



**HAL**  
open science

## Partitioning of the genetic trends of French dairy sheep in Mendelian samplings and long-term contributions

S. Antonios, A. Legarra, R. Pong-Wong, J.M. Astruc, S.T. Rodríguez-Ramilo,  
Z.G. Vitezica

### ► To cite this version:

S. Antonios, A. Legarra, R. Pong-Wong, J.M. Astruc, S.T. Rodríguez-Ramilo, et al.. Partitioning of the genetic trends of French dairy sheep in Mendelian samplings and long-term contributions. *Journal of Dairy Science*, 2023, 106 (9), pp.6275-6287. 10.3168/jds.2022-23009 . hal-04323394

**HAL Id: hal-04323394**

**<https://hal.science/hal-04323394>**

Submitted on 5 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## Partitioning of the genetic trends of French dairy sheep in Mendelian samplings and long-term contributions

S. Antonios,<sup>1\*</sup> A. Legarra,<sup>1†</sup> R. Pong-Wong,<sup>2</sup> J. M. Astruc,<sup>3</sup> S. T. Rodríguez-Ramilo,<sup>1</sup> and Z. G. Vitezica<sup>1</sup>

<sup>1</sup>GenPhySE, INPT, INRAE, ENVT, F-31326, Castanet Tolosan, France

<sup>2</sup>The Roslin Institute and R(D)SVS, University of Edinburgh, Easter Bush Campus, Edinburgh, EH25 9RG, United Kingdom

<sup>3</sup>Institut de l'Élevage, 149 rue de Bercy, F-75595 Paris, France

### ABSTRACT

The genetic trend of milk yield for 4 French dairy sheep breeds (Lacaune, Basco-Béarnaise, Manech Tête Noire, and Manech Tête Rousse) was partitioned in Mendelian sampling trends by categories of animals defined by sex and by selection pathways. Five categories were defined, as follows: (1) artificial insemination (AI) males (after progeny testing), (2) males discarded after progeny testing, (3) natural mating males, (4) dams of males, and (5) dams of females. Dams of males and AI males were the most important sources of genetic progress, as observed in the decomposition in Mendelian sampling trends. The yearly contributions were more erratic for AI males than for dams of males, as AI males are averaged across a smaller number of individuals. Natural mating males and discarded males did not contribute to the trend in terms of Mendelian sampling, as their estimated Mendelian sampling term is either null (natural mating males) or negative (discarded males). Overall, in terms of Mendelian sampling, females contributed more than males to the total genetic gain, and we interpret that this is because females constitute a larger pool of genetic diversity. In addition, we computed long-term contributions from each individual to the following pseudo-generations (one pseudo-generation spanning 4 years). With this information, we studied the selection decisions (selected or not selected) for females, and the contributions to the following generations. Mendelian sampling was more important than parent average to determine the selection of individuals and their long-term contributions. Long-term contributions were greater for AI males (with larger progeny sizes than females) and in Basco-Béarnaise than in Lacaune (with the latter being a larger population).

**Key words:** genetic gain, genetic trend, Mendelian sampling, long-term contributions, selection scheme

### INTRODUCTION

In animal genetic improvement, selection is based on EBV of candidates for selection. The methodological focus on EBV prediction may hide the role of Mendelian sampling in creating genetic progress. The breeding value of an individual can be expressed as the average of its parental breeding values plus a Mendelian sampling term. As the Mendelian sampling is the unique portion of genetic variation that the individual brings into the population (Woolliams, 2007), this term captures the “originality” of the individual with respect to its parents. In fact, the Mendelian sampling represents the deviation arising from recombination and segregation of parental chromosomes. The breeding value of an individual can be decomposed as a sum of Mendelian samplings of all its ancestors. Thus, the genetic progress can be understood as the selection of positive ancestors’ Mendelian samplings to their descendants and also candidates’ own Mendelian samplings (Woolliams and Thompson, 1994).

The long-term genetic contribution of an individual is its proportional contribution to the long-term genetic background of the population (Bijma, 2000). The contribution,  $r_{i(j)}$ , is defined as the expected fraction of genes from ancestor  $i$  (born at time  $t_1$ ) transmitted to descendant  $j$  (born at a later time,  $t_2$ ; Woolliams et al., 1993; Woolliams, 2007). When  $t_2 - t_1 \rightarrow +\infty$ , this distant future implies that contributions fluctuate at the beginning and stabilize in the long run. The long-term genetic contribution of an individual depends on the genetic superiority for selection of the animal; for example, when selecting based on EBV, the individuals with highest intrageneration EBV will be selected to have more offspring, which increases their long-term genetic contribution (Woolliams et al., 1999).

Received November 9, 2022.

Accepted February 28, 2023.

\*Corresponding author: [simona.antonios@inrae.fr](mailto:simona.antonios@inrae.fr)

†Current address: Council on Dairy Cattle Breeding, Bowie, MD 20716.

The study of Mendelian sampling trends by groups and categories of animals allows understanding of selection decisions (García-Cortés et al., 2008), whereas analyzing long-term contributions of ancestors to descendants allows understanding of the gene flow across generations and by categories of animals (Woolliams et al., 1999). For a better understanding of the selection schemes and their implementation, the dynamics of the contributions, and the sources of “originalities,” we used a retrospective analysis of Mendelian sampling trends and long-term genetic contributions in selected populations.

Partitioning the genetic gain by categories of animals and Mendelian samplings as described above has mainly been evaluated in breeding schemes importing animals from foreign populations. Examples can be found in the literature dealing with cattle (Gorjanc et al., 2011) and pig (Škorput et al., 2015) breeding schemes. However, to date, this method has not been used in closed breeds within country. Our work focuses in 4 French dairy sheep pure breeds: Lacaune (**LAC**) and the Western Pyrenees breeds Basco-Béarnaise (**BB**), Manech Tête Noire (**MTN**), and Manech Tête Rousse (**MTR**), which are local dairy sheep breeds with no introduction from other breeds or countries. Breeding schemes of these breeds were based on performance recording and progeny testing via AI rams (Astruc et al., 2002) until 2016, from which time they have been based on performance recording and genomic selection, with early selection of rams and widespread use of AI. The breeding goals include, depending on the breed, milk yield, fat and protein yields, fat and protein contents, somatic cell score, udder morphology, and other traits (Astruc et al., 2018); however, milk yield is highly selected in all breeds. Although some of the traits (e.g., fat and protein yields) are correlated with this trait, we will focus on milk yield alone, as otherwise the analyses become too complex and difficult to compare across breeds.

The objective of this work was to fine partition genetic trends in Mendelian samplings by categories of animals defined by sex and by selection pathways, and to similarly characterize long-term genetic contributions. We analyzed genetic gain for milk yield in 4 French dairy sheep breeds: LAC, BB, MTN, and MTR.

## MATERIALS AND METHODS

Because no human or animal subjects were used, this analysis did not require approval by an Institutional Animal Care and Use Committee or Institutional Review Board.

### Decomposition of Genetic Trend in Mendelian Samplings and Long-Term Genetic Contributions

The breeding or additive genetic value for an animal  $i$  ( $u_i$ ) can be expressed in terms of its parent (sire and dam) breeding values ( $u_s$  and  $u_d$ , respectively), whose average is called parent average, plus the Mendelian sampling of animal  $i$  ( $\phi_i$ ; Thompson, 1979) as

$$u_i = \frac{1}{2}u_s + \frac{1}{2}u_d + \phi_i. \quad [1]$$

In matrix notation, the vector of breeding values ( $\mathbf{u}$ ) can be written as a linear combination of Mendelian sampling terms and their ancestors as  $\mathbf{u} = \mathbf{T}\boldsymbol{\phi}$  (where vector  $\boldsymbol{\phi}$  also includes the breeding values of the pedigree founders). The triangular matrix  $\mathbf{T}$  describes the flow of genes through pedigree (Woolliams et al., 1999) and shows the relatedness between each individual and his ancestor (García-Cortés et al., 2008). Accordingly,  $\boldsymbol{\phi}$  is the vector of Mendelian sampling terms, assumed distributed as  $\boldsymbol{\phi} \sim N(0, \mathbf{D}\sigma_u^2)$  (García-Cortés et al., 2008), where  $\mathbf{D}$  is a diagonal matrix with Mendelian sampling variances for nonfounders and additive genetic variances for founders. The vector of breeding values is thus distributed as  $\mathbf{u} \sim N(0, \mathbf{A}\sigma_u^2)$ , where  $\mathbf{A} = \mathbf{TDT}'$  is the additive genetic relationship matrix and  $\sigma_u^2$  is the additive genetic variance.

Equation [1] and its matrix notation reveal that breeding values are linear combinations of Mendelian sampling terms. The breeding value of an individual  $i$  born can be expressed using Equation [1] in terms of its parents' breeding values and the Mendelian sampling term. The sire's breeding value can also be written using Equation [1] and so on. Thus, Equation [1] can be used recursively going back through the pedigree, and we can partition each breeding value (for an individual

$i$ ) as  $u_i = \sum_{j=1}^m T_{ij} \phi_j$  ( $m$  total number of animals), where  $u_i$

is decomposed into a sum of  $i$  independent terms or contributions that involve the breeding value of founders and the Mendelian sampling of nonfounders. The coefficient  $T_{ij}$  is the genetic contribution of an ancestor  $j$  to an individual  $i$ , called  $r_{j(i)}$  by Woolliams et al. (1999). Woolliams et al. (1999) defined the long-term genetic contribution [ $r_{j(i)}$ , the same as  $T_{ij}$ ] as the proportion of the genes in individual  $i$  transmitted from  $j$ . Using this definition, for a defined cohort (generation), the average of contributions  $r_{j(i)}$  of each ancestor  $i$  across all individuals (descendants or not)  $j$  belonging to a defined cohort, defines its long-term contribution to that cohort,  $r_i$  (as per Howard et al., 2018), as

$$r_i = \frac{1}{n} \sum_{j=1}^n r_{i(j)}. \quad [2]$$

The sum is computed was over all individuals  $j$  ( $j = 1, \dots, n$ ) born in the defined cohort, and  $n$  was the total number of animals in the defined cohort.

Equation [1] also holds for unbiased predictors, such as the BLUP (Henderson, 1973) of the breeding or additive genetic values ( $\hat{\mathbf{u}}$ ; García-Cortés et al., 2008) as follows:

$$\hat{\mathbf{u}} = \mathbf{T}\hat{\boldsymbol{\phi}}, \quad [3]$$

where  $\hat{\mathbf{u}}$  is a vector that contains the EBV, and the Mendelian samplings can be estimated as  $\hat{\boldsymbol{\phi}} = \mathbf{T}^{-1}\hat{\mathbf{u}}$ . García-Cortés et al. (2008) proposed to define a set of  $k$  partitions such that  $\mathbf{P}_1 + \mathbf{P}_2 + \dots + \mathbf{P}_k = \mathbf{I}$ . These partitions are based on groups or categories of interest; for instance, males and females. From this idea and using the Equation [3], we can write

$$\hat{\mathbf{u}} = \mathbf{T}(\mathbf{P}_1 + \mathbf{P}_2 + \dots + \mathbf{P}_k)\hat{\boldsymbol{\phi}} = \mathbf{TP}_1\hat{\boldsymbol{\phi}} + \mathbf{TP}_2\hat{\boldsymbol{\phi}} + \dots + \mathbf{TP}_k\hat{\boldsymbol{\phi}},$$

where  $\mathbf{P}_i$  (with  $i = 1, \dots, k$ ) is a diagonal matrix containing zeros and ones, and it selects the corresponding columns of the  $\mathbf{T}$  matrix; in other words, the corresponding ancestor effects for a given  $\hat{\mathbf{u}}$  (García-Cortés et al., 2008). Replacing the Mendelian sampling term by  $\mathbf{T}^{-1}\hat{\mathbf{u}}$ , we obtain

$$\hat{\mathbf{u}} = \mathbf{TP}_1\mathbf{T}^{-1}\hat{\mathbf{u}} + \mathbf{TP}_2\mathbf{T}^{-1}\hat{\mathbf{u}} + \dots + \mathbf{TP}_k\mathbf{T}^{-1}\hat{\mathbf{u}};$$

$$\hat{\mathbf{u}} = \hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2 + \dots + \hat{\mathbf{u}}_k.$$

Note that  $\hat{\mathbf{u}}_k$  is a part of the breeding value contributed by the category  $k$ . From these partitions ( $\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_k$ ), the partial genetic response can be obtained for a group or category. For example, the genetic gain can be partitioned using a category defined by sex, where  $\mathbf{P}_m$  and  $\mathbf{P}_f$  are the partition matrices defined for males and females, respectively. The matrix  $\mathbf{P}_m$  (or  $\mathbf{P}_f$ ) is a diagonal matrix with ones in the male (female) positions and zeros in the female (male) positions. The method presented above enables the calculation of partial genetic trends directly from the EBV and allows us to inspect the contributions of each category to the genetic gain. Note that by doing this, we have split, for each individual, the parent average into contributions due to Mendelian sampling of their ancestors, which, in turn, then we assign into categories. This is why García-Cortés et al. (2008) consistently use the term “Mendelian sampling” for this partition.

### Phenotypic and Pedigree Data

This study considered the breeding schemes of the BB, MTN, MTR, and LAC breeds until 2016, at which time genomic selection was introduced. Based on the 2021 official national genetic evaluation (single-step Genomic BLUP). In a single-step Genomic BLUP evaluation, the addition of genomic data to the evaluation can be seen as “more information,” and genomic EBV (**GEbV**) can be seen as the EBV obtained at the same time but with genomic information (Macedo et al., 2020). Thus, it is possible to analyze the selection decisions in an individual and its ancestors based on GEbV. The EBV or GEbV for milk yield (liters per ewe per lactation) of animals born between 1985 and 2016 and their pedigree were used in the analyses (Table 1). By 1985 all breeds had ongoing, routine pedigree and milk yield recordings. The breed with the greatest population size is LAC, followed by MTR, BB, and MTN. Pedigree completeness of rams was very high: as these are elite animals, all AI rams have at least 2 parents and 4 grandparents known. Because 3% of LAC, 11% of MTR, 18% of BB, and 20% of MTN dams have unknown sires, the genetic evaluation model includes unknown parent groups (**UPG**; see Table 1), to take into account the different genetic levels of missing sires (Quaas, 1988). Use of UPG allows for realistic estimates of genetic trends. We did not attempt alternative definitions of UPG. Macedo et al. (2022) showed that an adequate genetic trend requires correct definition of UPG, which was the case here. Because EBV in models with UPG are not estimable functions, in order to obtain meaningful partitions, we shifted the EBV by a constant such that the average EBV of the first cohorts of animals was 0. To estimate animals’ Mendelian sampling, we discounted the solution of the UPG from its progeny; but otherwise, we did not include UPG solutions when presenting contributions or trends.

We decomposed the total genetic gain in 2 different manners: first, into contributions from males and females; second, into contributions from categories defined by different selection pathways. Thus, animals were classified into 5 categories (Table 2): (1) AI Males, the best males kept after progeny testing; (2) Discarded AI Males, the males discarded after progeny testing; (3) Natural Mating (NM) Males; (4) Dams of Males, which are the elite females chosen to generate AI Males; and (5) Dams of Females, those that were not chosen as elite females, because their estimated genetic merit was lower than that of Dams of Males. In dairy sheep selection schemes, there are females (recorded for milk yield and other traits) and males, which can be either AI or NM Males. Artificial insemination is used for spread of genetic improvement

**Table 1.** Number of animals in pedigree per breed

Breed <sup>1</sup>	Number of animals	Number of females	Number of males	Number of dams	Number of sires	Number of UPG <sup>2</sup>
BB	145,047	142,614	2,433	66,077	2,249	32
MTN	100,764	99,328	1,436	43,822	1,353	40
MTR	489,098	480,157	8,941	230,632	8,409	40
LAC	1,474,543	1,452,032	22,511	721,567	17,346	17

<sup>1</sup>BB = Basco-Béarnaise; MTN = Manech Tête Noire; MTR = Manech Tête Rousse; LAC = Lacaune.

<sup>2</sup>UPG = unknown parent groups.

and (until 2016) for selection of elite males by progeny testing. The AI Males produce 76% of the females, with the remaining 24% coming from NM Males (on average of all breeds). In turn, elite females are also selected to give birth to best males. All males are offspring of AI Males and AI-born females; in other words, there is no AI Male son (or grandson) of an NM Male. The categories of females and males without progeny (“Females No Progeny” and “Males No Progeny,” respectively) were present in the pedigree, but they were not analyzed, as their genes do not contribute to the genetic progress. Table 2 shows the number of animals per breed and defined category, as well as the proportion selected in the 4 classical pathways (sire of sire, sire of dam, dam of sire, and dam of dam) in the 4 breeds. Note that these categories describe correctly the pre-2016 scheme even if some (e.g., Discarded AI Males) are no longer valid in the current scheme, which is a genomic one with no formal progeny testing (although eventually daughters’ performances contribute greatly to the EBV of the AI Males).

Finally, to understand some of the results, we computed generation intervals (average age of the ancestor when its selected offspring was born) and family size per category of animals for the BB and LAC breeds. Table 3 shows the generation intervals and the family size per category of animals for BB (as representative of Western Pyrenees breeds) and LAC breeds. Generation intervals for LAC are lower (~3 yr) than the generation intervals for BB (~4 yr). Family size of the LAC AI Males is larger compared with BB. The numbers of

progenies of AI Males that are males and females are greater for LAC than for BB.

### Analyzing Mendelian Samplings

Although pedigree and performance recording started earlier in some breeds (notably LAC), our analyses begin in 1985, when the dairy sheep selection schemes started to extensively use performance and pedigree recording.

The AlphaPart R package (Obšteter et al., 2021), available at <http://CRAN.R-project.org/package=AlphaPart>, which implements the method of García-Cortés et al. (2008), was used for the Mendelian sampling decomposition of EBV and genetic trends. The output of AlphaPart contains for each individual  $i$  the EBV ( $u_i$ ), which is partitioned in 2 different manners. First, the EBV is split into the parent average plus the Mendelian sampling of the individual; in our notation,  $\hat{u}_i = \widehat{PA}_i + \hat{\phi}_i = \frac{1}{2}\hat{u}_{sire(i)} + \frac{1}{2}\hat{u}_{dam(i)} + \hat{\phi}_i$ . Second, the EBV is split into a series of values (EBV\_“category” in the AlphaPart output) per each of the  $k$  categories, such that the sum across categories yields the EBV; in our notation,  $\hat{u}_i = \hat{u}_{(i)1} + \hat{u}_{(i)2} + \dots + \hat{u}_{(i)k}$ . In turn these  $\hat{u}_{(i)1}$ ,  $\hat{u}_{(i)2}$ , ...  $\hat{u}_{(i)k}$  are aggregated across (and not within) individuals into Mendelian sampling trends for each of the  $k$  categories (e.g., Figure 1). Note that the decomposition into trends due to Mendelian samplings sum to the overall genetic trend. In addition, it is of interest to obtain the genetic gain per year per category, in terms of Mendelian samplings. In practice this was carried

**Table 2.** Number of animals and selection proportions per breed and category

Breed <sup>1</sup>	Category							Selection proportion			
	AI Males	Discarded AI Males <sup>2</sup>	Natural Mating Males	Dams of Males	Dams of Females	Females, No Progeny	Males, No Progeny	Sire of sire	Sire of dam	Dam of sire	Dam of dam
BB	633	866	750	1,941	64,136	76,537	184	0.47	0.66	0.26	0.94
MTN	402	507	444	1,183	42,639	55,506	83	0.46	0.65	0.48	0.89
MTR	2,843	2,923	2,643	6,979	223,653	249,525	532	0.34	0.50	0.35	0.90
LAC	8,841	6,136	2,369	17,916	703,651	730,465	5,165	0.06	0.14	0.32	0.91

<sup>1</sup>BB = Basco-Béarnaise; MTN = Manech Tête Noire; MTR = Manech Tête Rousse; LAC = Lacaune.

<sup>2</sup>Discarded AI Males are the males discarded after progeny testing.

**Table 3.** Generation intervals (years) and number of offspring (males and females) per category<sup>1</sup>

Category <sup>2</sup>	BB				LAC			
	GIM	GIF	NPM	NPF	GIM	GIF	NPM	NPF
AI Males	4.3	3.1	8.3	86.2	3.9	3.1	13.8	109.0
Discarded AI Males	3.4	2.6	1.2	22.2	3.9	2.6	1.6	15.2
NM Males	3.2	2.7	1	4.4	3.5	2.9	1	4.5
DM	4.0	4.6	1.2	2.5	3.3	3.7	1.2	2.3
DF	—	4.2	—	1.8	—	3.4	—	1.9

<sup>1</sup>BB = Basco-Béarnaise; LAC = Lacaune. GIM = generation interval of the individual when the male progeny was born; GIF = generation interval of the individual when the female progeny was born; NPM = number of progeny that are males; NPF = number of progeny that are females.

<sup>2</sup>NM = natural mating; DM = dam of males; DF = dam of females.

out after processing the output of AlphaPart: the contribution in terms of Mendelian sampling of year  $t$  in category  $j$  was obtained as the difference between yearly averages:  $MSgain_{j(\text{year}=t)} = \hat{u}_{j(\text{year}=t)} - \hat{u}_{j(\text{year}=t-1)}$ .

### Analyzing Long-Term Genetic Contribution

Long-term genetic contributions ( $r$ ) were computed using self-made software in BB (again, as representative of Western Pyrenees breeds) and LAC breeds. This concept (and value) of long-term genetic contribution represents the contribution of an individual’s Mendelian sampling term to the long-term genetic pool (Woolliams, 2007). Note that the gene pool of the population has contributions from all ancestors and not just the founders.

To determine the value of  $r_i$  for individual  $i$ , a scale of pseudo-generations (hereinafter “generations,” for simplicity) was used to group the individuals depending on their year of birth. In dairy sheep, the generation interval is around 4 years. Base generation ( $G_0$ ) consisted of individuals born from 1985 to 1988. Seven more generations were defined, with the last generation ( $G_7$ ) consisting of individuals born between 2013 and 2016. The contributions  $r_i$  are highly left-skewed, and, for ease of presentation, we use histograms or boxplots of  $\log_{10}(r_i)$  (i.e.,  $-2$  means 0.01).

In addition to  $r_i$ , a selection score ( $x_i$ ) was determined for all individuals (Howard et al., 2018). If the individual had offspring,  $x_i = 1$ , and  $x_i = 0$  otherwise. Note that  $x_i = 0$  implies  $r_i = 0$ , and therefore 3 categories can be established: unselected ( $x_i = 0, r_i = 0$ ), selected but with no contribution to the defined generation ( $x_i = 1, r_i = 0$ ), and selected with contribution to the defined generation ( $x_i = 1, r_i > 0$ ). The case ( $x_i = 1, r_i = 0$ ) occurs when individual  $i$  produces offspring in the next generation but its descendants at some point do not produce further offspring in the following generations; its genetic line goes extinct. Table 4 shows the number of individuals in each category per breed (BB

and LAC) and sex. However, the scores were not used in our further analyses for males, because males that are recorded into the pedigree files are live, breeding males (AI or NM rams) and therefore selected, whereas females are recorded if they have at least 1 lactation and may then be selected or not. Thus, pedigree recording is highly biased in males toward selected animals. For females, 2 further analyses, detailed below, were performed with scores.

The relationships between long-term genetic contributions, estimated Mendelian sampling terms ( $\hat{\phi}$ ), and EBV were also studied. First, we studied the importance of parent average  $\widehat{PA}_i = (EBV_i - \hat{\phi}_i)$  or  $\hat{\phi}_i$  as source of selective advantage. Second, we examined the maintenance of contributions over time for selected individual. We used bivariate regressions to determine whether  $\widehat{PA}_i$  (corrected by year of birth to compare properly across successive years) or  $\hat{\phi}_i$  were involved in the selection of an individual ( $x_i = 1$  for individuals whose offspring were kept) and in the persistence of the contribution over time [which parents made a nonzero contribution ( $r_i > 0$ ) to the population at the last,  $G_7$ , pseudo-generation].

A generalized linear model with a binomial distribution and a logistic link function was fitted to assess the weight placed on  $\widehat{PA}_i$  versus  $\hat{\phi}_i$ . Let the probability of selection for individual  $i$  be  $\mu_i = E[x_i]$ , then if  $f(\cdot)$  is the logistic link function, the following model was fitted:

**Table 4.** Number of females and males in each breed<sup>1</sup> depending on the selection score ( $x_i$ ) and the long-term genetic contribution ( $r_i$ ) to generation 7

Constraint		BB		LAC	
$x_i$	$r_i$	Females	Males	Females	Males
$\geq 0$	$\geq 0$	142,614	2,433	1,452,032	22,511
$> 0$	$\geq 0$	66,077	2,249	721,567	17,346
$> 0$	$> 0$	31,567	1,882	308,514	15,419

<sup>1</sup>BB = Basco-Béarnaise; LAC = Lacaune.

$f^{-1}(\mu_i) = \alpha + \beta yr + \beta_{PA} \widehat{PA}_i + \beta_{\phi} \widehat{\phi}_i + \varepsilon_i$ , where  $\alpha$  is an intercept,  $yr$  is the year of birth, and  $\beta_{PA}$  and  $\beta_{\phi}$  are the respective regression coefficients on  $\widehat{PA}_i$  and  $\widehat{\phi}_i$ . The odds ratios were computed as  $\left[\exp(\widehat{\beta})\right]$ . This analysis was conducted using the “binomial” option in the “glm” package of R.

## RESULTS

Between 1985 and 2016, the total genetic progress for milk yield (estimated from the regression of the EBV on the year of birth) was 4.6 l/yr, 3.3 l/yr, 2.9 l/yr, and 2.7 l/yr for the LAC, MTR, MTN, and BB breeds, respectively. This implies an increment of more than 100 L for all the evaluated breeds between 1985 and 2016. The use of AI differs among breeds. In LAC flocks, AI rate reaches 85% (420 progeny-tested males by year), whereas this rate is 55% (with 30–130 males tested by year) in BB, MTN, and MTR (Astruc et al., 2002). This leads to different selection proportions among categories, as observed in Table 2; most notably, LAC has greater selection pressures in the “sire of” pathways. For all the following results, the genetic progress was stable for all 4 breeds only after approximately 1992, so some fluctuations in the results are apparent in the first years.

Figure 1 shows the decompositions of genetic trends of milk yield, in Mendelian sampling terms, contributed by males and contributed by females for the 4 breeds, in addition to the overall genetic trend. It can be verified that the genetic trend is the sum of the trends of Mendelian samplings contributed by males and females. Looking at the trends of contribution of Mendelian sampling by sex, the highest difference among breeds was between LAC and other breeds, as the contribution of the Mendelian sampling terms of females was greater than the contributions of Mendelian sampling terms of males in the LAC breed across the entire period. In 2016, LAC females and males’ Mendelian sampling contributions were (at the end of the genetic trends in 2016) 101 (69%) and 45 L (31%), respectively. The Western Pyrenees dairy sheep breeds (BB, MTN, and MTR) showed quite similar shapes of trends, where in all cases the Mendelian sampling contributions cumulated around 50% for females and males in 2016.

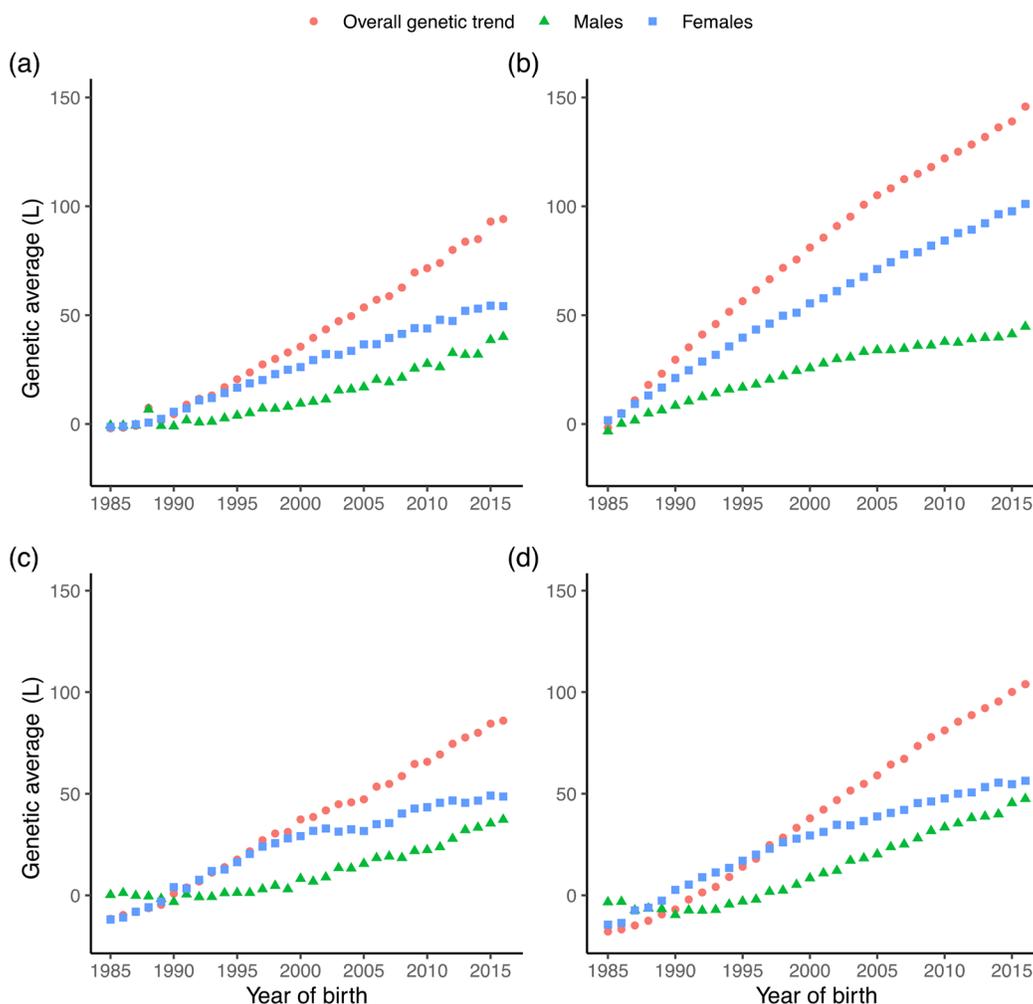
Figure 2 presents the decomposition of genetic trends of milk yield to the contributions of each category previously defined, for each breed. The categories’ contributions to the total genetic gain in the LAC breed differed from those of the Pyrenees dairy sheep breeds. In LAC, the contributions of the AI Males, Dams of Males, and Dams of Females were of the same order.

Looking at the year 2016, AI Males contributed 33% of the total genetic gain, and the females’ categories contributed to the total genetic gain as follows: dam of males contributed 39%, and dam of females contributed 27%. In the Western Pyrenees breeds, the AI Males’ category made the highest contribution to the total genetic gain at the end in 2016, with 47%, 47%, and 49% for the BB, MTN, and MTR breeds, respectively. The contribution to the total genetic gain by AI Males was followed by that made by Dams of Males and then by the contribution by Dams of Females. Dams of Males contributed 36%, 35%, and 39% to the total genetic gain for the BB, MTN, and MTR breeds, respectively. Finally, the dam of females category contributed 19%, 18%, and 10% to the total genetic gain in the BB, MTN, and MTR breeds, respectively. As expected, the categories that Discarded AI Males and NM Males did not contribute to the genetic gain in Mendelian sampling terms.

Figure 3 shows the average gains in terms of Mendelian samplings (selected for) by year of birth of the categories dam of males, dam of females, and AI Males. For example, in the AI Males category, the average contribution due to Mendelian sampling in 2001 is the difference between the point at the year 2001 in the trend of Mendelian samplings for AI Males in Figure 2, minus the value from the year 2000. The average contribution due to Mendelian sampling for the 2 categories of dams (dam of males and dam of females) varied between 0 and 2 L for the Western Pyrenees breeds and between 1 and 3 L for the LAC breed. The generated gain was relatively constant over time. For AI Males, the average was between –1 and 4 L. Depending on the breed, males contributed equally to or more than females, although their contribution was more irregular because at any point the number of males included in the averages shown in Figure 3 is smaller than the number of females. Similarly, less dispersion occurred in LAC compared with other breeds, due to a greater number of animals included in the averages.

The average contribution to gain due to Mendelian sampling terms for the 4 breeds varied between –1 and 1 L for Discarded AI Males (randomly, as their number is small), and it is around 0 for NM Males (with a larger number). The Discarded AI Males are not “good enough” after progeny testing, and they stop contributing their genes. Accordingly, after initial selection based on parent average, NM Males were not further selected. Thus, no contribution was obtained through their progeny to the selection response.

Now we present results on genetic contributions (measured on a scale from 0 to 1) for the BB and LAC breeds. Figure 4 shows the distribution of the  $\log_{10}$  of the contributions of  $G_0$  to the following generations.



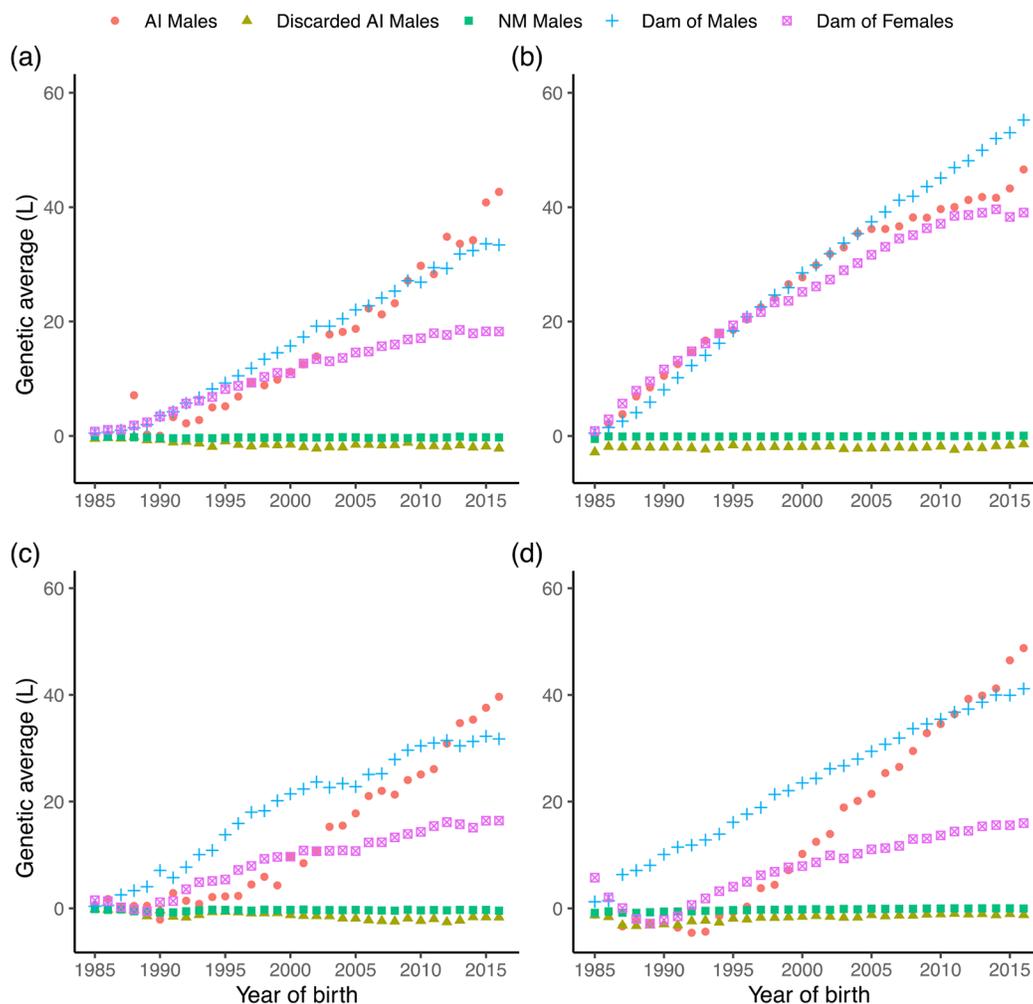
**Figure 1.** Total genetic trends for milk yield (L) and partitioning into contributions due to Mendelian samplings by sex. Breeds: (a) Basco-Béarnaise, (b) Lacaune, (c) Manech Tête Noire, (d) Manech Tête Rousse.

For instance, in  $G_7$ , the boxplot describes the final contributions from individuals in  $G_0$  (cohort 1985–1988) to individuals in  $G_7$  (cohort 2013–2016), as obtained from Equation [2] (i.e., their average contribution  $\sim 28$  yr later). Males had greater average individual genetic contributions to the genetic pool of the population than females, in both breeds, and that is because of males’ larger numbers of offspring because of AI. Greater contributions were observed for BB, because the number of animals at each generation is around 10 times lower than the number of LAC-contributed animals. For example, in the last generation ( $G_7$ ), the total number of individuals is 21,656 animals for BB and 185,247 for LAC. Therefore contributions in LAC tend to average more rapidly with time.

Genetic contributions of different categories (AI Males, Dams of Males, and Dams of Females) are

presented for both breeds in Supplemental Figure S1 (<https://doi.org/10.6084/m9.figshare.22325413.v1>, Antonios et al., 2023a). In the LAC and BB breeds, AI Males rated the highest average individual genetic contributions among all categories, followed by the genetic contributions of Dams of Males and then by the genetic contributions of Dams of Females. As mentioned above, the number of animals in the last generation ( $G_7$ ) in LAC is higher (10 times) than in the BB breed.

The distribution of the  $\log_{10}$  of contributions  $r_i$  is shown in Figure 5 (with different scales on the Y-axis). The contributions of the males are more shifted to the right (the mean of  $r_i$  is approximately  $10^{-4}$  for BB males and  $10^{-5}$  for LAC males), which indicates higher contributions for males than for females, where the females’ mean of  $r_i$  is approximately  $10^{-5}$  for BB and  $10^{-6}$  for LAC. In the real 0-to-1 scale, the maximal



**Figure 2.** Yearly partition of the genetic trends for milk yield into contributions due to Mendelian samplings by categories. Breeds: (a) Basco-Béarnaise, (b) Lacaune, (c) Manech Tête Noire, (d) Manech Tête Rouse. NM = natural mating.

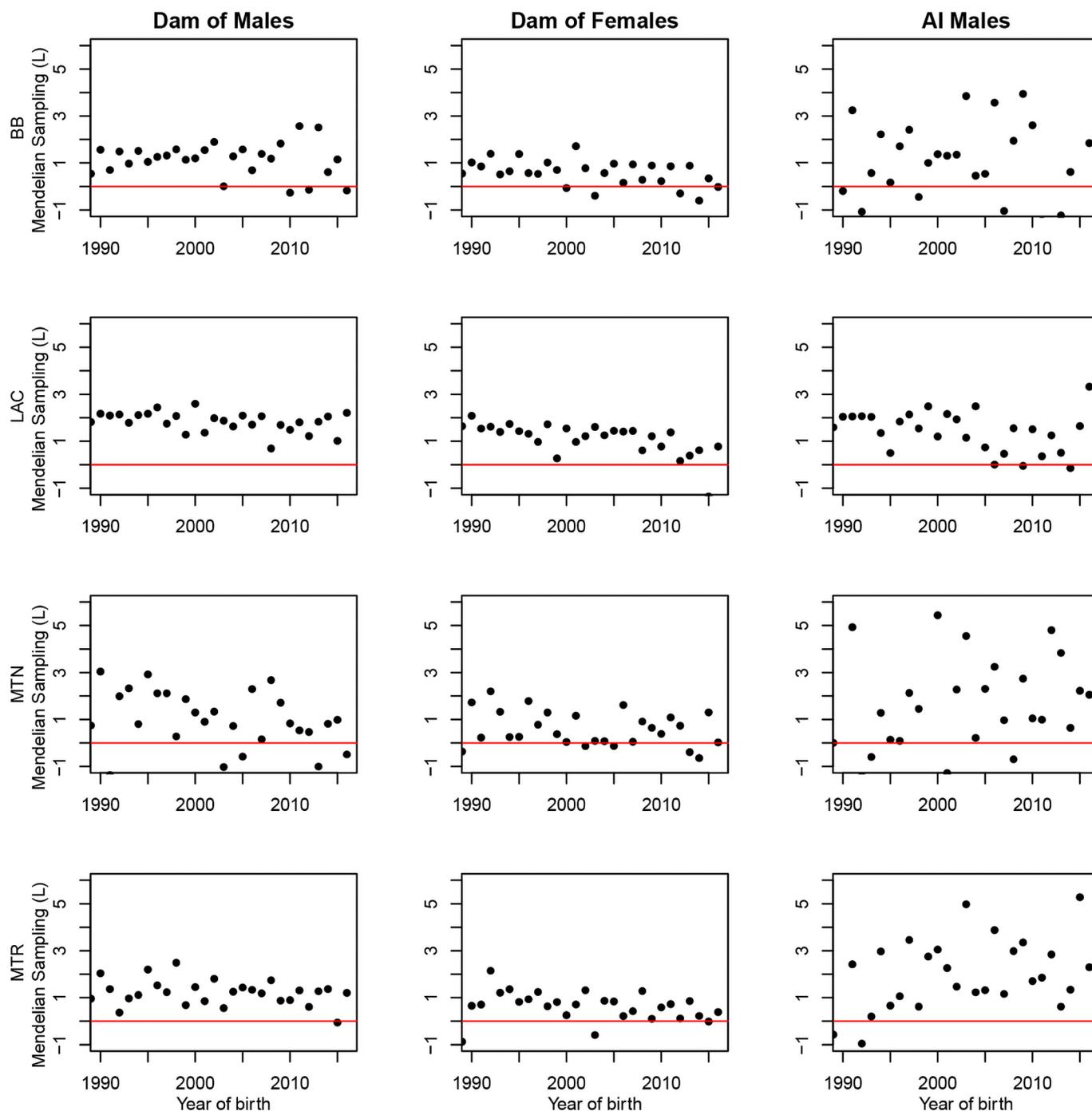
contributions of the males are approximately 0.09 for BB and 0.02 for LAC. For the females, the maximal contributions are approximately 0.03 and 0.01 for BB and LAC, respectively.

Long-term contributions ( $\log_{10}$  of the contributions) were plotted against the EBV of the first-generation individuals to their EBV in the BB and LAC breeds (see Supplemental Figure S2; <https://doi.org/10.6084/m9.figshare.21973358.v1>, Antonios et al., 2023b) and followed the expected relationship that individuals with highest EBV intrageneration will have more offspring, increasing their long-term contribution.

The logistic regression of the selection score  $x_i$  on  $\widehat{PA}_i$  and  $\hat{q}_i$  for females showed that both  $\widehat{PA}_i$  and  $\hat{q}_i$  were significant factors ( $P < 0.001$ ) in promoting selection as a parent. Estimates of odds ratio (95% CI) for Mendelian samplings were 1.0309 (1.0299–1.0317) for BB and

1.0203 (1.0200–1.0205) for LAC and, for parent average, were 1.0154 (1.0146–1.0162) for BB and 1.0149 (1.0147–1.0151) for LAC. Odds ratio for Mendelian sampling (1.0309 for BB and 1.0203 for LAC) is greater than  $\widehat{PA}_i$  (1.0154 for BB and 1.0149 for LAC). This means that for a one-unit (1-L) increase in  $\hat{q}_i$ , the odds of being selected increase 3% and 2% in BB and LAC, respectively. In males, it was not possible to perform this analysis because males registered in the pedigree are registered because they have offspring, and therefore have already been selected ( $x_i = 1$ ), whereas the others go to the slaughterhouse.

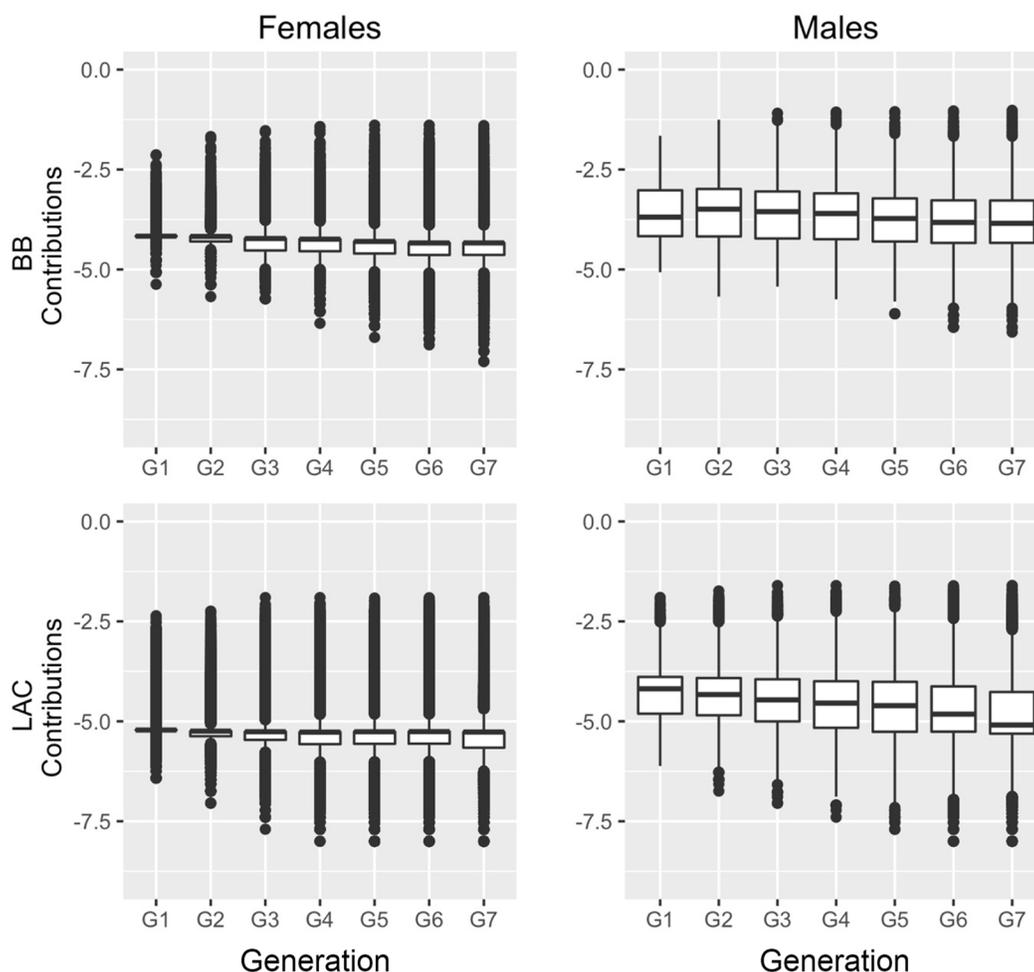
Among selected animals, we analyzed which animals contributed over time to  $G_7$  (if  $r_i > 0$  in  $G_7$ , then the binomial variable is equal to 1; otherwise, it is 0) in both sexes. Odds ratios from the logistic regression on  $\widehat{PA}_i$  and  $\hat{q}_i$  are shown in Table 5. In both sexes,  $\widehat{PA}_i$  and



**Figure 3.** Average contribution due to Mendelian sampling for milk yield by year of birth of the categories Dams of Males, Dams of Females, and AI Males. The red horizontal lines indicate the zero value on the contribution due to Mendelian sampling, which occurs when no additional gain is contributed to the next generation. Breeds: BB = Basco-Béarnaise; LAC = Lacaune; MTN = Manech Tête Noire; MTR = Manech Tête Rousse.

$\hat{\phi}_i$  were positively associated with maintaining a contribution across generations conditionally on being initially selected as parent. In BB, the effect of  $\widehat{PA}_i$  is greater than the effect of  $\hat{\phi}_i$  in both sexes. However, the

opposite was observed in LAC. This result is in agreement with Figure 1. In LAC,  $\hat{\phi}_i$  increases more in females than in males, and the odds of maintaining the contributions over time are 1% higher for LAC-selected females relative to males.



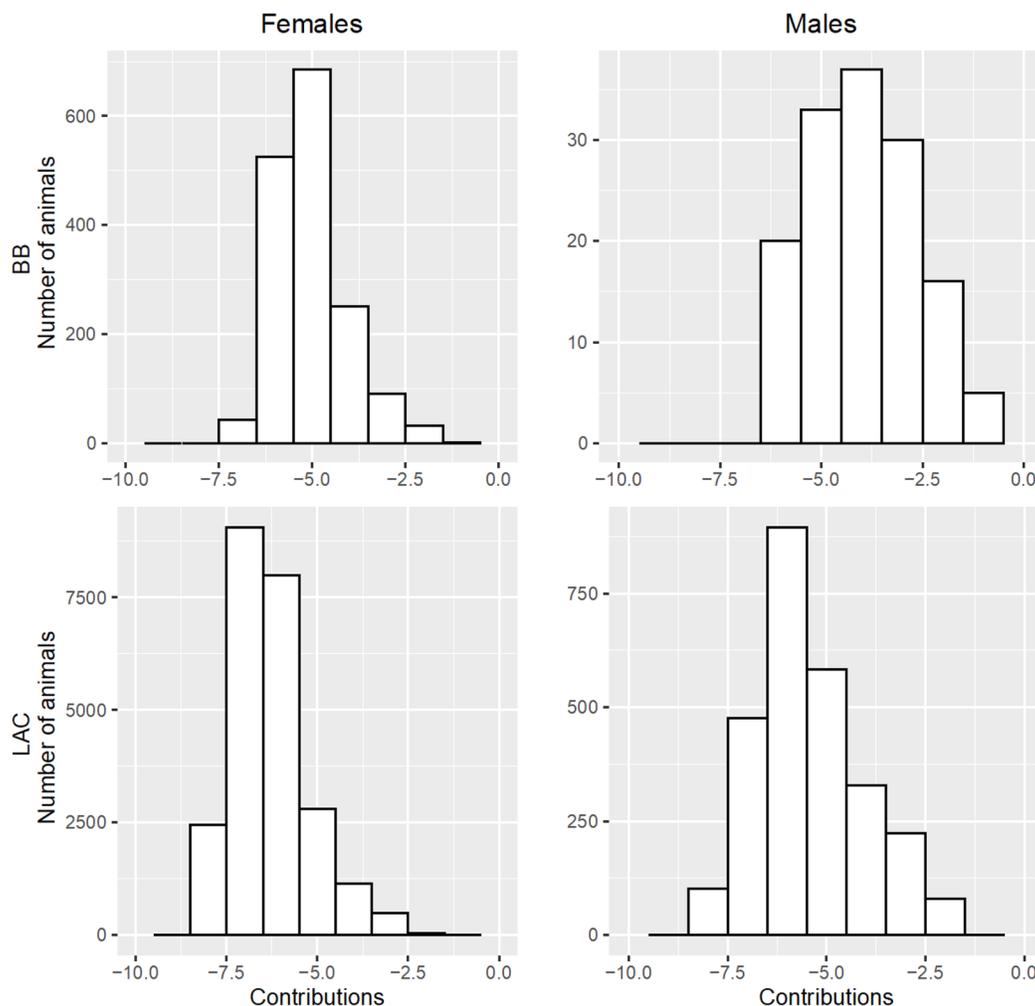
**Figure 4.** Boxplots of  $\log_{10}$  of contributions of males and females per generation from  $G_0$  (base generation) to  $G_7$  (the last generation) in Basco-Béarnaise (BB) and Lacaune (LAC) breeds. Lower whisker = minimum observation above 25% quantile  $- 1.5 \times$  interquartile range (IQR); lower line of the box = 25% quantile (Q1); horizontal line inside the box = median; upper line of the box = 75% quantile (Q3); upper whisker = maximum observation below 75% quantile  $+ 1.5 \times$  IQR; dots outside the box = outliers, values which either exceed  $Q3 + 1.5$  IQR or fall below  $Q1 - 1.5$  IQR.

### DISCUSSION

The decomposition of genetic trend into Mendelian sampling contributions by sex or categories showed that females play an important role in selection schemes. Across generations, females contribute to genetic progress as much as the males (BB, MTN, and MTR) and even more (LAC) to the final genetic gain, as shown in Figure 1. We interpret that, even if males are more heavily selected, half of their genetics come from their dam, whose positive Mendelian samplings are eventually selected and spread (often through AI Males) throughout the population. Indeed, this was observed by García-Cortés et al. (2008).

Most females are Dams of Males or Dams of Females (Figure 2). Dams of Males are progeny of AI Males, whereas Dams of Females can be progeny of AI or NM

Males. Dams of Males are expected to contribute more descendants than Dams of Females because only the sons of Dams of Males are selected. However, a dam of females can be the maternal grandmother of an AI male. The long-term contribution of an individual is given by the sum over its male and female selected offspring (Woolliams et al., 1999). The Mendelian sampling term in Dams of Males is estimated from their milk records, and also from granddaughter records when a son is selected as an AI male. In Dams of Females, their own phenotype, the daughters' milk records when they are selected as Dams of Males, and great-granddaughters' records (if a dam of females is the grandmother of an AI male) contribute to estimation of Mendelian sampling term. Accurate estimates of Mendelian samplings occur when the animal has a large number of progeny (e.g., Dams of Males) or even grand-progeny (e.g., Dams of



**Figure 5.** Histograms of the  $\log_{10}$  of the genetic contributions of males and females to the  $G_7$  (last generation) by sex in Basco-Béarnaise (BB) and Lacaune (LAC) breeds.

Males and Dams of Females). This shows that dams are as important as AI Males in driving changes in milk yield.

The decomposition of genetic trend was different in the LAC breed compared with the Western Pyrenees

**Table 5.** Estimates of odds ratios, with 95% CI, from the bivariate logistic regression of maintenance of nonzero contributions ( $r_i > 0 \mid x_i = 1$ ) on estimated parent average ( $\widehat{PA}$ ) and estimated Mendelian sampling terms ( $\hat{\phi}$ ) for females and males per breed<sup>1</sup>

Variable	BB	LAC
<b>Females</b>		
$\widehat{PA}$	1.0334 (1.0319–1.0348)	1.0057 (1.0054–1.0061)
$\hat{\phi}$	1.0303 (1.0288–1.0317)	1.0202 (1.0199–1.0205)
<b>Males</b>		
$\widehat{PA}$	1.0496 (1.0372–1.0622)	1.0035 (1.0000–1.0069)
$\hat{\phi}$	1.0061 (0.9995–1.0131)	1.0134 (1.0109–1.0158)

<sup>1</sup>BB = Basco-Béarnaise; LAC = Lacaune.

breeds. The difference between female and male contribution is greater in the LAC breed (Figure 1). In LAC, the fact that both categories of dams contribute to genetic gain in the same order as AI Males may be due to a shorter generation interval (~3 yr in LAC and ~4 yr in BB in average for dams and AI Males, as observed from averages of values in Table 3) and to the intense use of AI in the nucleus flock (greater numbers of female progenies for the AI Males in LAC than in BB, Table 3). In LAC, AI Males produced a greater number of female offspring than in the other breeds (Table 3). Thus, the EBV of LAC AI Males is more accurate, and the selection for originality (Mendelian sampling) will be more efficient. Additionally, the greater offspring numbers of LAC AI Males make possible a more intense within-family selection, as shown in Table 3. The difference in the relative importance of the Mendelian sampling and  $\widehat{PA}_i$  between breeds across generations could

reflect different selection strategies in practice. Mendelian sampling relies on within-family selection, whereas parent average is more related to across-family variation. A larger family size in LAC than in BB is in agreement with a greater importance of within-family selection in LAC, which could be because it is possible to exercise greater selection intensities within family when families are larger, as is the case.

To improve a population, we need certain animals to be better than their parents. For this, the originality of the animals with respect to their parents (i.e., the Mendelian sampling) has to be captured. The latter is estimated from different sources of information: (1) own phenotype (milk yield for females, collected in the selection scheme), (2) phenotypes of the progeny of males (progeny test; i.e., in the AI Males in these schemes), or (3) by genomic evaluation (not studied in this work, because genomic selection was not yet implemented in 2016). Although the precision in capturing Mendelian sampling is greater in males (by progeny testing), the final genetic progress is largely given by the originality (or genetic variability) of the females, even if their prediction (by own phenotype) is less accurate and their progeny number is smaller. This is true in this study and possibly in most ruminant selection schemes based on intensive use of AI. Note that a selected female can be the mother or grandmother of an AI male and consequently have a large number of grand-offspring. Moreover, because the number of eligible females is very high, the selection intensity to be a dam of males can be very high. Therefore, fine partitioning in Mendelian sampling trends can be interpreted as females being a large and easily accessible reservoir of genetic variability for selection. Still, AI Males are crucially important because they allow the spread of best genetics in the population (Figure 4 and Supplemental Figure S1) and genetically connect the herds. The smaller the number of the animals, the greater the contributions of those animals will be for future years. Additionally, males had greater contributions than females in both breeds because the AI Males, despite their smaller number compared with the dams, contribute more than dams to the gene pool of the population due to their larger number of offspring.

## CONCLUSIONS

This study decomposed the genetic gain in Mendelian samplings by categories of animals defined by sex and by selection pathways, and explored long-term genetic contributions. This allowed identification of the different contributions of categories of individuals, and better understanding of the selection scheme. Dams contribute largely to the final genetic gain by their origi-

inality, whereas AI males are crucially important in the spread of the best genetics in the population and make a proportionally greater contribution to the gene pool of the population. The relative weights of parent average and Mendelian sampling in selection decisions seem to depend on family sizes. In this study, the Mendelian sampling term was observed to be the most important factor determining the selection of a female to become dam and maintaining genetic contributions over time.

## ACKNOWLEDGMENTS

The authors thank the breeders who provided the data and the anonymous referees for useful comments on the manuscript. This study received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 772787 (SMARTER), and project ARDI (grant agreement EFA 208/16; Derio [Bizkaia], Spain) from POCTEFA funds. The GenoToul bioinformatics platform Toulouse Midi-Pyrenees provided computing and storage resources. The authors have not stated any conflicts of interest.

## REFERENCES

- Antonios, S., A. Legarra, R. Pong-Wong, J.-M. Astruc, S. T. Rodríguez-Ramilo, and Z. G. Vitezica. 2023a. Supplemental Figure S1. Boxplot of log10 of contributions of AI Males, Dam of Males and Dam of Females per generation from G0 (base generation) to G7 (the last generation) in Basco-Béarnaise (BB) and Lacaune (LAC) breeds. *figshare*. Journal contribution. <https://doi.org/10.6084/m9.figshare.22325413.v1>.
- Antonios, S., A. Legarra, R. Pong-Wong, J.-M. Astruc, S. T. Rodríguez-Ramilo, and Z. G. Vitezica. 2023b. Supplemental Figure S2. Long-term contribution (log10 of the contributions) plotted against EBV of the first-generation (G0) individuals in Basco-Béarnaise (BB) and Lacaune (LAC) breeds. *figshare*. Journal contribution. <https://doi.org/10.6084/m9.figshare.21973358.v1>.
- Astruc, J. M., F. Barillet, A. Barbat, V. Clement, and D. Boichard. 2002. Genetic evaluation of dairy sheep in France. Page 45 in Proc. 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France. World Congress on Genetics Applied to Livestock Production.
- Astruc, J.-M., D. Buisson, H. Larroque, and G. Lagriffoul. 2018. New breeding goals and role of genomics on adaptation and resilience traits in French dairy sheep. Page 705 in Proc. 69th Annual Meeting of the European Federation of Animal Science, Dubrovnik, Croatia. Wageningen Academic Publishers.
- Bijma, P. 2000. Long-term genetic contributions. Prediction of rates of inbreeding and genetic gain in selected populations. PhD thesis. Wageningen University, the Netherlands.
- García-Cortés, L. A., J. C. Martínez-Ávila, and M. A. Toro. 2008. Partition of the genetic trend to validate multiple selection decisions. *Animal* 2:821–824. <https://doi.org/10.1017/S175173110800205X>.
- Gorjanc, G., K. Potočnik, L. A. García-Cortés, J. Jakobsen, and J. Dürr. 2011. Partitioning of international genetic trends by origin in Brown Swiss bulls. In Proc. 44th Interbull Meeting, Cork, Ireland. *Interbull Bulletin*, 81-6.
- Henderson, C. R. 1973. Sire evaluation and genetic trends. *J. Anim. Sci.* 1973(Symposium):10–41. <https://doi.org/10.1093/ansci/1973.Symposium.10>.
- Howard, D. M., R. Pong-Wong, P. W. Knap, V. D. Kremer, and J. A. Woolliams. 2018. Selective advantage of implementing optimal contributions selection and timescales for the convergence of long-

- term genetic contributions. *Genet. Sel. Evol.* 50:24. <https://doi.org/10.1186/s12711-018-0392-z>.
- Macedo, F. L., J. M. Astruc, T. H. E. Meuwissen, and A. Legarra. 2022. Removing data and using metafounders alleviates biases for all traits in Lacaune dairy sheep predictions. *J. Dairy Sci.* 105:2439–2452. <https://doi.org/10.3168/jds.2021-20860>.
- Macedo, F. L., O. F. Christensen, J.-M. Astruc, I. Aguilar, Y. Masuda, and A. Legarra. 2020. Bias and accuracy of dairy sheep evaluations using BLUP and SSGBLUP with metafounders and unknown parent groups. *Genet. Sel. Evol.* 52:47. <https://doi.org/10.1186/s12711-020-00567-1>.
- Obšteter, J., J. Holl, J. M. Hickey, and G. Gorjanc. 2021. AlphaPart—R implementation of the method for partitioning genetic trends. *Genet. Sel. Evol.* 53:30. <https://doi.org/10.1186/s12711-021-00600-x>.
- Quaas, R. L. 1988. Additive genetic model with groups and relationships. *J. Dairy Sci.* 71:91–98. [https://doi.org/10.1016/S0022-0302\(88\)79986-5](https://doi.org/10.1016/S0022-0302(88)79986-5).
- Škorput, D., G. Gorjanc, A. Kasap, and Z. Luković. 2015. Partition of genetic trends by origin in Landrace and Large-White pigs. *Animal* 9:1605–1609. <https://doi.org/10.1017/S1751731115001056>.
- Thompson, R. 1979. Sire evaluation. *Biometrics* 35:339–353. <https://doi.org/10.2307/2529955>.
- Woolliams, J. 2007. Genetic contributions and inbreeding. Pages 147–165 in *Utilisation and Conservation of Farm Animal Genetic Resources*. K. Oldenbroek, ed. Wageningen Academic Publishers.
- Woolliams, J., and R. Thompson. 1994. A theory of genetic contributions. Pages 127–134 in *Proc. 5th World Congress on Genetics Applied to Livestock Production*, Guelph, Canada. World Congress on Genetics Applied to Livestock Production.
- Woolliams, J. A., P. Bijma, and B. Villanueva. 1999. Expected genetic contributions and their impact on gene flow and genetic gain. *Genetics* 153:1009–1020. <https://doi.org/10.1093/genetics/153.2.1009>.
- Woolliams, J. A., N. R. Wray, and R. Thompson. 1993. Prediction of long-term contributions and inbreeding in populations undergoing mass selection. *Genet. Res.* 62:231–242. <https://doi.org/10.1017/S0016672300031943>.

## ORCID

- S. Antonios  <https://orcid.org/0000-0002-2265-5118>
- A. Legarra  <https://orcid.org/0000-0001-8893-7620>
- J. M. Astruc  <https://orcid.org/0000-0001-7685-1301>
- S. T. Rodríguez-Ramilo  <https://orcid.org/0000-0001-7150-0692>
- Z. G. Vitezica  <https://orcid.org/0000-0002-3162-7258>