkers. 2015. Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genomics* 16:542.
- Zhang, Y., J. M. Young, C. Wang, X. Sun, A. Wolc, and J. C. M. Dekkers. 2014. Inbreeding by pedigree and genomic markers in selection lines of pigs. Page 484 in *Proceedings of the 10th World Congress of Genetics Applied to Livestock Production, Canada*. Am. Soc. Anim. Sci., Champaign, IL.