the **genetics**society

ARTICLE

An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree

Setegn Worku Alemu [1] · Naveen Kumar Kadri[1] · Chad Harland[1] · Pierre Faux [1] · Carole Charlier [1] · Armando Caballero [2] · Tom Druet [1]

Abstract

The estimation of the inbreeding coefficient (F) is essential for the study of inbreeding depression (ID) or for the management of populations under conservation. Several methods have been proposed to estimate the realized F using genetic markers, but it remains unclear which one should be used. Here we used whole-genome sequence data for 245 individuals from a Holstein cattle pedigree to empirically evaluate which estimators best capture homozygosity at variants causing ID, such as rare deleterious alleles or loci presenting heterozygote advantage and segregating at intermediate frequency. Estimators relying on the correlation between uniting gametes (F_{UNI}) or on the genomic relationships (F_{GRM}) presented the highest correlations with these variants. However, homozygosity at rare alleles remained poorly captured. A second group of estimators relying on excess homozygosity (F_{HOM}), homozygous-by-descent segments (F_{HBD}), runs-of-homozygosity (F_{ROH}) or on the known genealogy (F_{PED}) was better at capturing whole-genome homozygosity, reflecting the consequences of inbreeding on all variants, and for young alleles with low to moderate frequencies ($0.10 < . < 0.25$). The results indicate that F_{UNI} and F_{GRM} might present a stronger association with ID. However, the situation might be different when recessive deleterious alleles reach higher frequencies, such as in populations with a small effective population size. For locus-specific inbreeding measures or at low marker density, the ranking of the methods can also change as F_{HBD} makes better use of the information from neighboring markers. Finally, we confirmed that genomic measures are in general superior to pedigree-based estimates. In particular, F_{PED} was uncorrelated with locus-specific homozygosity.

Introduction

Inbreeding results from the mating of related individuals and is associated with negative consequences such as inbreeding depression (ID), the reduction in fitness due to

✉ Tom Druet
tom.druet@uliege.be

1 Unit of Animal Genomics, GIGA-R & Faculty of Veterinary Medicine, University of Liège, Liège, Belgium

2 Centro de Investigación Mariña, Departamento de Bioquímica, Genética e Inmunología, Edificio CC Experimentais, Universidade de Vigo, Campus de Vigo, As Lagoas, Marcosende, 36310 Vigo, Spain

increased homozygosity (Hedrick and Garcia-Dorado 2016). Inbreeding depression is common in livestock species (Leroy 2014) and many recessive disorders associated with increased inbreeding have been identified in intensively selected cattle breeds. There are two main hypotheses to explain ID (Charlesworth and Willis 2009). The first is increased homozygosity at partially recessive deleterious alleles (e.g., Charlesworth and Charlesworth 1999) and the second is reduced heterozygosity (equivalent to increased homozygosity) at loci presenting heterozygote advantage (overdominance). Although Charlesworth (2015) concluded that the second hypothesis might be important in Drosophila, most empirical evidences suggest that the first one might be more important for other species (Charlesworth and Willis 2009; Hedrick 2012). The two proposed mechanisms have distinct consequences on allele frequencies. Purging selection maintains deleterious alleles at low frequency or removes them from the population, so that such alleles are mostly young and segregate at low frequency. In contrast, overdominant alleles are maintained at

intermediate frequency due to balancing selection (Charlesworth and Willis 2009).

The inbreeding coefficient (F) is a tool for the management of populations and for the study of ID. It also provides information on relatedness among parents, mating systems, population structure and recent demographic events (see, e.g., Caballero 2020, Chap 4). Wright (1922) defined the coefficient of inbreeding in terms of correlations between the parents' uniting gametes. Malécot (1948) offered an alternative definition based on the probability that two homologous alleles in an individual are identical-by-descent (IBD), i.e., they are copies of an allele from a common ancestor. Pedigree-based estimators of the inbreeding coefficient rely on this definition and require the choice of a reference population (e.g., an earlier generation), often determined by the available genealogy. The choice of this reference generation is somewhat arbitrary as individuals from that generation must be considered unrelated. Genomic inbreeding measures can also be estimated using genetic markers. Several authors concluded that genomic measures are better than pedigree-based estimates (e.g., Keller et al. 2011; Wang 2016), mainly because they provide estimators of the realized inbreeding, are robust to pedigree errors and do not require a genealogy.

Numerous genomic estimators of the inbreeding coefficient have been proposed, and there is no consensus on which is the most appropriate (Goudet et al. 2018). Inbreeding measures can for instance be estimated by maximum likelihood approaches (Milligan 2003; Wang 2007), by methods-of-moment (Ritland 1996; Purcell et al. 2007), from the diagonal elements of a genomic relationship matrix (GRM) (VanRaden 2008), from simple heterozygosity or homozygosity measures (Szulkin et al. 2010; Bjelland et al. 2013), based on genotypic correlations (Ackerman et al. 2017) or from the proportion of the genome within runs-of-homozygosity (ROH) (McQuillan et al. 2008; Ferenčaković et al. 2013). These different estimators of the inbreeding coefficient can be evaluated using different approaches. First, their theoretical properties can be derived, as Yengo et al. (2017) did for bias and standard errors, although this was not possible for all estimators. Alternatively, empirical comparisons between estimators can be performed on real datasets, that have been genotyped at low to moderate density with neutral markers selected on their minor allele frequency (MAF), and without knowledge of the true inbreeding coefficient or relatedness (Santure et al. 2010; Bjelland et al. 2013; Pryce et al. 2014; Zhang et al. 2015a; Goudet et al. 2018; Kardos et al. 2018a). In such cases, the coefficients are often compared to the pedigree-based estimates, although the latter should not be considered as the golden standard (Speed and Balding 2015). Performances of estimators can also be evaluated empirically on real data, by testing which measure presents

the highest correlation with fitness traits (Kardos et al. 2016). As such, several authors have used statistical criteria to determine which inbreeding coefficients best fit recorded phenotypes (Grueber et al. 2011; Ferenčaković et al. 2017; Clark et al. 2019). Simulation studies can also be carried out (e.g., Milligan 2003; Keller et al. 2011; Druet and Gautier 2017; Nietlisbach et al. 2019) but they rely on an arbitrary definition of the true inbreeding coefficient and other assumptions that are sometimes unrealistic. For instance, they might assume unlinked loci, absence of selection, random mating, equal parent contributions, non-overlapping generations, homogeneous recombination rates, or a simplified population history, such as a constant effective population size (N_e) or a single bottleneck. Previous comparisons concluded that the best method varied according to parameters such as the number of markers, the number of alleles, the number of individuals, the relatedness within the population, the mating structure or the intended application (e.g., Milligan 2003; Wang 2011; Goudet et al. 2018).

Whole-genome sequence data might provide a good opportunity to empirically evaluate different inbreeding estimators, since genome-wide heterozygosity can be measured extremely accurately (Kardos et al. 2016). With resequencing data, genotypes are available for almost all variants, including those contributing to ID. In addition, genotypes for markers segregating at all frequencies and for alleles from different functional categories, including deleterious variants or those under balancing selection, are present in the data. Consequently, they offer a complementary empirical strategy to evaluate the properties and accuracies of different inbreeding measures, for applications such as the study of ID or the evaluation of conservation and selection programs. For instance, Yengo et al. (2017) used true genotypes available for more than nine million SNPs to simulate ID and compare inbreeding measures in humans. In such an approach, allele frequencies, linkage disequilibrium (LD) patterns and heterogeneity along the genome matched reality. Some assumptions, however, were still required regarding the architecture of ID, such as the class of variants causing it, or the relationship considered between effect size and allele frequency. Thus, the way in which the phenotypes were simulated and how metrics were evaluated had been subject of debate (Kardos et al. 2018b; Yengo et al. 2018; Nietlisbach et al. 2019). We herein propose to follow a similar empirical strategy to evaluate different inbreeding measures in cattle, a livestock species with a very different demographic history compared to human populations. To that end, we used whole-genome sequence data available for a Dutch Holstein cattle pedigree. Inbreeding coefficients were estimated from subsets of markers and the resequencing data was used to estimate homozygosity at different groups of markers. The latter homozygosity measures could for instance serve as proxies

for homozygosity at alleles causing ID. Furthermore, performance evaluations were also realized for inbreeding coefficients estimated for specific positions in the genome. Such locus-specific measures would be useful to identify regions contributing to inbreeding depression (e.g., Pryce et al. 2014) or to manage inbreeding at specific loci.

Material and methods

Data

We used whole-genome sequences (WGS) from 245 Dutch Holstein cattle sequenced with a coverage higher than 15x. These corresponded to a set of 145 parents and their 100 sequenced offspring. The animals are part of a pedigree containing 743 individuals and sequenced at variable coverage (Harland et al. 2017). The data processing to generate the final Variant Call Format (VCF) file is described in Kadri et al. (2016). This includes description of DNA extraction, library preparation, reads alignment to the reference genome (Bos Taurus UMD 3.1), base quality calibration, variant calling and variant quality score recalibration. For this recalibration of variant quality, we used a set of trusted SNPs from the BovineHD (Illumina) and Axiom Genome-Wide BOS 1 (Affymetrix) commercial genotyping arrays as reference training set.

We selected 12,735,685 autosomal bi-allelic SNPs based on the variant quality score recalibration procedure from GATK (DePristo et al. 2011). The selected SNPs had a variant quality score above the threshold defined to conserve 97.5% from the variants in the reference training set. The inbreeding coefficients were estimated with a subset of 37,675 SNPs present on the Illumina BovineSNP50 Bead-Chip. Markers located in putative map errors as defined by Kadri et al. (2016) were excluded.

We extracted a genealogy including all the 743 sequenced individuals and their ancestors. The generated pedigree file contained 12,238 individuals, and the 743 sequenced individuals had on average 99.9, 97.2 and 84.0% known ancestors in their 5th, 8th and 10th pedigree generation, respectively.

Estimation of inbreeding coefficients

The levels of genomic inbreeding were estimated with several measures and using the set of 37,675 SNPs from the commercial array (248 monomorphic SNPs were additionally filtered out for estimation of genome-wide inbreeding coefficients).

A set of estimators of the inbreeding coefficient were obtained from individual SNP data. The first measure (F_{UNI}) was based on the correlation between uniting gametes

(Yang et al. 2011) and is equivalent to the method proposed by Li and Horvitz (1953) or Ritland (1996). The second measure (F_{GRM}) was obtained from the diagonal elements of the genomic relationship matrix (GRM) computed using the first method proposed by VanRaden (2008). These two measures were estimated with GCTA (Yang et al. 2011). The third measure was the excess of homozygosity (F_{HOM}), a moment estimator based on the expected and observed individual heterozygosity, implemented in PLINK (Purcell et al. 2007) and proposed by Li and Horvitz (1953). The fourth measure (F_{ML}) was the maximum likelihood estimator from Wang (2007) using genotypes of triad of individuals to estimate the nine condensed IBD states (Jacquard 1974). The method is implemented in COANCESTRY (Wang 2011) and the related R package (Pew et al. 2015).

A second set of estimators of F was based on sequences of consecutive homozygous SNPs. Runs of Homozygosity (ROH) were detected with PLINK with the following options: a minimum of 50 SNPs per ROH, at least 1 SNP per 100 Kb, a scanning window of 50 SNPs, a total length > 2 Mb, spacing between successive SNPs <500 Kb and no heterozygous SNPs. These ROH were then used to calculate F_{ROH}, defined as the proportion of the genome in ROH (McQuillan et al. 2008). A distinction between different ROH length classes (2–5 Mb, 5–10 and >10) were considered, as described in more detail in Supplementary Text S1. Estimators were also obtained from the proportion of the genome in homozygous-by-descent (HBD) segments (Druet and Gautier 2017), closely related to F_{ROH}. A comparison of these two last approaches is available in Solé et al. (2017). A hidden Markov model with four HBD classes with rates equal to 5, 25, 125 and 525 was run with RZooROH (Bertrand et al. 2019). These rates are associated to the length of HBD segments in each HBD class: the expected length of HBD segments being equal to $1/R_k$ Morgans, where R_k is the rate of the class k (corresponding to ancestors present approximately $0.5 \times R_k$ generations ago; Hayes et al. 2003). A more complete description of this model can be found in Druet and Gautier (2017) or Solé et al. (2017). The inbreeding coefficient F_{HBD}, was estimated as the probability to belong to any of the HBD classes averaged over the whole genome.

In addition to genomic measures, we also estimated the inbreeding coefficient based on the pedigree data (F_{PED}; Wright 1922), including a distinction between recent inbreeding, from contributions of ancestors present in the last five generations of the pedigree, and ancient inbreeding, from earlier contributors (see Supplementary Text S1).

Properties of inbreeding coefficients

The inbreeding coefficient has been defined in terms of correlations between the parents' uniting gametes by Wright

(1922) and as the probability that two homologous alleles in an individual are IBD by Malécot (1948). It also measures the fraction by which the heterozygosity has been reduced due to inbreeding (Crow and Kimura 1970). The different inbreeding coefficients used in the present study match often more closely to one of these definitions. For instance, F_{UNI} is directly related to the definition of Wright (1922), F_{PED} fits that from Malécot (1948) and F_{HOM} measures the heterozygosity reduction. Different estimators thus have different properties, some of which are summarized below. When defined as an IBD probability, estimators must range between 0 and 1 and F_{PED}, F_{ROH}, F_{HBD} and F_{ML} fall in that category. The other measures, F_{UNI}, F_{GRM} and F_{HOM}, can take negative values and behave more like correlations (e.g., Wang 2014). A base population where individuals are considered unrelated must also be defined when relying on IBD probabilities. For F_{PED}, this corresponds obviously to the founders of the pedigree, whereas for F_{HBD} and F_{ROH}, the base population depends on the shortest identified HBD segments as their size is related to the number of generations to the common ancestor. For other coefficients, the base population is indirectly defined by the set of individuals used to estimate the allele frequencies (e.g., Wang 2014). Interestingly, F_{HOM} and F_{ROH} weight all alleles equally, whereas F_{UNI} and F_{GRM}, two methods relying on correlations or covariances between genetic effects, give more weight to homozygosity at rare alleles (VanRaden 2008; Keller et al. 2011). Allele frequencies are used differently by F_{ML} and F_{HBD}, which rely on the probabilities to observe genotypes conditionally on F and Hardy-Weinberg proportions. In both cases, genotypes come from a mixture of two distributions (autozygous *vs* allozygous) and F is estimated as the value maximizing the likelihood function. We show that F_{HBD} and F_{ML} are indeed equivalent when the SNPs are considered independent in F_{HBD} (see Supplementary Text S2). Independence between SNPs would correspond to homozygosity-by-descent resulting from very distant ancestors separated by many generations of recombination. Hardy-Weinberg proportions are also used in F_{HOM} to estimate the expected total number of homozygous genotypes. Finally, F_{HBD} and F_{ROH} exploit information from neighboring SNPs, by identifying sequences of homozygous markers. The length of these homozygous stretches is informative about the number of generations to the common ancestor. These estimators also provide the ability to estimate locus-specific inbreeding coefficients. Overall, although the different metrics share some properties, their connections remain relatively complex.

SNP annotation

To evaluate the properties of different inbreeding coefficients, we compared them for different groups of SNPs from the WGS data. Therefore, we started by classifying the SNPs according to different criteria mainly related with their putative deleteriousness, such as their frequency, age and predicted functional effect.

Marker allele frequency

Allele frequency (AF) was selected as a criterion since it is linked to the age of the alleles and to their possible selection coefficient, i.e., their deleterious effect.

Age of alleles

Deleterious alleles are expected to be young since purifying selection eliminates them relatively rapidly. Unfortunately, we did not know the true age of the alleles identified in our dataset, yet some indicators were available. First, allele frequency can be utilized as a proxy for relative allele age (Kelleher et al. 2019). Secondly, alleles observed in multiple populations (or breeds) can be assumed to be on average older than alleles observed only in our Holstein pedigree. Thus, alternate alleles not observed in a sample of 50 whole-genome sequenced Belgian Blue Beef cattle used in Charlier et al. (2016), hereafter referred to as 'private alleles', allowed us to enrich our set of variants in young alleles.

To validate these hypotheses, we used the approach of Albers and McVean (2020) for dating genomic variants implemented in GEVA (Genealogical Estimation of Variant Age). We first phased variants from *Bos taurus* autosome (BTA) 25 with Beagle 4.0 (Browning and Browning 2007) using the pedigree option. We then ran GEVA and relied on the recombination clock to estimate the age of alleles.

Functional annotation

Deleterious or beneficial variants are more likely to be coding or regulatory variants. Therefore, the VCF file was annotated into different functional categories using Variant Effect Predictor (VEP) (McLaren et al. 2016). VEP predicts consequences of variants on protein sequence and uses Sorting Intolerant From Tolerant (SIFT) scores (Ng and Henikoff 2003) to determine which amino acid substitutions are deleterious or tolerated. Three classes of SNPs were then created using this information: synonymous variants, tolerated missense variants and deleterious missense variants. Variants classified as 'low confidence' by VEP were excluded from the analysis.

Empirical evaluation of the properties of inbreeding coefficients using metrics computed from WGS data

We compared different measures of inbreeding estimated with the 37,675 array-like SNPs with homozygosity

measured at different groups of alleles from the WGS data including 12,735,685 autosomal SNPs. These latter homozygosity measures were used to mimic homozygosity at alleles causing ID or the impact of inbreeding on whole-genome homozygosity. Correlations between different estimators and these homozygosity scores were used to evaluate the performances of the different methods. The standard errors of these correlations were obtained using the Fisher transformation.

Homozygosity for different allele frequency groups

We computed homozygosity of alternate alleles grouped in different frequency classes (e.g., 0.0–0.05, 0.05–0.10, etc.) to understand how efficiently inbreeding coefficients captured homozygosity in these different classes. Similarly, we estimated marker homozygosity (i.e., homozygosity at both reference and alternate alleles) per class of MAF. Note that this measure also reflects heterozygosity, as both measures sum to one. All these metrics helped us to compare the general properties of inbreeding coefficients, but also their association with subsets of SNPs potentially associated with ID. For instance, homozygosity at low frequency alleles would be related to the homozygosity at partially recessive detrimental alleles, whereas heterozygosity at intermediate frequency alleles would be associated with the heterozygosity at overdominant alleles.

Whole-genome sequence homozygosity

We next considered the total WGS homozygosity as another metric, capturing the genome-wide impact of inbreeding. Inbreeding increases homozygosity at all loci of the genome simultaneously (e.g., Szulkin et al. 2010; Wang 2014). Consequently, the inbreeding coefficient can be measured as the fraction by which heterozygosity has been reduced (Crow and Kimura 1970), and WGS homozygosity has been suggested as a measure to empirically evaluate different inbreeding estimators (Kardos et al. 2016).

Homozygous mutation load (HML)

Following Keller et al. (2011), we counted the number of rare or low-frequency alleles that were homozygous per individual, considered to be a proxy for ID. We computed HML at different allele frequency (AF) thresholds (0.05, 0.10 to 0.15) to determine whether results were sensitive to frequency of the alleles included in the HML score.

A weighted HML (wHML) was also computed using the inverse of allele frequency as weights, as deleterious effects are expected to be stronger for rarer alleles (e.g., Yengo et al. 2017). HML scores were also computed specifically for non-synonymous, tolerated and deleterious missense variants.

Regional scores (locus-specific)

To study the properties of regional inbreeding coefficients at a specific locus, we estimated regional homozygosity and regional HML scores using all SNPs present in non-overlapping 1 Mb windows, and compared them with regional measures of the inbreeding coefficient. F_{UNI} or F_{GRM} were computed using only the markers from the Illumina BovineSNP50 BeadChip present in the respective 1 Mb window. For F_{HBD} and F_{ROH}, we averaged the HBD probabilities and the ROH status (0/1), respectively, from SNPs in the window. For regional homozygosity, windows with less than 5 SNPs on the 50 K array were excluded from the analysis, whereas for regional HML, windows with less than 2000 variants from the WGS data, less than 5 SNPs on the 50 K array or less than 8 individuals with a non-zero score, were excluded.

Results

Inbreeding measures

Descriptive statistics of the estimated inbreeding coefficients are reported in Table S1. F_{ML}, F_{HBD}, F_{ROH} and F_{PED} values were always positive. With F_{ML}, 81 out of 145 parents had a null inbreeding coefficient whereas no null values were reported for F_{PED}, F_{HBD} or F_{ROH}. With the exception of F_{ML}, genomic measures presented higher variances than F_{PED}, with the largest values observed for F_{GRM} followed by F_{HOM} and F_{UNI}. The correlations between all measures are available in Table S2. The HBD-based measure was highly correlated with F_{ROH} (0.95) and with the excess of homozygosity (F_{HOM}) measure (0.96). Their correlations with F_{PED} were relatively high, equal to 0.76, 0.77 and 0.83 for F_{HBD}, F_{HOM} and F_{ROH}, respectively. The correlation between F_{UNI} and F_{GRM} was also strong ($r = 0.88$). The correlation between F_{UNI} and F_{ML} was 0.76, but increased to 0.96 when only the individuals with $F_{ML} > 0$ were considered (Supplementary Fig. S1). Both measures had high or moderate correlations with all other measures. Measures of F_{PED}, F_{ROH} and F_{HBD} considering all pedigreed generations and all fragment lengths showed larger correlation with the other F estimators than measures assuming only recent/ancient generations, or short/long fragments (Table S2). Thus, only full measures will be considered in the following, where $F_{HBD-525}$ will be referred to as F_{HBD}. Additional details on the results obtained with these partitioned F measures are given in Supplementary Text S1 and Supplementary Figs S2-5.

Age of alleles

Using the software GEVA we predicted the age of 231,111 alternate alleles located on BTA25 using the recombination

Table 1 Average age of alternate alleles from BTA25, expressed in Time to the Most Recent Common Ancestor (TMRCA), estimated with GEVA and using the recombination clock.

	Alternate alleles		Private alternate alleles	
Allele frequency	Number of alleles	Average TMRCA	Number of alleles	Average TMRCA
$0 <. \leq 0.05$	59,943	604	19,108	361
$0.05 <. \leq 0.15$	50,596	1284	6136	564
$0.15 <. \leq 0.30$	43,931	1820	1633	772
$. > 0.30$	76,281	2946	180	4851

The time is measured in generations. Private alleles refer to alleles absent from a whole-genome sequence data sample available for another breed (Belgian Blue cattle). Private alleles are selected to enrich the set of SNPs in young alleles.

clock. The Time to the Most Recent Common Ancestor (TMRCA) was estimated in generations and represented a relative measure that allowed us to compare categories of alleles. Private alternate alleles had clearly lower average TMRCA (462) than alternate alleles in general, private alleles included (1758). Table 1 provides the average TMRCA for both types of alleles classified according to their allele frequency. Alleles segregating at lower frequency were younger and more so if they were not present in the Belgian Blue cattle sample (i.e., private). Interestingly, private variants were enriched in low frequency alleles, with very few alleles having a frequency >0.30. In summary these results confirmed that private alleles segregating at low frequency were enriched in young alleles.

Genome-wide comparisons of estimated inbreeding coefficients

Correlations with homozygosity measured in different allele frequency (AF) classes

The correlations between the inbreeding coefficients and the homozygosity at alternate alleles grouped according to their frequency (20 classes) are plotted in Fig. 1a. For the least frequent alleles, correlations ranged from as low as 0.05 (F_{PED}) to 0.76 (F_{GRM}). Alleles with slightly larger frequencies (from 0.05 to 0.15) presented higher correlations with all metrics, indicating that the lowest frequency alleles might be more difficult to capture with inbreeding coefficients. For allele frequencies below 0.25, the two methods giving more weights to rare alleles (F_{UNI} and F_{GRM}) performed the best followed by likelihood-based methods (F_{ML} and F_{HBD}). Metrics giving equal weight to all alleles such as F_{HOM} and F_{ROH} were less efficient for rare alleles, but better for homozygosity at frequent alleles for which F_{UNI}, F_{GRM} and F_{ML} were clearly less useful.

The pedigree-based measure (F_{PED}) achieved the lowest correlations and presented patterns similar to F_{HOM} and F_{ROH}, estimators relying on whole-genome homozygosity. For instance, all these estimators presented lower correlations when homozygosity was estimated using rare alleles.

Correlations with homozygosity at private alleles

When homozygosity was computed using private alleles (enriched in younger alleles), F_{UNI}, F_{GRM} and F_{ML} were still the best estimators to capture homozygosity at rare alleles (AF < 0.10), particularly for the lowest frequency class (Fig. 1b), but these correlations were lower than in the previous section. Conversely, the correlations increased for other inbreeding measures, which became more efficient when homozygosity was measured specifically at private alleles. Consequently, the methods presented smaller differences in terms of correlations. For alleles with frequencies ranging from 0.10 to 0.25, F_{HBD}, F_{HOM}, F_{ROH} and F_{PED} were even more efficient than F_{UNI}, F_{GRM} and F_{ML} for which correlations were strongly reduced. Finally, for the class of variants with an AF > 0.25, the small number of private alleles reaching these frequencies reduced the reliability of the analysis.

Correlations with marker homozygosity measured in different MAF classes

Correlations between inbreeding coefficients and marker homozygosity are in line with observations for allele homozygosity metrics (Fig. 1c). Indeed, marker homozygosity at SNPs with low MAF results mainly from homozygosity at frequent alleles. Consequently, methods that captured well homozygosity at frequent alleles (F_{HOM}, F_{ROH} and F_{PED}) had the strongest correlations with marker homozygosity at SNPs with low MAF. Conversely, F_{UNI}, F_{GRM} and F_{HBD} performed better when MAF was higher than 0.25. Overall, most inbreeding coefficients were better at capturing marker homozygosity for alleles segregating at intermediate frequency (MAF > 0.15) than homozygosity at private and rare alleles (AF < 0.15).

Correlations with whole-genome homozygosity

When whole-genome homozygosity was estimated for all alleles (Fig. 1d, Supplementary Fig. S6), irrespective of their MAF or age, F_{ROH} and F_{HOM} presented the highest correlation
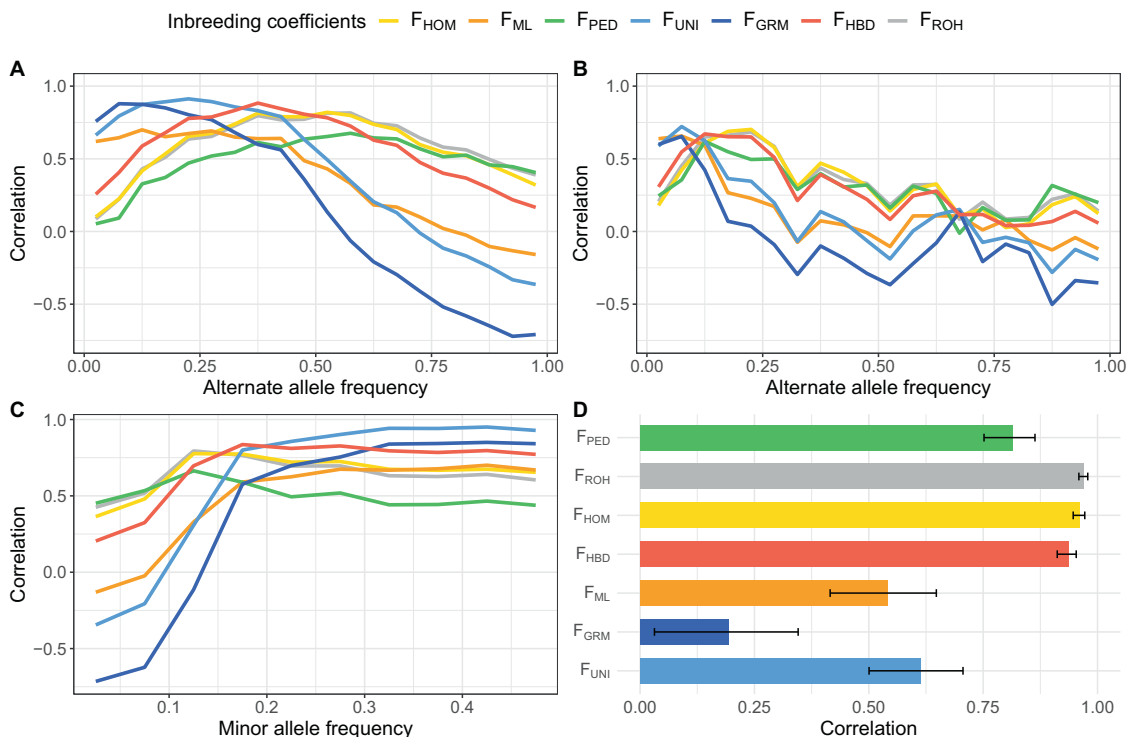
Fig. 1 **Correlation coefficients between individual inbreeding measures estimated with 37,675 SNPs and scores obtained from the whole-genome sequence data in 145 individuals. a** Correlation with homozygosity at alternate alleles grouped according to their allele frequency. **b** Correlation with homozygosity at private alleles (young alleles) grouped according to their allele frequency. **c** Correlation with global marker homozygosity (counted at both reference and alternate alleles) as a function of minor allele frequency. **d** Correlation with whole-genome homozygosity. The error bars represent the 95% confidence intervals.

(0.97 and 0.96, respectively), closely followed by F_{HBD} (0.94). Interestingly, F_{PED} was also highly correlated (0.81), even more than the remaining genomic measures, which give more weight to rare alleles, while F_{GRM} presented a relatively weak correlation with this score (0.19).

Correlations with homozygous mutations load (HML)

Methods giving more weight to rare alleles such as F_{UNI} and F_{GRM}, better captured HML (Fig. 2a, Supplementary Figs S7–9) in agreement with their better correlation with homozygosity at rare alleles. The differences with other estimators were larger for lower AF thresholds. We subsequently weighted alleles by the inverse of their frequency, as rare alleles are more likely to have strong deleterious effects (e.g., Yengo et al. 2017). In that case, correlations varied little for the lowest frequency threshold (0.05), whereas for higher thresholds the correlations were somehow reduced as expected (Fig. 2b). The HML was then computed for synonymous, missense tolerated and missense deleterious variants with an AF threshold set at 0.15 (Fig. 2c, Supplementary Figs S10–12). Alleles in the most damaging classes were less frequent. Correlations obtained with metrics such as F_{UNI}, F_{GRM} and F_{ML} decreased, more so for more deleterious classes. Interestingly, the other

measures had the opposite behavior: higher correlations for these specific classes than for general HML and better performances for more deleterious alleles. Nevertheless, their performance was still below that from the first group of methods. F_{ROH} had correlations similar to those obtained with F_{HOM}.

As deleterious alleles are expected to be rare and young, we re-estimated the HML and wHML using only private alleles that are enriched in young alleles. As observed before, F_{UNI} and F_{GRM} had lower correlations with private alleles, whereas the other methods performed better (Fig. 2d–e, Supplementary Figs S13–15). As a result, differences between methods were smaller and F_{ML} performed better than F_{GRM}. Interestingly, all inbreeding coefficients had correlations higher than 0.50 with AF threshold set at 0.15. When HML was computed with synonymous or missense variants, F_{UNI} still presented the highest correlations but F_{HBD} was now second, for all three sets of variants (Fig. 2f, Supplementary Figs S16–18). With private alleles, homozygosity at more deleterious alleles was more difficult to capture irrespective of the method. As before, correlations obtained with F_{UNI}, F_{GRM} and F_{ML} dropped when considering only the variants in coding regions whereas correlations with other metrics were less impacted. As a result, smaller differences were observed between methods, in
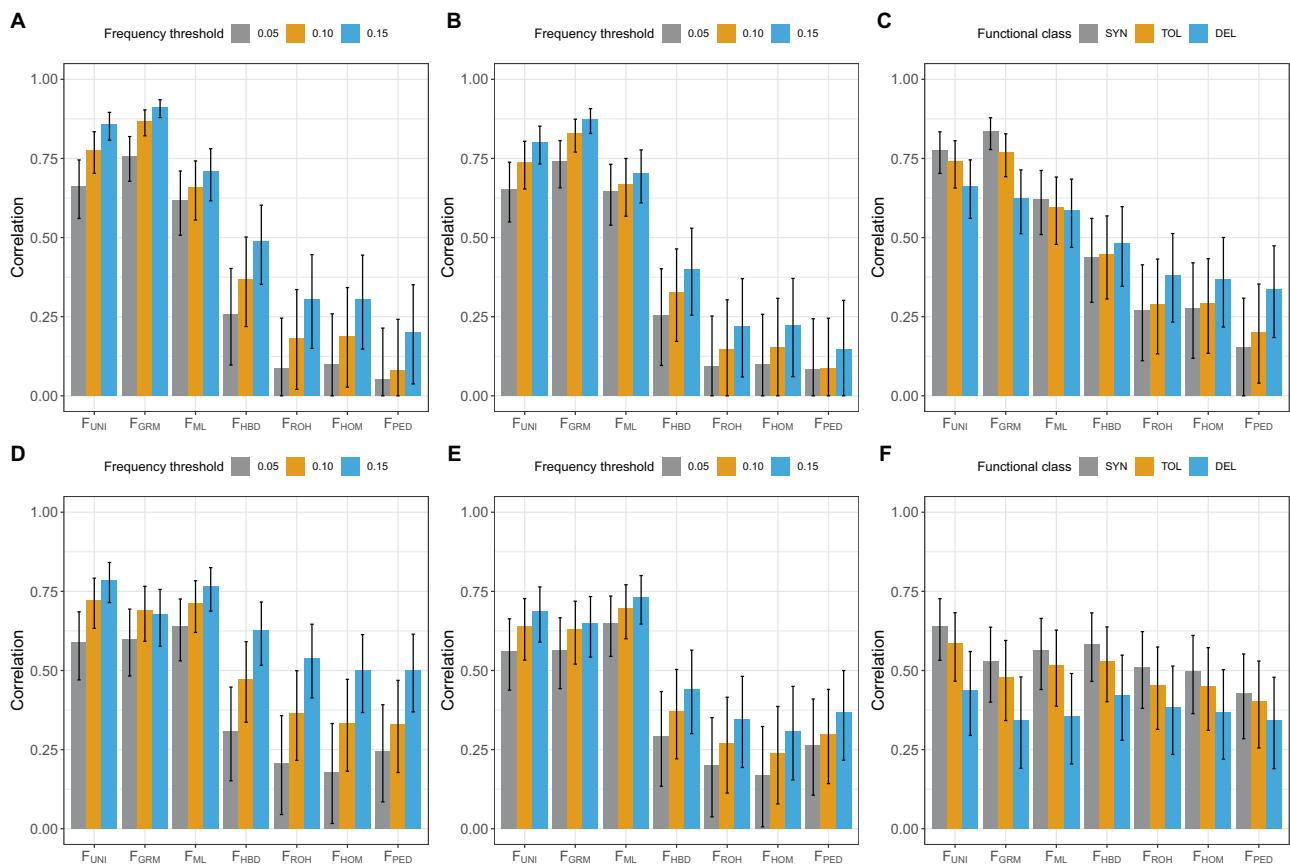
Fig. 2 Correlation coefficients between individual inbreeding measures estimated with 37,675 SNPs and homozygous mutation load (HML) obtained from the whole-genome sequence data in 145 individuals. HML was computed using alternate (**a**, **b**, **c**) and private alternate (**d**, **e**, **f**) alleles. **a** and **d** Correlation with HML estimated with allele frequency thresholds of 0.05, 0.10 and 0.15. **b** and **e** Correlation with weighted HML estimated with allele frequency thresholds of 0.05, 0.10 and 0.15. **c** and **f** Correlation with HML estimated with synonymous (SYN), tolerated (TOL) and deleterious (DEL) missense variants and using an allele frequency threshold of 0.15. The error bars represent the 95% confidence intervals and are truncated at 0.

particular when HML was computed with deleterious missense variants only. With these HML scores derived from private alleles, F_{ROH} performed better than F_{HOM}. Note that when private alleles with specific annotations were used, HML scores were derived from fewer variants. Therefore, these correlations should be interpreted cautiously.

Comparisons with regional scores

Inbreeding measures were also evaluated for their association with regional scores estimated in 1 Mb non-overlapping windows (see methods). F_{ML} was excluded from the comparisons due to its long running times and because it did not perform best in genome-wide comparisons.

We started by computing regional homozygosity measured at all alleles irrespective of their frequency (Fig. 3a). When averaged over all windows, correlations with genomic inbreeding coefficients were relatively high for F_{HBD} (0.75), F_{HOM} (0.74), F_{UNI} (0.67) and F_{ROH} (0.62) but somewhat lower with F_{GRM} (0.44). Pedigree-based estimators were clearly below all genomic measures with an average correlation close to zero (0.08). There was nevertheless considerable variation between regions of the genome, in particular with F_{GRM} (Fig. 3a).

Regional HML was then computed using alternate alleles with $AF \leq 0.15$ (Fig. 3b). Correlations between local inbreeding measures and regional HML were lower than those obtained with whole-genome HML scores and showed a larger variation. For instance, they ranged from -0.12 to 0.88 for F_{UNI}. On average, F_{UNI} performed best (0.43), followed by F_{GRM} (0.41), F_{HBD} (0.39), F_{ROH} (0.34) and F_{HOM} (0.25), whereas F_{PED} had almost null average correlations (0.02). The ranking of the methods changed however from window to window.

Discussion

We utilized cattle whole-genome sequence data to empirically evaluate different estimators of the inbreeding
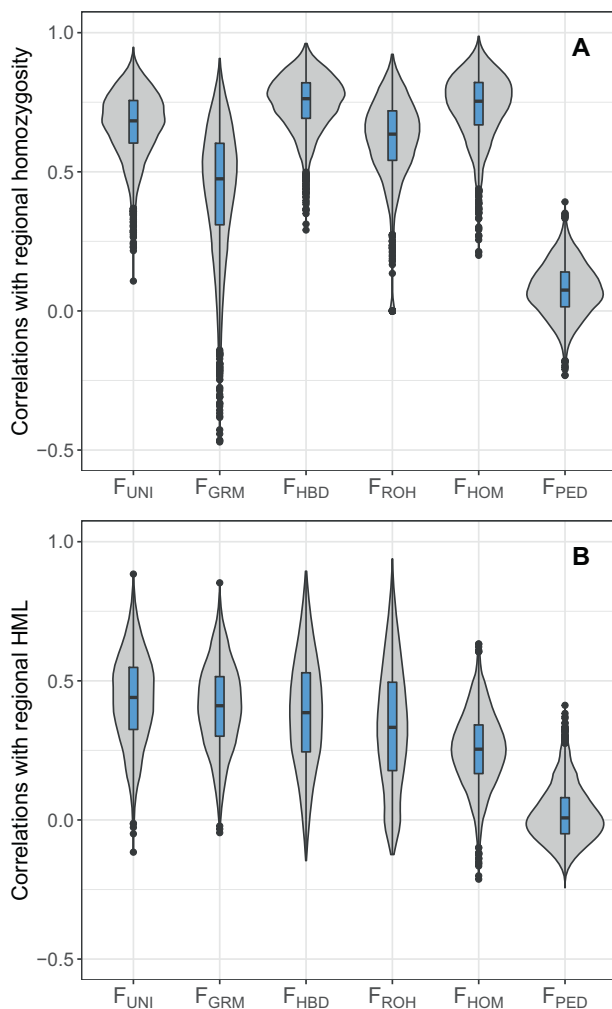
Fig. 3 Correlation coefficients between individual regional inbreeding measures and regional scores in 1 Mb windows computed from the whole-genome sequence data in 145 individuals. The regional inbreeding coefficients were estimated only with markers present among the 37,675 SNPs from the bovine genotyping array (see Methods for more details). The correlations for ~2500 windows are presented as a violin plot combined with an inner boxplot. **a** Correlation with regional homozygosity. **b** Correlation with regional homozygous mutation load (HML).

rare alleles and marker homozygosity at SNPs with moderate to high MAF. A second group of metrics based on the number of homozygous SNPs that give equal weights to all alleles, including F_{HOM} and F_{ROH}, achieved the highest correlations with whole-genome homozygosity, but were less efficient to capture homozygosity at rare alleles. When homozygosity was measured for private sets of alleles that were shown to be enriched in young alleles, the performance of the latter measures improved whereas it decreased for the first group of estimators. Interestingly, the properties observed for F_{PED} matched those of the second group. The first group of methods relies on correlations between parental gametes (F_{UNI}) or variances of genotypes within an individual (F_{GRM}) and better fits the definition of the inbreeding coefficient in terms of correlation proposed by Wright ([1922]). Conversely, the second group behaving similar to F_{PED} would correspond to the definition by Malécot ([1948]), relying on the probability that two homologous alleles in an individual are IBD (without imposing any constraint on locus position or on allele frequency). Indeed, they performed better when alleles are young (i.e., more likely to be IBD) and measure the increased proportion of homozygosity (correlated with the increased proportion of autozygosity) at all variants irrespective of their frequency. The last two measures, F_{ML} and F_{HBD}, both relying on likelihood maximization (see Methods), presented intermediate properties. We observed that F_{ML} was highly correlated with F_{UNI} for positive inbreeding coefficients (Supplementary Fig. S1) and thus behaved in a manner similar to the first group. In contrast, F_{HBD} was closer to the properties of the second group. Although F_{HBD} uses allele frequencies to compute HBD probabilities, homozygous genotypes that are in long HBD segments receive the same weight irrespective of their AF, as it occurs with F_{ROH} or F_{HOM}.

Our approach can also be used to investigate other aspects related to inbreeding coefficients. For instance, we also studied the ability from different methods to work regionally. Such locus-specific estimators could be useful for performing homozygosity mapping experiments to identify regions associated with recessive diseases or ID (Abney et al. [2002]; Leutenegger et al. [2006]). Similarly, the approach allows the study of the properties of inbreeding coefficients estimated at lower marker density (Supplementary Text S3). Robustness at low marker density is important for applications in agricultural species, where such low-density arrays are sometimes used to reduce genotyping costs, but also for non-model species where high-density arrays might not be available. For both applications, the ranking between methods and their properties remained in line with the high-density results (Supplementary Figs S19–21). As expected, regional homozygosity or HML were more difficult to capture than genome-wide

coefficient. This sample represents a population with small N_e and under intense selection. It brings therefore complementary information to studies relying on populations with large N_e, such as humans (e.g., Yengo et al. [2017]). It is informative for agricultural species but also for wild species with small N_e, including for populations in conservation programs. Our results must be interpreted cautiously, in particular for the rarest alleles, as the sample size was relatively small. Nevertheless, this approach revealed some properties of the inbreeding coefficient estimators. A first group of methods that give higher weight to homozygosity at rare alleles, including F_{UNI} and F_{GRM}, presented the strongest correlations with both genome homozygosity at

scores. With fewer markers, correlations were also lower, but this reduction remained limited for most methods. Interestingly, F_{HBD} was still efficient at low marker density and appeared to be a good compromise in that case, particularly for regional scores (Supplementary Fig. S21). The method uses local information from neighboring SNPs and the genetic map in a probabilistic framework that accounts for uncertainty; two important elements at low density. With the same approach, properties of recent and ancient inbreeding could also be revealed. For instance, estimators obtained with long *versus* short ROH or using recent *versus* ancient pedigree generations can be compared (Supplementary Text S1). In both cases, inbreeding coefficients using all ROH or all pedigree-generations presented the highest correlations with homozygosity measures (Supplementary Figs S2–3). Nevertheless, the longest ROH (>5 Mb) or the five last pedigree generations accounted for most of the variation between individual inbreeding levels (Supplementary Table S1). Associated estimators performed relatively well, even when ROH were restricted to >10 Mb. Conversely, inbreeding coefficients associated with short ROH (<5 Mb) or with more ancient pedigree generations presented limited variation. Likewise, HBD-measures including all HBD segments better captured homozygosity at rare alleles or HML than related measures considering only the longest segments associated with recent ancestors (Supplementary Figs S4–5). Finally, we also investigated the properties of inbreeding coefficients predicted in offspring thanks to parental genotypes (Supplementary Text S4). Such predictions are important to manage inbreeding levels in livestock species or in conservation programs. With these predicted values, correlations with scores computed from the WGS data were lower than when the inbreeding coefficient was estimated using the genotypes from the individual, as expected (Supplementary Figs S22–24). The same dichotomy between methods predicting well homozygosity at rare alleles and those capturing better whole-genome homozygosity was observed.

The properties highlighted by our empirical approach can also contribute to understand properties from heterozygosity-fitness correlation (HFC) approaches (e.g., Pemberton 2004; Szulkin et al. 2010). The absence of HFC in certain studies has generated debate in the past (David 1998; Pemberton 2004; Szulkin et al. 2010). Several hypotheses have previously been proposed to explain this observation (e.g., David 1998; Slate and Pemberton 2002; Szulkin et al. 2010). For instance, it was postulated that heterozygosity at a few markers (most often micro-satellites) might not capture heterozygosity at other variants, in particular those causing ID (e.g., Balloux et al. 2004; Grueber et al. 2011). It was recommended to use identity disequilibrium measures (Balloux et al. 2004; Slate et al.

2004) to evaluate the correlation between homozygosity at different loci and to assess whether marker heterozygosity was expected to capture differences in genome-wide heterozygosity levels resulting from inbreeding. Here, we observed that with 6000 SNPs (Supplementary Text S3, Supplementary Fig. S19) the genome-wide homozygosity was highly correlated with inbreeding coefficients related to marker homozygosity (F_{ROH}, F_{HOM}, F_{HBD}). However, the homozygosity at rare (deleterious) alleles proved more difficult to capture. Therefore, the correlation with fitness or ID might still be low even when identity disequilibrium is high, for instance if identity disequilibrium is measured among frequent alleles and does not reflect correlation with rare deleterious alleles. Several studies also reported that pedigree measures might present higher correlations with fitness than marker heterozygosity, and recommended F_{PED} as the inbreeding measure to use (e.g., Pemberton 2004; Grueber et al. 2011; Nietlisbach et al. 2017). These results were however most often obtained with relatively few markers (e.g., Grueber et al. 2011; Nietlisbach et al. 2017) and several authors subsequently stated that genomic measures were superior to pedigree-based estimators (e.g., Keller et al. 2011; Wang 2016). Here, we confirm that marker-based inbreeding coefficients performed better than pedigree-based ones, in particular for regional scores that had almost null correlations with F_{PED}.

Inbreeding depression is mainly caused by an accumulation of partially recessive deleterious mutations (Charlesworth and Charlesworth 1999) which, in general, are young and remain at low frequency (e.g., Pritchard 2001). Accordingly, HML has been proposed by Keller et al. (2011) as a proxy for ID. They showed in their study that the homozygosity at alleles with a frequency below 0.05 was indeed similar to homozygosity at recessive deleterious alleles. However, the optimal AF threshold depends on the population demographic history and its N_e. When N_e is low, as for livestock species, domestic animals or endangered species, alleles with larger selection coefficients can remain effectively neutral (as long as $N_e s \ll 1$) and deleterious alleles can reach higher frequencies compared to human populations (Kimura 1983). Since selection is less effective in small populations, mildly deleterious mutation can accumulate (Keller and Waller 2002) and even become fixed (Frankham 1995). When $N_e \le 100$, as in several cattle breeds, mildly deleterious variants might reach frequencies around 0.15. Furthermore, mildly deleterious alleles might also segregate at high frequencies as a result of population bottlenecks experienced during domestication or breed creation, and as a result of artificial selection for linked favorable variants, through genetic hitch-hiking (see Bosse et al. 2019). As an illustration, genetic variants causing recessive defects reached frequencies above 0.10, and even higher in the most extreme cases, in Belgian Blue cattle

(Fasquelle et al. 2009; Sartelet et al. 2012; Druet et al. 2014; Charlier et al. 2016), but these alleles provided a potential heterozygous advantage. In the present study, F_{UNI}, F_{GRM} and F_{ML} captured HML better than other metrics, more so for rare alleles, suggesting that these methods could be more suited to estimate ID and to avoid fitness reduction associated with inbreeding in mating designs. When HML was computed with private alleles enriched for young alleles, the second group of estimators started to behave better and differences between methods were smaller. In particular, when HML was estimated for young and deleterious alleles with properties similar to those of variants causing ID, F_{HBD}, F_{ROH} and F_{HOM} had higher correlations than F_{GRM} or F_{ML}. Among the methods from the second group, F_{HBD} performed best, notably for regional scores and estimations at lower marker density. In humans, long ROH are enriched in homozygous deleterious alleles (Szpiech et al. 2013) whereas Zhang et al. (2015b) observed the opposite in cattle. Here, we show that both in terms of estimations or predictions, higher correlations with homozygosity at rare and young deleterious variants are obtained when also including shorter HBD segments or ROH (Supplementary Text S1). This is in agreement with recommendations from Kardos et al. (2018a, 2018b) and indicates that at least some of the deleterious variants are present in short HBD segments. However, it is important to keep in mind that HML is an imperfect proxy of ID and that all these correlations with HML must be interpreted cautiously. It is not known which variants are truly deleterious and whether alleles have favorable or negative effects. Ideally, we should use the variants causing ID, weighted by their effect. Finally, note that HML, and more particularly regional HML, could also somehow be related to the d² metric, which measures the distance between microsatellites alleles to capture their time of coalescence (Coulson et al. 1998). The number of homozygous SNPs reflects to a certain extent how closely related the uniting gametes were for that locus.

Overall, our empirical results illustrate that the best inbreeding coefficient estimator might depend on the frequency, age and effect size of alleles contributing to inbreeding depression and the population demographic history. Even for evaluating ID caused by recessive deleterious alleles, the present and past effective population size and the size of the allele effects will result in a different distribution of AF. Although F_{UNI} and F_{GRM} presented high correlations with homozygosity at rare alleles, other metrics might perform better for other groups of alleles. In case the contribution from heterozygosity at loci presenting heterozygous advantage to ID is important, as suggested by Charlesworth (2015), inbreeding coefficients should capture homozygosity at these loci segregating at intermediate frequency (see Fig. 1c). Overall, inbreeding coefficients presented higher correlations with homozygosity at such loci

than with homozygosity at rare alleles. Therefore, in that scenario, inbreeding coefficients would present higher correlations with ID, and F_{UNI} or F_{HBD} would perform best as they had the highest correlations with homozygosity at the target loci (Fig. 1c). The fact that different metrics capture homozygosity at SNPs with different properties makes the conclusions from different simulation studies difficult to interpret. Indeed, Yengo et al. (2017) had strong conclusions in favor of F_{UNI} as a preferred measure to estimate ID with human data, whereas Keller et al. (2011) or Nietlisbach et al. (2019) presented results in favor of F_{ROH}. However, phenotypes were simulated with different approaches and in populations with different structures. More recently, Caballero et al. (2020) have shown that the results of these papers are not contradictory. In scenarios of large population sizes, such as in human populations, F_{UNI} can be an appropriate inbreeding measure to estimate ID, whereas in scenarios of small population sizes, F_{ROH} may be more appropriate. Therefore, the inbreeding coefficient achieving the highest correlation with ID might differ according to the scenarios and populations considered.

The optimal inbreeding coefficient estimator varies also according to the intended application. When the inbreeding coefficient is used to measure the heterozygosity reduction at all alleles, irrespective of their frequencies or their age, the use of the second group of methods, which are more related to the proportions of autozygous genotypes (F_{HBD}, F_{HOM}, F_{ROH} and F_{PED}) is recommended. This information is important when the objective is to determine the extinction risk of a population, to assess whether a conservation program is efficiently implemented, to understand the recent demographic history from a population, or to estimate the effective population size. Similarly, these measures are useful to investigate mating systems in a population or to identify consanguineous matings. They might also be used to minimize inbreeding in small captive populations and to maintain diversity at all variants. In this group, F_{HBD} (or F_{ROH}) performed best and should be preferred to F_{HOM} or F_{PED}, in agreement with Keller et al. (2011). These measures are, in addition, easier to interpret as they have positive values and represent autozygosity accumulated relative to a base population. F_{HBD} also behaves well at lower marker densities and can be used to estimate locus-specific inbreeding coefficients or to perform homozygosity mapping experiments to identify regions associated with recessive diseases or ID (Abney et al. 2002; Leutenegger et al. 2006).

The results we have reported present limitations since they relied on some approximations. In particular, our sample size was relatively modest, and this could influence some results. It contained healthy adult animals and did not include individuals that suffered problems earlier in life. Ideally, such an evaluation of inbreeding measures should

be performed on larger samples of unselected animals. Measuring directly inbreeding depression in a large cohort of individuals as done by Yengo et al. (2017) would represent a complementary and valuable empirical evaluation of different inbreeding coefficients. Indeed, Szulkin et al. (2010) and Kardos et al. (2016) suggested that the most precise inbreeding measures should present the strongest association with ID.

Conclusions

Using an empirical approach relying on whole-genome sequence data from a small cattle pedigree, we studied the properties from different inbreeding coefficients. For instance, F_{UNI} was shown to have the highest correlations with rare alleles and might therefore present a strong association with ID when it results from the action of rare recessive deleterious alleles. Nevertheless, ID might remain difficult to capture when associated with rare missense variants. For locus-specific inbreeding measures, the ranking of the methods might change since F_{HBD} makes better use of the information from neighboring markers. Measures related to homozygosity (F_{HBD}, F_{ROH} or F_{HOM}) were more efficient to capture the proportion of the genome that is IBD, irrespective of allele frequency or age of alleles. Since F_{UNI} and F_{HBD}/F_{ROH} present complementary properties, they might both be used when testing for ID. Finally, we confirmed that genomic measures are superior to pedigree-based estimates. In particular, F_{PED} was uncorrelated with locus-specific scores.

Data availability

The genotypes used in the present study are available from the Dryad Digital Repository https://doi.org/10.5061/dryad.vx0k6djq8. The repository also contains the annotation of all the variants and the genotypes for subset of markers present of the bovine genotyping arrays. Original VCFs files are also available at https://www.ebi.ac.uk/ena/browser/view/PRJEB38336, under the name BPWG.vcf.gz.

Compliance with ethical standards

References

Abney M, Ober C, McPeek MS (2002) Quantitative-trait homozygosity and association mapping and empirical genomewide significance in large, complex pedigrees: fasting serum-insulin level in the Hutterites. Am J Hum Genet 70:920–934

Ackerman MS, Johri P, Spitze K, Xu S, Doak TG, Young K et al. (2017) Estimating seven coefficients of pairwise relatedness using population-genomic data. Genetics 206:105–118

Albers PK, McVean G (2020) Dating genomic variants and shared ancestry in population-scale sequencing data. PLoS Biol 18: e3000586

Balloux F, Amos W, Coulson T (2004) Does heterozygosity estimate inbreeding in real populations? Mol Ecol 13:3021–3031

Bertrand AR, Kadri NK, Flori L, Gautier M, Druet T (2019) RZooRoH: an R package to characterize individual genomic autozygosity and identify homozygous-by-descent segments. Methods Ecol Evol 10:860–866

Bjelland DW, Weigel KA, Vukasinovic N, Nkrumah JD (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

Bosse M, Megens H-J, Derks MF, de Cara ÁM, Groenen MA (2019) Deleterious alleles in the context of domestication, inbreeding, and selection. Evol Appl 12:6–17

Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet 81:1084–1097

Caballero A (2020) Quantitative Genetics. Cambridge University Press, Cambridge, UK

Caballero A, Villanueva B, Druet T (2020) On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. Evol Appl 00:1–13. https://doi.org/10.1111/eva.13126

Charlesworth B (2015) Causes of natural variation in fitness: evidence from studies of Drosophila populations. Proc Natl Acad Sci 112:1662–1669

Charlesworth B, Charlesworth D (1999) The genetic basis of inbreeding depression. Genet Res 74:329–340

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nat Rev Genet 10:783

Charlier C, Li W, Harland C, Littlejohn M, Coppieters W, Creagh F et al. (2016) NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. Genome Res 26:1333–1341

Clark DW, Okada Y, Moore KH, Mason D, Pirastu N, Gandin I et al. (2019) Associations of autozygosity with a broad range of human phenotypes. Nat Commun 10:1–17

Coulson TN, Pemberton JM, Albon SD, Beaumont M, Marshall TC, Guinness FE et al. (1998) Microsatellites reveal heterosis in red deer. Proc R Soc Lond B Biol Sci 265:489–495

Crow JF, Kimura M (1970) An introduction to population genetics theory. Harper & Row, Publishers, New York, Evanston and London

David P (1998) Heterozygosity–fitness correlations: new perspectives on old problems. Heredity 80:531–537

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491

Druet T, Ahariz N, Cambisano N, Tamma N, Michaux C, Coppieters W et al. (2014) Selection in action: dissecting the molecular underpinnings of the increasing muscle mass of Belgian Blue Cattle. BMC Genomics 15:796

Druet T, Gautier M (2017) A model-based approach to characterize individual inbreeding at both global and local genomic scales. Mol Ecol 26:5820–5841

Fasquelle C, Sartelet A, Li W, Dive M, Tamma N, Michaux C et al. (2009) Balancing selection of a frame-shift mutation in the MRC2 gene accounts for the outbreak of the Crooked Tail Syndrome in Belgian Blue Cattle. PLoS Genet 5:e1000666

Ferenčaković M, Hamzić E, Gredler B, Solberg TR, Klemetsdal G, Curik I et al. (2013) Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. J Anim Breed Genet 130:286–293

Ferenčaković M, Sölkner J, Kapš M, Curik I (2017) Genome-wide mapping and estimation of inbreeding depression of semen quality traits in a cattle population. J Dairy Sci 100:4721–4730

Frankham R (1995) Conservation genetics. Annu Rev Genet 29:305–327

Goudet J, Kay T, Weir BS (2018) How to estimate kinship. Mol Ecol 27:4121–4135

Grueber CE, Waters JM, Jamieson IG (2011) The imprecision of heterozygosity-fitness correlations hinders the detection of inbreeding and inbreeding depression in a threatened species. Mol Ecol 20:67–79

Harland C, Charlier C, Karim L, Cambisano N, Deckers M, Mni M, et al. (2017) Frequency of mosaicism points towards mutation-prone early cleavage cell divisions. Preprint at https://www.biorxiv.org/content/10.1101/079863v1

Hayes BJ, Visscher PM, McPartlan HC, Goddard ME (2003) Novel multilocus measure of linkage disequilibrium to estimate past effective population size. Genome Res 13:635–643

Hedrick PW (2012) What is the evidence for heterozygote advantage selection? Trends Ecol Evol 27:698–704

Hedrick PW, Garcia-Dorado A (2016) Understanding inbreeding depression, purging, and genetic rescue. Trends Ecol Evol 31:940–952

Jacquard A (1974) The genetic structure of populations. Springer-Verlag, New-York, NY

Kadri NK, Harland C, Faux P, Cambisano N, Karim L, Coppieters W et al. (2016) Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle. Genome Res 26:1323–1332

Kardos M, Åkesson M, Fountain T, Flagstad Ø, Liberg O, Olason P et al. (2018a) Genomic consequences of intensive inbreeding in an isolated wolf population. Nat Ecol Evol 2:124

Kardos M, Nietlisbach P, Hedrick PW (2018b) How should we compare different genomic estimates of the strength of inbreeding depression? Proc Natl Acad Sci 115:E2492–E2493

Kardos M, Taylor HR, Ellegren H, Luikart G, Allendorf FW (2016) Genomics advances the study of inbreeding depression in the wild. Evol Appl 9:1205–1218

Kelleher J, Wong Y, Wohns AW, Fadil C, Albers PK, McVean G (2019) Inferring whole-genome histories in large population datasets. Nat Genet 51:1330–1338

Keller MC, Visscher PM, Goddard ME (2011) Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics 189:237–249

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. Trends Ecol Evol 17:230–241

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge, UK

Leroy G (2014) Inbreeding depression in livestock species: review and meta-analysis. Anim Genet 45:618–628

Leutenegger A-L, Labalme A, Genin E, Toutain A, Steichen E, Clerget-Darpoux F et al. (2006) Using genomic inbreeding coefficient estimates for homozygosity mapping of rare recessive traits: application to Taybi-Linder syndrome. Am J Hum Genet 79:62–66

Li CC, Horvitz DG (1953) Some methods of estimating the inbreeding coefficient. Am J Hum Genet 5:107

Malécot G (1948) Mathématiques de l'hérédité. Masson et Cie, Paris

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A et al. (2016) The ensembl variant effect predictor. Genome Biol 17:122

McQuillan R, Leutenegger A-L, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L et al. (2008) Runs of homozygosity in European populations. Am J Hum Genet 83:359–372

Milligan BG (2003) Maximum-likelihood estimation of relatedness. Genetics 163:1153–1167

Ng PC, Henikoff S (2003) SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812–3814

Nietlisbach P, Keller LF, Camenisch G, Guillaume F, Arcese P, Reid JM et al. (2017) Pedigree-based inbreeding coefficient explains more variation in fitness than heterozygosity at 160 microsatellites in a wild bird population. Proc R Soc B Biol Sci 284:20162763

Nietlisbach P, Muff S, Reid JM, Whitlock MC, Keller LF (2019) Nonequivalent lethal equivalents: models and inbreeding metrics for unbiased estimation of inbreeding load. Evol Appl 12:266–279

Pemberton J (2004) Measuring inbreeding depression in the wild: the old ways are the best. Trends Ecol Evol 19:613–615

Pew J, Muir PH, Wang J, Frasier TR (2015) related: an R package for analysing pairwise relatedness from codominant molecular markers. Mol Ecol Resour 15:557–561

Pritchard JK (2001) Are rare variants responsible for susceptibility to complex diseases? Am J Hum Genet 69:124–137

Pryce JE, Haile-Mariam M, Goddard ME, Hayes BJ (2014) Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genet Sel Evol 46:71

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

Ritland K (1996) Estimators for pairwise relatedness and individual inbreeding coefficients. Genet Res 67:175–185

Santure AW, Stapley J, Ball AD, Birkhead TR, Burke T, Slate J (2010) On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. Mol Ecol 19:1439–1451

Sartelet A, Druet T, Michaux C, Fasquelle C, Géron S, Tamma N et al. (2012) A splice site variant in the bovine RNF11 gene

compromises growth and regulation of the inflammatory response. PLoS Genet 8:e1002581

Slate J, David P, Dodds KG, Veenvliet BA, Glass BC, Broad TE et al. (2004) Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. Heredity 93:255–265

Slate J, Pemberton JM (2002) Comparing molecular measures for detecting inbreeding depression. J Evol Biol 15:20–31

Solé M, Gori A-S, Faux P, Bertrand A, Farnir F, Gautier M et al. (2017) Age-based partitioning of individual genomic inbreeding levels in Belgian Blue cattle. Genet Sel Evol 49:92

Speed D, Balding DJ (2015) Relatedness in the post-genomic era: is it still useful? Nat Rev Genet 16:33

Szpiech ZA, Xu J, Pemberton TJ, Peng W, Zöllner S, Rosenberg NA et al. (2013) Long runs of homozygosity are enriched for deleterious variation. Am J Hum Genet 93:90–102

Szulkin M, Bierne N, David P (2010) Heterozygosity-fitness correlations: a time for reappraisal. Evol Int J Org Evol 64:1202–1217

VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91:4414–4423

Wang J (2007) Triadic IBD coefficients and applications to estimating pairwise relatedness. Genet Res 89:135–153

Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. Mol Ecol Resour 11:141–145

Wang J (2014) Marker-based estimates of relatedness and inbreeding coefficients: an assessment of current methods. J Evol Biol 27:518–530

Wang J (2016) Pedigrees or markers: which are better in estimating relatedness and inbreeding coefficient? Theor Popul Biol 107:4–13

Wright S (1922) Coefficients of inbreeding and relationship. Am Nat 56:330–338

Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88:76–82

Yengo L, Zhu Z, Wray NR, Weir BS, Yang J, Robinson MR et al. (2017) Detection and quantification of inbreeding depression for complex traits from SNP data. Proc Natl Acad Sci 114:8602–8607

Yengo L, Zhu Z, Wray NR, Weir BS, Yang J, Robinson MR et al. (2018) Estimation of inbreeding depression from SNP data. Proc Natl Acad Sci 115:E2494–E2495

Zhang Q, Calus MP, Guldbrandtsen B, Lund MS, Sahana G (2015a) Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. BMC Genet 16:88

Zhang Q, Guldbrandtsen B, Bosse M, Lund MS, Sahana G (2015b) Runs of homozygosity and distribution of functional variants in the cattle genome. BMC Genomics 16:542