



# Article Using Pedigree and Genomic Data toward Better Management of Inbreeding in Italian Dairy Sheep and Goat Breeds

Matteo Cortellari <sup>1,†</sup><sup>®</sup>, Alessio Negro <sup>1,2,†</sup><sup>®</sup>, Arianna Bionda <sup>1,\*</sup><sup>®</sup>, Silverio Grande <sup>2</sup>, Alberto Cesarani <sup>3,4</sup><sup>®</sup>, Antonello Carta <sup>5</sup>, Nicola Macciotta <sup>3</sup>, Stefano Biffani <sup>6</sup><sup>®</sup> and Paola Crepaldi <sup>1</sup><sup>®</sup>

- <sup>1</sup> Dipartimento di Scienze Agrarie e Ambientali—Produzione, Territorio, Agroenergia, Università degli Studi di Milano, 20133 Milan, Italy
- <sup>2</sup> Ufficio Studi, Associazione Nazionale della Pastorizia (Asso.Na.Pa.), 00187 Rome, Italy
- <sup>3</sup> Dipartimento di Agraria, Università degli Studi di Sassari, 07100 Sassari, Italy
- <sup>4</sup> Department of Animal and Dairy Science, University of Georgia, Athens, GA 30602, USA
- <sup>5</sup> Unità di Ricerca di Genetica e Biotecnologie, Agris Sardegna, 07100 Sassari, Italy
- <sup>6</sup> Istituto di Biologia e Biotecnologia Agraria (IBBA), Consiglio Nazionale delle Ricerche, 20133 Milano, Italy
- \* Correspondence: arianna.bionda@unimi.it
- † These authors contributed equally to this work.

**Simple Summary:** The inbreeding coefficient is relevant for managing livestock and safeguarding biodiversity, especially in small populations. Small ruminant breeders mainly rely on pedigree information, but genomics is increasingly gaining ground as a tool to face possible pedigree inaccuracies. This study investigates pedigree-based ( $F_{PED}$ ) and genomic ( $F_{ROH}$  and  $F_{GRM}$ ) inbreeding in a representative number of Italian sheep and goat populations. Indeed, even though it has been widely studied in cattle, there is still little knowledge about the relationship between these coefficients in small ruminants, which are characterized by a different population structure, often with unconnected farms. Mean inbreeding values were low,  $F_{ROH}$  being the highest, with breed differences due to different managements; the correlation between  $F_{PED}$  and  $F_{ROH}$  was the strongest and directly related to pedigree depth. Moreover, we estimated  $F_{PED}$  from  $F_{ROH}$  using a linear regression model. Since massive genotyping is not affordable to small ruminant breeders, it is important to understand the distinction and relationship between differently calculated inbreeding coefficients, also in view of the introduction of genomic enhanced breeding values. Our study highlights the importance of accurate pedigree information and, especially if not obtainable, of genotyping animals. Correct data contribute to mitigate inbreeding depression and loss of genetic variability.

Abstract: The inbreeding coefficient is an important parameter for livestock management. Small ruminant breeders and associations mainly rely on pedigree information, but genomic tools are gaining relevance, overcoming possible pedigree inconsistencies. This study investigates the relationship between pedigree-based and genomic inbreeding in two goat and four sheep dairy breeds. Pedigree and genomic data (medium-density SNPchip) were obtained for 3107 goats and 2511 sheep. We estimated pedigree depth (number of fully traced generations, FullGen) and inbreeding (FPED), as well as two genomic inbreeding indexes, using runs of the homozygosity (FROH) and genomic relationship matrix (F<sub>GRM</sub>). The correlation between the inbreeding coefficients was assessed. A linear regression model (LRM) was fitted for estimating FPED from FROH. After quality control on genomic data, we retained 5085 animals. Mean inbreeding values were low, with higher  $F_{ROH}$  than  $F_{PED}$  and F<sub>GRM</sub>. Breed differences can partially depend on different managements. The correlation between FPED and FROH was the highest and directly related to pedigree depth. The best LRM was chosen for FullGen  $\geq$ 4 and  $\geq$ 6 for goats and sheep, respectively; after excluding animals with extreme residuals, a new refined regression equation was calculated. Since massive genotyping is not affordable to small ruminant breeders, it is important to understand the distinction and relationship between differently calculated inbreeding coefficients, also in view of the introduction of genomic enhanced breeding values. Our study highlights the importance of accurate pedigree information and, especially if not obtainable, of calculating genomic-based inbreeding coefficients. A better estimation of animals' relatedness contributes to improve animal breeding and conservation.



Citation: Cortellari, M.; Negro, A.; Bionda, A.; Grande, S.; Cesarani, A.; Carta, A.; Macciotta, N.; Biffani, S.; Crepaldi, P. Using Pedigree and Genomic Data toward Better Management of Inbreeding in Italian Dairy Sheep and Goat Breeds. *Animals* 2022, *12*, 2828. https:// doi.org/10.3390/ani12202828

Academic Editors: Dimitrios Loukovitis and Emilia Bagnicka

Received: 4 August 2022 Accepted: 17 October 2022 Published: 18 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). Keywords: small ruminant; inbreeding; SNP; pedigree; genomic inbreeding; runs of homozygosis;  $F_{ROH}$ 

#### 1. Introduction

The accurate estimate of the inbreeding value of an individual could be of extreme importance for all the domesticated species, both for commercial and conservation purposes [1–3]. With regards to small ruminants, pedigree information continues to be the most widely used tool to compute inbreeding coefficients, even though several authors have pointed out that pedigree-based inbreeding values can be severely underestimated due to the high level of inconsistencies in these species [4,5]. Indeed, in goats and sheep, the correct parentage can be difficult to assess due to several factors, such as the high missing information rate, the limited use of artificial insemination, the creation of mating groups with the simultaneous presence of more males [6] and the extensive farming system [7]. According to Legarra et al. (2014) [8], the unknown fatherhood rate reaches 50% and 20% in Latxa and Manech/Basco-Béarnaise sheep breeds, respectively.

Today, genomic data obtained by high-throughput SNP genotyping represent a new instrument that should be integrated with the traditional pedigree information to more precisely estimate the relationship between individuals [9]. Moreover, genomic information allows for a more accurate inbreeding estimation by overcoming pedigree inconsistencies and by accounting for the Mendelian sampling. This allows to better decipher the genetic variation between the individuals, which is of great interest primarily in small and endangered populations, and to eventually estimate unbiased genetic values [10,11].

There are different ways to calculate inbreeding using SNP chip data, and those based on runs of homozygosity (ROH) and the genomic relationship matrix (GRM) are among the most used [12,13]. ROH are segments of homozygous genotypes identified for a single individual and composed of a series of identical haplotypes. ROH provide information about a subject's level of autozygosity, age of the inbreeding events and origin, but most importantly, they can be used to assess a reliable inbreeding coefficient ( $F_{ROH}$ ) [4]. Indeed, recombination events can disrupt long stretches of DNA over generations (or meiosis), allowing to estimate the age of the inbreeding events: Short ROH derive from more distant ancestors, whereas long ROH suggest that the autozygosity of an individual comes from more recent meiosis events, leading to the presence of a recent common ancestor [14].

Even if SNP genotyping has been routinely implemented in dairy cattle [15], its costs are still too high compared to the value of a single sheep or goat; thus, small ruminant breeders and breeder associations cannot afford to genotype a large number of animals. This supports the need to find a way to integrate traditional pedigree data with the genomic data and eventually estimate reliable inbreeding coefficients. Some previous studies have already investigated the genomic variability of Italian sheep and goat breeds [16–19]. However, their results were based on a limited number of animals and/or did not elucidate the relationship between traditional pedigree-based and innovative genomic-based parameters. For this reason, our work studies the relationship between  $F_{PED}$  and pedigree depth and the genomic inbreeding  $F_{ROH}$  and  $F_{GRM}$  in six Italian small ruminant populations focusing, in particular, on  $F_{PED}$  and  $F_{ROH}$ .

#### 2. Materials and Methods

#### 2.1. Breeds

For this work, we selected a total of six breeds, namely four local dairy sheep breeds— Comisana, Delle Langhe, Massese and Sarda—and two cosmopolite dairy goat breeds— Camosciata delle Alpi and Saanen. All the breeds are mainly used for cheese production, being both representative of different production systems and geographical areas. The official herd books of these breeds are managed by the Italian Sheep and Goat Breeders Association (Rome, Italy), hereinafter referred to as Asso.Na.Pa. In all the breeds, male selection candidates are DNA-tested for parentage assignment. Artificial insemination (AI) is used mainly in Camosciata delle Alpi and Saanen breeds, where from 10 to 15% of annual births are from semen of foreign AI bucks. All breeds are officially milk-recorded, and a routine genetic evaluation for milk yields is in place. The selection scheme for Comisana and Massese sheep breeds is based on two different closed nucleus flocks reared at the Genetic Center of Asciano (Siena, Italy); instead, the Sarda sheep is managed using a pyramidal scheme, with a nucleus flock at the apex [20]. Finally, selection in Delle Langhe sheep aims to improve milk production through the estimate of the breeding value for milk yield (expressed in kg); a software is available for breeders to optimize genetic breeding values and manage mating to reduce offspring inbreeding.

#### 2.2. Datasets and Quality Control

Pedigree records and genomic data were obtained from the 'Conservation, Health and Efficiency Empowerment of Small Ruminant' (CHEESR) repository project, a database created by Asso.Na.Pa. in the framework of the Italian National Rural Development Plan (PSRN)—sub-measure 10.2 project and supported by the European Agricultural Fund for Rural Development (EAFRD). In total, we analyzed 3107 goats (2164 Camosciata delle Alpi and 943 Saanen) and 2511 sheep (1498 Sarda, 534 Comisana, 375 Massese and 104 Delle Langhe). All animals were genotyped with the medium density chip, Ovine SNP50 BeadChip for ewes and Goat SNP50 BeadChip for goats. A quality control analysis was performed on both species fitting the same parameters and using PLINK 1.9 software [21]: A missing genotype frequency  $\geq 0.05$  and an individual call rate below 0.95 were the thresholds to exclude SNP or animals, respectively. In addition, all the unplaced markers (according to the assembly versions ARS1 for goat and OAR4 for sheep) and those located on the X chromosome were excluded in accordance with the majority of the Literature and given the small number of SNPs (about 3% of the total).

# 2.3. Pedigree Analysis

All the pedigree analyses were carried out using the *optiSel* package of the R software [22]. The function 'summary.Pedig' was used to calculate the pedigree-based inbreeding (F<sub>PED</sub>), the number of fully traced generations (i.e., whose ancestors are all known, FullGen), the number of maximum generation (MaxGen), the number of equivalent complete generations (EquiGen), the percentage of animals with a complete first generation and the effective size for each sheep or goat breed according to Wright's formula [23].

#### 2.4. Genomic Inbreeding Calculation

We used the function '-homozyg' of PLINK 1.9 for the calculation of the ROH applying the following criteria: (i) a sliding window of 20 SNPs, (ii) no heterozygous genotype allowed, (iii) no more than two missing genotypes, (iv) a minimum number of SNPs in a ROH equal to 20, (v) a minimum ROH length of 1 Mb, (vi) a minimum SNP density of 1 SNP per 500 kb and (vii) a maximum gap of 500 kb between two consecutive homozygous SNPs. In order to minimize false positives discovery within regions of low marker density, rather stringent criteria were selected. ROH distribution was characterized for each breed by estimating the number of individuals without ROH, the mean number of ROH per individual, the mean total length of ROH per individual, the mean length of a ROH per individual and the genomic inbreeding coefficient ( $F_{ROH}$ ) in each individual. The  $F_{ROH}$ coefficients were computed using the following equation:

$$F_{\text{ROH,i}} = \frac{L_{\text{ROH,i}}}{L_{\text{AUTO}}}$$
(1)

where  $L_{ROH}$  is the sum of the total length of ROH in individual *i* and  $L_{AUTO}$  is the total length of the autosomes covered by the SNPs [24]. In addition, we categorized all the estimated ROH in five different ROH length classes: from 1 to 2 Mb, from 2 to 4 Mb, from 4 to 8 Mb, from 8 to 16 Mb and over 16 Mb. ROH length can represent the starting point

for the evaluation of how many generations ago an inbreeding event occurred: Indeed, recent ROH segments are longer because of the less probability of being interrupted by recombination events; on the other hand, more ancient ones are shorter. In particular,  $F_{ROH}$  are expected to correspond to the ancestral population dating 50 (1 < ROH < 2), 20 (2 < ROH < 4), 12.5 (4 < ROH < 8), 6 (8 < ROH < 16) and 3 (ROH > 16) generations ago [25].

The computation of the GRM and the consequent  $F_{GRM}$  values was performed through the use of the '-ibc' parameter implemented in the GCTA v1.93.3 software [26]. Whereas no minor allele frequency (MAF) threshold was applied to allow a better estimation of ROH [27], before the estimation of the  $F_{GRM}$  we excluded the SNPs with a MAF < 0.05, as this parameter is strongly influenced by the frequencies of rare alleles [28].

# 2.5. $F_{PED}$ - $F_{ROH}$ Estimate

Using the functions 'cor' and 'lm' implemented in the R base package, we calculated the correlation coefficient between  $F_{ROH}$ ,  $F_{GRM}$  and  $F_{PED}$ , and we fitted a linear model to estimate the most probable  $F_{PED}$  value from genomic data ( $F_{ROH}$ ). This calculation was performed on all the individuals of both species together considering increasing FullGen classes (i.e., increasing number of known ancestors).

Then, for each species, we chose the linear model of the FullGen class that maintained at least 600 individuals of the species of interest and, at the same time, had both the highest correlation coefficient between  $F_{PED}$  and  $F_{ROH}$  and the highest linear model  $R^2$ . After that, we applied the coefficients of the chosen FullGen model to estimate the  $F_{PED}$  value only on the individuals of the correspondent species with the selected minimum FullGen. We calculated the standardized residuals (Z-score of the differences between the predicted and observed  $F_{PED}$  values) and excluded all the animals with a standardized residual falling under the 1st or over the 3rd quartile. Lastly, we recalculated the linear model using only the retained subjects and re-estimated their  $F_{PED}$  values with the coefficient of this last refined model.

#### 3. Results

#### 3.1. Dataset Creation

We applied the same thresholds and filtering parameters to both sheep and goat datasets. After the quality control, we retained 3086 (2147 Camosciata delle Alpi and 939 Saanen) individuals and 51,097 markers for goats, and 2484 (1480 Sarda, 529 Comisana, 371 Massese and 104 Delle Langhe) individuals and 46,723 markers for sheep. All the 485 individuals for which we could not calculate  $F_{GRM}$ ,  $F_{ROH}$  or  $F_{PED}$  were excluded. Therefore, the final dataset used for estimating the correlation coefficients and the linear model included 5085 animals (3028 goats and 2057 sheep) (Table 1). The results of pedigree structure of each of the six breeds considered are reported in Supplementary Table S1.

**Table 1.** Number of subjects in the final dataset for each species and breed, relative mean value of pedigree, ROH and GRM-based inbreeding coefficients ( $F_{PED}$ ,  $F_{ROH}$  and  $F_{GRM}$ ) and their correlation coefficients.

Species/Breed	N. Subjects	Mean F <sub>PED</sub>	Mean F <sub>ROH</sub>	Mean F <sub>GRM</sub>	F <sub>PED</sub> -F <sub>ROH</sub> Correlation Coefficient (p-Value)	F <sub>PED</sub> -F <sub>GRM</sub> Correlation Coefficient (p-Value)	F <sub>ROH</sub> -F <sub>GRM</sub> Correlation Coefficient (p-Value)
Goats	3028	$0.017\pm0.039$	$0.059\pm0.032$	$0.012\pm0.046$	0.278 (<0.0001)	0.272 (<0.0001)	0.604 (<0.0001)
Camosciata delle Alpi	2093	$0.016\pm0.038$	$0.057\pm0.033$	$0.013\pm0.043$	0.260 (<0.0001)	0.228 (<0.0001)	0.671 (<0.0001)
Saanen	935	$0.020\pm0.042$	$0.062\pm0.030$	$0.010\pm0.051$	0.312 (<0.0001)	0.351 (<0.0001)	0.492 (<0.0001)
Sheep	2057	$0.062\pm0.076$	$0.097 \pm 0.063$	$0.006\pm0.064$	0.817 (<0.0001)	0.365 (<0.0001)	0.477 (<0.0001)
Sarda	1053	$0.093 \pm 0.095$	$0.135\pm0.065$	$0.023 \pm 0.073$	0.804 (<0.0001)	0.346 (<0.0001)	0.489 (<0.0001)
Delle Langhe	104	$0.060\pm0.037$	$0.099\pm0.038$	$-0.025 \pm 0.127$	0.436 (<0.0001)	0.225 (0.022)	0.354 (<0.0001)
Comisaña	529	$0.018\pm0.010$	$0.044\pm0.016$	$-0.008 \pm 0.021$	0.379 (<0.0001)	-0.130(0.003)	0.229 (<0.0001)
Massese	371	$0.039\pm0.010$	$0.067\pm0.017$	$-0.013 \pm 0.025$	0.318 (<0.0001)	-0.070(0.181)	0.224 (<0.0001)

For each sheep and goat breed, we calculated  $F_{PED}$ ,  $F_{GRM}$  and  $F_{ROH}$  for different ROH length classes. An  $F_{PED}$  value equal to 0 was observed in less than 1% of sheep and 37% of goats. As shown in Figure 1a,b, the mean  $F_{ROH}$  values were higher than the other inbreeding coefficients in both the species. Instead,  $F_{PED}$  values were higher than  $F_{GRM}$  in sheep, whereas the two were similar in goats. It is worth noticing that  $F_{GRM}$  showed some negative values in all the breeds and a particularly spread distribution in the Delle Langhe breed.





Regarding the  $F_{ROH}$  distribution per ROH length class (Figure 1c), the two goat breeds had approximately the same distribution for all the  $F_{ROH}$  classes apart from the one derived from ROH > 16 Mb (suggesting recent inbreeding), which was higher in Saanen than in Camosciata delle Alpi. In the sheep species, Sarda and Delle Langhe breeds presented the highest  $F_{ROH}$  value and proportion of long ROH classes. In any case, all inbreeding coefficients were relatively low in both species, with the highest values in the Sarda sheep.

The correlation coefficients between  $F_{GRM}$ ,  $F_{ROH}$  and  $F_{PED}$  were much higher in sheep than in goats, and, in particular, in the Sarda breed. In addition, the correlation between  $F_{PED}$  and  $F_{ROH}$  was higher than between  $F_{PED}$  and  $F_{GRM}$ ; therefore, we decided to use a linear regression model to estimate  $F_{PED}$  from  $F_{ROH}$  (Table 1).

The correlation coefficients between  $F_{PED}$  and  $F_{ROH}$  and the  $R^2$  of the linear model, reported in Table 2, showed a progressive increment as the minimum FullGen increased until we reached FullGen > 6. All the correlations were statistically significant (p < 0.0001). Based on these results and the assumptions explained in the material and methods section

(highest correlation value and  $R^2$ , and at least 600 subjects), we identified the model calculated for FullGen  $\ge 4$  as the best one for the goats, and the one calculated on Fullgen  $\ge 6$  as the best for the sheep.

**Table 2.** Number of goats and sheep,  $F_{ROH}$ - $F_{PED}$  correlation coefficients and *p*-value, and linear regression model (LRM) coefficients and R<sup>2</sup> per minimum class of fully traced generations in the pedigree (FullGen).

Minimum FullGen	N. Animals	N. Goats	N. Sheep	Correlation Coefficient	LRM Intercept	LRM Slope	LRM R <sup>2</sup>
0	5085	3028	2057	0.712	-0.028	0.860	0.507
1	4549	2493	2056	0.725	-0.028	0.888	0.526
2	3911	1877	2034	0.735	-0.027	0.910	0.540
3	3311	1358	1953	0.753	-0.026	0.937	0.567
4	2522	717	1805	0.782	-0.027	0.971	0.611
5	1602	167	1435	0.825	-0.028	1.010	0.681
6	927	18	909	0.849	-0.030	1.056	0.720
7	378	2	376	0.847	-0.029	1.085	0.718
8	107	0	107	0.773	0.008	0.997	0.597

Correlation was significant (*p*-Value < 0.0001) for all the classes of FullGen.

# 3.2.1. Goats

Following the method shown in Figure 2, we applied the coefficients of the linear model calculated on all the animals with FullGen  $\geq 4$  (intercept = -0.03, slope = 0.97) on the 717 goats with FullGen  $\geq 4$  to estimate  $F_{PED}$  values (Figure 3a). Then, we excluded the 358 goats with a standardized residual falling under the first and over the third quartiles, and we calculated the new regression equation on the remaining goats:  $F_{PED} = -0.03 + 0.91 \times F_{ROH}$  ( $R^2 = 0.86$ ,  $p < 2.2 \times 10^{-16}$ ) (Figure 3b).

Lastly, we used this last equation to estimate the definitive  $F_{PED}$  from  $F_{ROH}$  for all the goats with FullGen  $\geq$  4, and for each 0.05 class of estimated definitive  $F_{PED}$ , we calculated the corresponding mean and range of  $F_{ROH}$  (Table 3). All data are reported, but due to the low number of subjects in some of the estimated  $F_{PED}$  classes, only those with a relevant size are further commented. For example, it should be noted that for the class of estimated definitive  $F_{PED}$  0–0.05, the corresponding  $F_{ROH}$  values were higher in Saanen than in Camosciata delle Alpi.

**Table 3.** Number of goats,  $F_{ROH}$  range, mean, standard deviation (SD) and 95% confidence interval (CI) per each class of estimated  $F_{PED}$ .

Breed	Estimated F <sub>PED</sub> Class	N. Subjects	F <sub>ROH</sub> Mean	F <sub>ROH</sub> SD	F <sub>ROH</sub> 95% CI	F <sub>ROH</sub> Range
Goats	0.00-0.05	613	0.054	0.017	0.052-0.055	0.008-0.088
Camosciata delle Alpi	0.00-0.05	453	0.052	0.017	0.050-0.053	0.008-0.088
Saanen	0.00 - 0.05	160	0.058	0.015	0.056-0.061	0.024-0.088
Goats	0.05-0.10	81	0.103	0.013	0.101-0.106	0.088-0.138
Camosciata delle Alpi	0.05–0.10	49	0.102	0.011	0.099–0.105	0.089-0.132
Saanen	0.05-0.10	32	0.106	0.015	0.100-0.111	0.088-0.138
Goats	0.10-0.15	17	0.162	0.014	0.156-0.169	0.143-0.182
Camosciata delle Alpi	0.10-0.15	9	0.157	0.014	0.148-0.166	0.143-0.179
Saanen	0.10-0.15	8	0.168	0.012	0.159–0.176	0.153-0.182



**Figure 2.** Flowchart of the estimation of a linear regression model estimating  $F_{PED}$  from  $F_{ROH}$  for goats and sheep.



**Figure 3.** Linear regression models to estimate  $F_{PED}$  from  $F_{ROH}$ . (a) Linear regression model calculated on all the animals with FullGen  $\geq$  4. (b) Refined model (removing animals falling below the 1st and above the 3rd quartile of the standardized residual distribution) for goats with FullGen  $\geq$  4.

#### 3.2.2. Sheep

For sheep, we selected the coefficients relative to the regression model calculated for animals with a minimum FullGen equal to 6 (intercept = -0.03, slope = 1.06) and applied them on the 927 sheep with a FullGen  $\geq 6$  to calculate the estimated  $F_{PED}$  values (Figure 4a). We standardized the residuals and excluded 463 subjects falling under the first and over the third quartiles. Finally, we calculated the new equation using the 464 retained sheep:  $F_{PED} = -0.03 + 1.02 \times F_{ROH}$  ( $R^2 = 0.97$ ,  $p < 2.2 \times 10^{-16}$ ) (Figure 4b).



**Figure 4.** Linear regression models to estimate  $F_{PED}$  from  $F_{ROH}$ . (a) Linear regression model calculated on all the animals with FullGen  $\geq 6$ . (b) Refined model (removing animals falling below the 1st and above the 3rd quartile of the standardised residual distribution) for sheep with FullGen  $\geq 6$ .

The estimated  $F_{PED}$  intervals and relative  $F_{ROH}$  means and ranges for all the sheep with FullGen  $\geq 6$  are reported in Table 4. There was a significant difference in the breeds' mean  $F_{ROH}$  corresponding to an estimated  $F_{PED}$  of 0.00–0.05 and 0.05–0.10: Comisana had the lowest  $F_{ROH}$  values, followed by Massese, while Delle Langhe and Sarda present the highest values. As shown in Table 4, only Sarda animals were found in the higher  $F_{PED}$ classes (>0.20).

Breed	Estimated F <sub>PED</sub> Class	N. Subjects	F <sub>ROH</sub> Mean	F <sub>ROH</sub> SD	F <sub>ROH</sub> 95% CI	F <sub>ROH</sub> Range
Sheep	0.00-0.05	419	0.055	0.014	0.054-0.057	0.022-0.078
Comisana	0.00-0.05	163	0.045	0.013	0.043-0.047	0.022-0.078
Massese	0.00-0.05	198	0.061	0.010	0.059-0.062	0.029-0.078
Sarda	0.00-0.05	33	0.068	0.009	0.065-0.071	0.041 - 0.078
Delle Langhe	0.00-0.05	25	0.067	0.012	0.062-0.072	0.027-0.078
Sheep	0.05-0.10	231	0.097	0.014	0.095-0.098	0.078-0.127
Comisana	0.05-0.10	2	0.092	0.014	0.072-0.111	0.082-0.101
Massese	0.05-0.10	59	0.090	0.011	0.087-0.093	0.078-0.120
Sarda	0.05-0.10	122	0.099	0.015	0.096-0.101	0.078-0.127
Delle Langhe	0.05-0.10	48	0.100	0.013	0.096-0.103	0.079-0.124
Sheep	0.10-0.15	89	0.150	0.015	0.146-0.153	0.128-0.176
Comisana	0.10-0.15	1	0.162			
Sarda	0.10-0.15	79	0.149	0.014	0.146-0.152	0.128-0.176
Delle Langhe	0.10-0.15	9	0.152	0.018	0.140-0.164	0.128-0.175
Sheep	0.15-0.20	71	0.199	0.013	0.196-0.203	0.177-0.225
Sarda	0.15-0.20	68	0.199	0.013	0.196-0.203	0.177-0.225
Delle Langhe	0.15-0.20	3	0.203	0.022	0.178-0.227	0.179-0.222
Sheep (Sarda)	0.20-0.25	56	0.250	0.015	0.246-0.254	0.225-0.274
Sheep (Sarda)	0.25-0.30	27	0.293	0.014	0.288-0.298	0.275-0.321
Sheep (Sarda)	0.30-0.35	13	0.336	0.013	0.329-0.343	0.324-0.371

**Table 4.** Number of sheep,  $F_{ROH}$  range, mean, standard deviation (SD) and 95% confidence interval (CI) per each class of estimated  $F_{PED}$ .

Only Sarda sheep had estimated  $F_{PED}$  values > 0.20.

# 4. Discussion

The inbreeding coefficient is a fundamental tool for livestock husbandry: It is necessary for increasing the accuracy of the genomic and genetic breeding value of animals and the consequent improving of the breeds [29]; it is one of the main parameters to monitor the selection programs; it is also useful in biodiversity conservation, allowing to identify possible pitfalls in terms of loss of genetic variability; and it helps breeders to choose the best mating schemes to improve animal welfare and to avoid inbreeding depression (which is especially relevant when using the AI) [30]. As previously mentioned, genomic inbreeding coefficients are more reliable than pedigree-based ones, being free from possible registration inaccuracy [4,11]. This is the reason why the implementation of genomic tools at the field level has increased, leading to genomic selection programs not only in cattle but also in other farm animals such as sheep and goats [31–33].

Nevertheless, the genotyping costs, although consistently reduced in recent years, cannot yet be afforded by the breeders of small ruminants. This makes it necessary to understand how to correctly interpret and compare inbreeding coefficients calculated from different data. Therefore, the present study aimed to explore the  $F_{PED}$  and genomic inbreeding coefficients ( $F_{GRM}$  and  $F_{ROH}$ ), as well as their relationship, in a representative number of animals belonging to six differently managed Italian dairy small ruminant breeds.

Our results showed that there were some differences in the variability distribution of inbreeding coefficients, both within and among breeds. As we know from the literature,  $F_{GRM}$  values tend to be more variable with respect to the other inbreeding parameters because they are highly influenced by the rare allele variants frequencies [12]; moreover, the particularly spread distribution found in the Delle Langhe breed could also depend on its smaller sample compared to the other breeds. Regarding the wide range of  $F_{PED}$  values identified by the analysis, we know that goat pedigrees are less complete compared to other species, similar to a lot of animals (almost 40%) with  $F_{PED} = 0$ , often due to lack of genealogic information. For the same reason, the correlation between  $F_{ROH}$  and  $F_{PED}$  coefficients is lower in goats. Regarding the percentage of  $F_{ROH}$  per ROH length class, the different management of the breeds can account for a portion of the observed variability: For example, the higher level of recent inbreeding found in Saanen compared

with Camosciata delle Alpi could partially depend on the higher use of AI in the former (33% vs. 25% in our samples, higher than the values, 11% and 12%, estimated in the whole populations enrolled to the herd books in the same years). The inbreeding estimated on these subjects was, for the most part, low. The lack of very inbred animals could be considered a limitation of the present study, but it should also be considered that this can depend either on good management of the population or from the low level of connection of the Italian small ruminant farms, with a minimal exchange of animals except when the breeders make use of the AI or receive animals from a genetic center. The diffusion of AI is still limited in both species; however, it is rising in goats due to the commercialization of foreign, mainly French, semen. Avoiding an excessive increase of inbreeding is a main goal of the breeders and their associations to prevent inbreeding depression, especially in small-sized populations, whose survival would be in danger if a reduction of the number of the breeding animals occurred. If the genomic inbreeding could be estimated for all the animals of a farm, or at least a representative number, it would be possible to warn the breeders who are at risk of inbreeding-related issues.

The comparison of inbreeding coefficients with those reported in literature is hampered by the several different parameters that can be defined to estimate them. The average  $F_{PED}$  in a population varies based on the considered pedigree depth (i.e., number of generations tracked back), the use of unknown parent groups and the involved computational algorithm [34], some of which can estimate non-zero inbreeding coefficients also for animals with unknown parents [35].  $F_{ROH}$  can be estimated starting from sliding (as in the present study) or consecutive (not overlapping regions) ROH, both varying according to the parameters used to define the regions. In the literature, results about this parameter have also considerably changed within the same breed. One example is the Sarda dairy sheep: In this study, the average  $F_{ROH}$  for this breed was 0.135 using sliding regions, whereas Cesarani et al. (2019) [32] reported a value of 0.059 using consecutive regions with more strict parameters and Mastrangelo et al. (2018) [18] estimated an average  $F_{ROH}$  of 0.041. In the latter study, the authors reported  $F_{ROH}$  values also for Delle Langhe (0.080 vs. 0.099 in this study), Comisana (0.016 vs. 0.044 in this study) and Massese (0.055 vs. 0.067 in this study). The different values between Cesarani et al. (2019) [18] and the present study are mainly due to the different sample size and the minimum number of SNP to call a ROH, which was 30 in the previous study and 20 in the present one. Finally, F<sub>GRM</sub> values change according to the method used to create the genomic relationship matrix [28,36,37] in which different scaling and weighting factors are used, and according to the quality control made on SNPs data, in particular allele frequency. In fact, some studies have suggested the use of fixed allele frequency (0.5) when building the GRM. According to Bjelland et al. (2013) [38], this kind of matrix is a homozygosity measure adjusted to conform to the distribution of the pedigree inbreeding and, therefore, the correlations between  $F_{PED}$  and  $F_{GRM}$  estimated with this fixed frequency are extremely high.

The analysis of the relationship among inbreeding coefficients showed that the correlation of  $F_{PED}$  was generally higher with  $F_{ROH}$  (0.30 in goats and 0.82 in sheep) than  $F_{GRM}$  (0.27 and 0.35, respectively), especially when the pedigrees were more complete. The results of our analyses are consistent with those previously reported in other studies [12,39]. Of interest is the negative correlation observed between  $F_{PED}$  and  $F_{GRM}$  in Comisana and Massese breeds and that already reported in the literature for other species [40]. Pedigree depth was directly related to the correlation between  $F_{ROH}$  and  $F_{PED}$ ; this means that in order to estimate  $F_{PED}$  accurately, between four and six full generations should be known. It would be important to bring this information to the attention of the breeders, who should especially genotype those animals with a small number of known ancestors. We used a simple regression model that aims to estimate a more accurate pedigree-based inbreeding coefficient on the basis of  $F_{ROH}$  and can be useful for the breeders to compare the inbreeding of genotyped and non-genotyped animals. This regression model could obviously be improved using a larger sample size per breed and including highly related animals. We noticed that goats had shallower pedigrees and a higher number of inbreeding coefficients

equal to zero. For this reason, since the two species share several characteristics, such as similar management and production aptitude (milk and cheese) and they are often raised together, the initial model included both species. The results of the model also highlighted that different breeds had different mean  $F_{ROH}$  per estimated  $F_{PED}$  class. This aspect should be considered when implementing the genomic inbreeding in the animal evaluation, for example, through the estimation of genetic breeding values. Indeed, previously published works have attempted to propose a new model based on ROH for the computation of genomic breeding values in cattle ( $G_{ROH}$ ), finding quite high levels of accuracy [41]. This could be a very useful solution in small ruminant populations in which we have less accurate pedigree information and a lower level of connection between farms.

#### 5. Conclusions

Pedigree information is the most used source of data to compute inbreeding coefficients, especially in small ruminants, due to the still high cost of genotyping. As emerged from our analyses on Italian sheep and goat populations, when all data were included, there were low values of correlation between inbreeding calculated from pedigree  $F_{PED}$  and genomic data  $F_{ROH}$ , particularly in goats (0.30). Indeed,  $F_{PED}$  values equal to zero (almost 40% in goats) occur often due to a lack of genealogic information. Excluding subjects with likely erroneous or too shallow pedigrees while keeping an adequate sample size allowed us to better estimate a reliable correlation between the two inbreeding coefficients. These results underline the importance for small ruminant breeders and associations to improve the accuracy of the collection of pedigree information and, especially when it is not possible, the significant advantages of estimating a genomic inbreeding coefficient. The correctness of the data, indeed, allows to improve small ruminant breeds in different ways: having more reliable EBV and creating better mating schemes in order to improve welfare, production and health-related traits; minimizing the risk of inbreeding depression; and better understanding the effect of their management on inbreeding trends. On the contrary, inaccurate information might limit the breed's genetic progress and lead to the reduction of the genetic variability, which is essential to safeguard livestock biodiversity and to face future climate challenges.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ani12202828/s1. Table S1: Breed pedigree information: maximum number of fully (FullGen), total (MaxGen), and equivalent complete (EquiGen) traced generations, % of subjects with a complete first generation and with MaxGen equal to 1, and effective population size (N<sub>e</sub>).

Author Contributions: Conceptualization, P.C., M.C., A.B. and A.N.; software, M.C., A.C. (Alberto Cesarani) and S.B.; formal analysis, A.N., M.C., A.B., S.B. and A.C. (Alberto Cesarani); resources, S.G., P.C. and A.C. (Antonello Carta); data curation, A.N., M.C. and A.B.; writing—original draft preparation, M.C., A.B., A.N. and P.C.; writing—review and editing, M.C., A.N., A.B., S.G., A.C. (Alberto Cesarani), A.C. (Antonello Carta), N.M., S.B. and P.C.; supervision, P.C., A.C. (Alberto Cesarani) and S.B.; funding acquisition, S.G. and P.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** The APC was funded by the Dipartimento di Scienze Agrarie e Ambientali—Produzione, Territorio, Agroenergia (Università degli Studi di Milano).

**Institutional Review Board Statement:** Datasets were obtained from pre-existing data based on routine animal recording procedures; moreover, DNA sampling for all individuals was conducted using nasal swabs and no invasive procedures were applied. Thus, in accordance to the 2010/63/EU guide and the adoption of the Law D.L. 04/03/2014, n.26 by the Italian Government, an ethical approval was not required for our study.

Informed Consent Statement: No informed consent was required since no personal data were included.

**Data Availability Statement:** The data presented in this study are not publicly available and are part of the official herd book information managed by the Associazione Nazionale della Pastorizia (Asso.Na.Pa.). For further information, please contact S.G. (direzione@assonapa.it).

Acknowledgments: The authors acknowledge the Italian Sheep and Goat Breeders Association of (Asso.Na.Pa.).

Conflicts of Interest: The authors declare no conflict of interest.

# References

- 1. Panetto, J.; Machado, M.A.; da Silva, M.V.G.; Barbosa, R.S.; dos Santos, G.G.; Leite, R.D.M.; Peixoto, M.G.C. Parentage assignment using SNP markers, inbreeding and population size for the Brazilian Red Sindhi cattle. *Livest. Sci.* 2017, 204, 33–38. [CrossRef]
- Schiavo, G.; Bovo, S.; Bertolini, F.; Tinarelli, S.; Dall'Olio, S.; Costa, L.N.; Gallo, M.; Fontanesi, L. Comparative evaluation of genomic inbreeding parameters in seven commercial and autochthonous pig breeds. *Animal* 2020, 14, 910–920. [CrossRef] [PubMed]
- Polak, G.; Gurgul, A.; Jasielczuk, I.; Szmatoła, T.; Krupiński, J.; Bugno-Poniewierska, M. Suitability of Pedigree Information and Genomic Methods for Analyzing Inbreeding of Polish Cold-Blooded Horses Covered by Conservation Programs. *Genes* 2021, 12, 429. [CrossRef]
- 4. Rodríguez-Ramilo, S.; Elsen, J.; Legarra, A. Inbreeding and effective population size in French dairy sheep: Comparison between genomic and pedigree estimates. *J. Dairy Sci.* 2019, *102*, 4227–4237. [CrossRef]
- Wang, Z.; Zhou, B.; Zhang, T.; Yan, X.; Yu, Y.; Li, J.; Mei, B.; Wang, Z.; Zhang, Y.; Wang, R.; et al. Assessing Genetic Diversity and Estimating the Inbreeding Effect on Economic Traits of Inner Mongolia White Cashmere Goats Through Pedigree Analysis. *Front. Vet. Sci.* 2021, *8*, 665872. [CrossRef] [PubMed]
- Hayes, B.J.; Goddard, M. Technical note: Prediction of breeding values using marker-derived relationship matrices. J. Anim. Sci. 2008, 86, 2089–2092. [CrossRef]
- Paiva, R.D.M.; de Sousa, J.E.R.; Ferreira, J.; Cunha, E.E.; de Paiva, M.P.S.L.M.; Gouveia, A.M.G.; Facó, O. Population structure and effect of inbreeding on milk yield of Saanen goats in Brazilian production systems. *Small Rumin. Res.* 2020, 192, 106194. [CrossRef]
- Legarra, A.; Baloche, G.; Barillet, F.; Astruc, J.; Soulas, C.; Aguerre, X.; Arrese, F.; Mintegi, L.; Lasarte, M.; Maeztu, F.; et al. Withinand across-breed genomic predictions and genomic relationships for Western Pyrenees dairy sheep breeds Latxa, Manech, and Basco-Béarnaise. J. Dairy Sci. 2014, 97, 3200–3212. [CrossRef]
- 9. Villanueva, B.; Fernández, A.; Saura, M.; Caballero, A.; Fernández, J.; Morales-González, E.; Toro, M.A.; Pong-Wong, R. The value of genomic relationship matrices to estimate levels of inbreeding. *Genet. Sel. Evol.* **2021**, *53*, 42. [CrossRef]
- Andonov, S.; Lourenco, D.; Fragomeni, B.; Masuda, Y.; Pocrnic, I.; Tsuruta, S.; Misztal, I. Accuracy of breeding values in small genotyped populations using different sources of external information—A simulation study. *J. Dairy Sci.* 2017, 100, 395–401. [CrossRef]
- Cortes-Hernández, J.; García-Ruiz, A.; Vásquez-Peláez, C.G.; Ruiz-Lopez, F.D.J. Correlation of Genomic and Pedigree Inbreeding Coefficients in Small Cattle Populations. *Animals* 2021, 11, 3234. [CrossRef] [PubMed]
- 12. Zhang, Q.; Calus, M.P.; Guldbrandtsen, B.; Lund, M.S.; Sahana, G. Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. *BMC Genet*. **2015**, *16*, 88. [CrossRef] [PubMed]
- 13. Caballero, A.; Villanueva, B.; Druet, T. On the estimation of inbreeding depression using different measures of inbreeding from molecular markers. *Evol. Appl.* **2021**, *14*, 416–428. [CrossRef]
- 14. Ceballos, F.C.; Hazelhurst, S.; Ramsay, M. Assessing runs of Homozygosity: A comparison of SNP Array and whole genome sequence low coverage data. *BMC Genom.* **2018**, *19*, 106. [CrossRef] [PubMed]
- 15. Weller, J.; Ezra, E.; Ron, M. Invited review: A perspective on the future of genomic selection in dairy cattle. *J. Dairy Sci.* 2017, 100, 8633–8644. [CrossRef] [PubMed]
- Ciani, E.; Crepaldi, P.; Nicoloso, L.; Lasagna, E.; Sarti, F.M.; Moioli, B.; Napolitano, F.; Carta, A.; Usai, G.; D'Andrea, M.; et al. Genome-wide analysis of Italian sheep diversity reveals a strong geographic pattern and cryptic relationships between breeds. *Anim. Genet.* 2014, 45, 256–266. [CrossRef]
- Nicoloso, L.; The Italian Goat Consortium; Bomba, L.; Colli, L.; Negrini, R.; Milanesi, M.; Mazza, R.; Sechi, T.; Frattini, S.; Talenti, A.; et al. Genetic diversity of Italian goat breeds assessed with a medium-density SNP chip. *Genet. Sel. Evol.* 2015, 47, 62. [CrossRef]
- 18. Mastrangelo, S.; Ciani, E.; Sardina, M.T.; Sottile, G.; Pilla, F.; Portolano, B. Runs of homozygosity reveal genome-wide autozygosity in Italian sheep breeds. *Anim. Genet.* **2018**, *49*, 71–81. [CrossRef]
- 19. Cortellari, M.; Barbato, M.; Talenti, A.; Bionda, A.; Carta, A.; Ciampolini, R.; Ciani, E.; Crisà, A.; Frattini, S.; Lasagna, E.; et al. The climatic and genetic heritage of Italian goat breeds with genomic SNP data. *Sci. Rep.* **2021**, *11*, 10986. [CrossRef]
- 20. Carta, A.; Casu, S.; Salaris, S. Invited review: Current state of genetic improvement in dairy sheep. J. Dairy Sci. 2009, 92, 5814–5833. [CrossRef]
- 21. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* 2015, *4*, 7. [CrossRef] [PubMed]

- 22. Wellmann, R. Optimum contribution selection for animal breeding and conservation: The R package optiSel. *BMC Bioinform*. **2019**, *20*, 25. [CrossRef] [PubMed]
- 23. Wright, S. The genetical structure of populations. Ann. Eugen. 1949, 15, 323–354. [CrossRef] [PubMed]
- McQuillan, R.; Leutenegger, A.-L.; Abdel-Rahman, R.; Franklin, C.S.; Pericic, M.; Barac-Lauc, L.; Smolej-Narancic, N.; Janicijevic, B.; Polasek, O.; Tenesa, A.; et al. Runs of Homozygosity in European Populations. *Am. J. Hum. Genet.* 2008, *83*, 359–372. [CrossRef] [PubMed]
- 25. Howrigan, D.P.; Simonson, M.A.; Keller, M.C. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genom.* **2011**, *12*, 460. [CrossRef] [PubMed]
- Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A Tool for Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* 2011, *88*, 76–82. [CrossRef]
- 27. Meyermans, R.; Gorssen, W.; Buys, N.; Janssens, S. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genom.* **2020**, *21*, 94. [CrossRef]
- 28. VanRaden, P. Efficient Methods to Compute Genomic Predictions. J. Dairy Sci. 2008, 91, 4414–4423. [CrossRef]
- 29. Aguilar, I.; Fernandez, E.N.; Blasco, A.; Ravagnolo, O.; Legarra, A. Effects of ignoring inbreeding in model-based accuracy for BLUP and SSGBLUP. J. Anim. Breed. Genet. 2020, 137, 356–364. [CrossRef]
- Maltecca, C.; Tiezzi, F.; Cole, J.; Baes, C. Symposium review: Exploiting homozygosity in the era of genomics—Selection, inbreeding, and mating programs. J. Dairy Sci. 2020, 103, 5302–5313. [CrossRef]
- Teissier, M.; Larroque, H.; Robert-Granié, C. Weighted single-step genomic BLUP improves accuracy of genomic breeding values for protein content in French dairy goats: A quantitative trait influenced by a major gene. *Genet. Sel. Evol.* 2018, 50, 31. [CrossRef] [PubMed]
- Cesarani, A.; Sechi, T.; Gaspa, G.; Usai, M.G.; Sorbolini, S.; Macciotta, N.P.P.; Carta, A. Investigation of genetic diversity and selection signatures between Sarda and Sardinian Ancestral black, two related sheep breeds with evident morphological differences. *Small Rumin. Res.* 2019, 177, 68–75. [CrossRef]
- Zhu, S.; Guo, T.; Yuan, C.; Liu, J.; Li, J.; Han, M.; Zhao, H.; Wu, Y.; Sun, W.; Wang, X.; et al. Evaluation of Bayesian alphabet and GBLUP based on different marker density for genomic prediction in Alpine Merino sheep. *G3 Genes Genomes Genet.* 2021, 11, jkab206. [CrossRef] [PubMed]
- 34. Colleau, J.-J. An indirect approach to the extensive calculation of relationship coefficients. *Genet. Sel. Evol.* **2002**, *34*, 409. [CrossRef] [PubMed]
- Aguilar, I.; Misztal, I. Technical Note: Recursive Algorithm for Inbreeding Coefficients Assuming Nonzero Inbreeding of Unknown Parents. J. Dairy Sci. 2008, 91, 1669–1672. [CrossRef]
- Amin, N.; Van Duijn, C.M.; Aulchenko, Y.S. A Genomic Background Based Method for Association Analysis in Related Individuals. PLoS ONE 2007, 2, e1274. [CrossRef]
- 37. Yang, J.; Benyamin, B.; McEvoy, B.P.; Gordon, S.; Henders, A.; Nyholt, D.; Madden, P.A.; Heath, A.C.; Martin, N.; Montgomery, G.; et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **2010**, *42*, 565–569. [CrossRef]
- Bjelland, D.; Weigel, K.; Vukasinovic, N.; Nkrumah, J. Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. J. Dairy Sci. 2013, 96, 4697–4706. [CrossRef]
- Alemu, S.W.; Kadri, N.K.; Harland, C.; Faux, P.; Charlier, C.; Caballero, A.; Druet, T. An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity* 2021, *126*, 410–423. [CrossRef] [PubMed]
- 40. Hidalgo, J.; Cesarani, A.; Garcia, A.; Sumreddee, P.; Larios, N.; Mancin, E.; García, J.; Núñez, R.; Ramírez, R. Genetic Background and Inbreeding Depression in Romosinuano Cattle Breed in Mexico. *Animals* **2021**, *11*, 321. [CrossRef] [PubMed]
- Luan, T.; Yu, X.; Dolezal, M.; Bagnato, A.; Meuwissen, T.H. Genomic prediction based on runs of homozygosity. *Genet. Sel. Evol.* 2014, 46, 64. [CrossRef] [PubMed]