L'impact de la sélection génomique sur la diversité et le progrès génétiques dans les races bovines laitières françaises de grand effectif

I - Résumé de l'étude

En France, la mise en place des évaluations génomiques dans les races bovines laitières de grand effectif a commencé en 2009 (Boichard et al., 2012), et a modifié radicalement les schémas de sélection. Il est donc indispensable d'en évaluer les conséquences réelles sur la diversité génétique. Cela permettra aux entreprises de sélection d'adapter éventuellement ces schémas de sélection dans le futur afin d'optimiser à la fois le progrès et la diversité génétiques.

Dans ce contexte, l'objectif de cette étude est de comprendre l'impact de la sélection génomique sur la diversité génétique dans trois races bovines laitières françaises : la Montbéliarde et la Normande, deux races nationales, et la Prim'Holstein (appelée Holstein dans l'article), une race internationale. Nous avons mesuré les changements relatifs du progrès génétique et de trois estimations différentes de la diversité génétique : (i) les taux de consanguinité et d'apparentement à partir du pedigree, (ii) le taux de consanguinité à partir des ROH, et (iii) l'évolution de la longueur moyenne des ROH ; avant et après la mise en place de la sélection génomique dans ces trois races. Nous avons comparé les estimations de la consanguinité à partir des pedigrees et des ROH et exploré les changements dans les schémas de sélection qui pourraient expliquer les différences observées entre les races nationales et internationales.

La sélection génomique a induit une augmentation du progrès génétique annuel de 50, 71 et 33% pour les taureaux de race Montbéliarde, Normande et Holstein, respectivement, et les intervalles de génération ont été réduits d'un facteur de 1,7, 1,9 et 2, respectivement. Nous n'avons constaté aucun changement significatif du taux de consanguinité annuel pour les deux races nationales, Montbéliarde et Normande, mais une augmentation significative du taux de consanguinité annuel pour la race internationale Holstein, qui atteint désormais 0,55 % par an à partir des ROH et 0,49 % par an à partir des données généalogiques (ce qui équivaut à un taux de 1,36 et 1,39 % par génération, respectivement). Les ROH étaient en moyenne plus longs pour les taureaux de la race Holstein que pour ceux des deux autres races.

Après la mise en place de la sélection génomique, le progrès génétique annuel estimé à partir de l'ISU a augmenté pour les taureaux de trois grandes races bovines laitières françaises : la Montbéliarde, la Normande et la Prim'Holstein. Cette augmentation était attendue suite aux résultats de précédentes études prospectives sur données simulées ou rétrospectives sur données réelles. Dans le même temps, le taux de consanguinité par an et par génération a augmenté pour

les taureaux de la race internationale Prim'Holstein, comme cela a été observé dans d'autres régions du monde comme les Pays-Bas ou l'Amérique de Nord. L'augmentation du taux de consanguinité chez la Prim'Holstein était en partie due à une augmentation de la consanguinité récente (moins de 5 générations). Aucune augmentation du taux de consanguinité n'a été observée pour les taureaux de races nationales Montbéliarde et Normande. Cette différence reflète des modifications différentes des schémas de sélection dans les trois races, suggérant que l'augmentation du ratio nombre efficace/nombre réel de taureaux dans la race Prim'Holstein n'a pas suffi à compenser l'introduction substantielle de consanguinité récente dans cette race. Cette augmentation de la consanguinité récente en Prim'Holstein peut s'expliquer par l'utilisation intensive de quelques individus seulement comme grands-pères des taureaux, principalement originaires des États-Unis.

II - Article publié : *The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds*

GSE Genetics Selection Evolution

RESEARCH ARTICLE

Check for updates

The impact of genomic selection on genetic diversity and genetic gain in three French dairy cattle breeds

Anna-Charlotte Doublet[1,2]* , Pascal Croiseau[1], Sébastien Fritz[1,2], Alexis Michenet[1,2], Chris Hozé[1,2], Coralie Danchin-Burge[3], Denis Laloë[1] and Gwendal Restoux[1]

Abstract

Background: In France, implementation of genomic evaluations in dairy cattle breeds started in 2009 and this has modified the breeding schemes drastically. In this context, the goal of our study was to understand the impact of genomic selection on the genetic diversity of bulls from three French dairy cattle breeds born between 2005 and 2015 (Montbéliarde, Normande and Holstein) and the factors that are involved.

Methods: We compared annual genetic gains, inbreeding rates based on runs of homozygosity (ROH) and pedigree data, and mean ROH length within breeds, before and after the implementation of genomic selection.

Results: Genomic selection induced an increase in mean annual genetic gains of 50, 71 and 33% for Montbéliarde, Normande and Holstein bulls, respectively, and in parallel, the generation intervals were reduced by a factor of 1.7, 1.9 and 2, respectively. We found no significant change in inbreeding rate for the two national breeds, Montbéliarde and Normande, and a significant increase in inbreeding rate for the Holstein international breed, which is now as high as 0.55% per year based on ROH and 0.49% per year based on pedigree data (equivalent to a rate of 1.36 and 1.39% per generation, respectively). The mean ROH length was longer for bulls from the Holstein breed than for those from the other two breeds.

Conclusions: With the implementation of genomic selection, the annual genetic gain increased for bulls from the three major French dairy cattle breeds. At the same time, the annual loss of genetic diversity increased for Holstein bulls, possibly because of the massive use of a few elite bulls in this breed, but not for Montbéliarde and Normande bulls. The increase in mean ROH length in Holstein may reflect the occurrence of recent inbreeding. New strategies in breeding schemes, such as female donor stations and embryo transfer, and recent implementation of genomic evaluations in small regional breeds should be studied carefully in order to ensure the sustainability of breeding schemes in the future.

Background

In France, genomic evaluations started to be implemented in dairy cattle breeds in 2009. After a brief transition period for the breeding companies to adopt this technology, these new evaluation methods have drastically changed dairy cattle breeding schemes since 2011. Genomic evaluations have allowed breeding companies

to increase the number of male candidates for selection and the use of new bulls [1, 2]. At the early selection stage, genomic selection (GS) allows exploiting a major part of the genetic variance compared to the previously used models based on parents' average values [1, 3]. In fact, Mendelian sampling variation, which is defined as the difference between the genetic value of an individual and the average of the parental genetic values, can be captured from molecular markers without any information on the parents' phenotype or the offspring's phenotype. This makes it possible to obtain breeding values

*Correspondence: anna-charlotte.doublet@inra.fr
[1] GABI, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France
Full list of author information is available at the end of the article

Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 2 of 13

with higher reliability sooner in the life of an individual compared with those based on pedigree data [4]. Indeed, the estimation of pedigree-based breeding values at birth depends mostly on information about related individuals, resulting in a low reliability and a high correlation between estimated breeding values of relatives. These high correlations decrease only later in the life of an animal when information based on its own performance or performances of the progeny becomes available. Consequently, GS reduces the number of related individuals chosen as candidates for selection [5]. Moreover, collecting genotype data is less expensive than collecting phenotype data, which means that more male candidates can be evaluated with high reliability. Furthermore, methods to balance inbreeding and genetic gain in breeding schemes, such as optimal contribution selection [6], allow for better conservation of genetic diversity when based on genomic information rather than on pedigree data [7, 8]. Thus, GS was expected to reduce the rate of inbreeding and to increase genetic gain per generation at the same time [2].

However, since animals with high genomic estimated breeding values (GEBV) can be selected as parents at a very young age, generation intervals are significantly shortened [9]. As a result, the annual inbreeding rate could then increase and this has already been observed in Dutch-Flemish Holstein–Friesian cattle [10] and in North American Holstein cattle [11].

Such an increase in inbreeding rate results in a loss of genetic diversity, with lower additive genetic variance leading to a lower response to selection [12]. Moreover, a lower overall genetic variability leads also to a loss of adaptive potential for selecting new breeding goals in the context of climate change [13], and more severe inbreeding depression [14–16].

Measuring genetic diversity is necessary in order to manage it. Typically, genetic diversity is estimated and managed with the help of inbreeding and kinship coefficients based on pedigree data (e.g. [17, 18]). These coefficients reflect the proportion of genome expected to be identical-by-descent (IBD) in an individual and between two individuals, respectively. The reliability of these estimates depends on the quality of the pedigree records and their depth [7, 19, 20]. In particular, pedigree-based inbreeding and kinship may be underestimated due to incomplete pedigree data and the assumption of arbitrary non-inbred and unrelated founder individuals, thus leading to an overestimation of genetic diversity.

Falling costs in the field of high-throughput genotyping and sequencing make it possible to obtain more and more genotyping data, and thus to estimate genetic diversity more precisely. These data allow the computation of inbreeding coefficients and genomic relationship matrices (GRM) [21, 22]. The calculation of inbreeding coefficients from GRM computed using all the allele frequencies set to 0.5 globally consists in computing observed and expected proportions of homozygous single nucleotide polymorphisms (SNPs) considered as identical-by-state (IBS).

Another method to estimate genetic diversity consists in detecting IBD segments, called runs of homozygosity (ROH), by considering that segments of consecutive homozygous SNPs originate from common ancestors [23, 24]. Thus, it is possible to compute ROH-based inbreeding coefficients defined as the proportion of the autosomal genome that is actually included in ROH above a given length threshold [24], corresponding to the realized proportion of genome that is IBD. This approach to estimate inbreeding coefficients seems to reflect IBD better than genomic inbreeding coefficients calculated from molecular data based on the GRM [11, 25–27]. The length of such ROH segments follows an inverse exponential distribution with expectation 1/2G Morgan, with G being the number of generations to the common ancestor from which the segment arises (e.g. [28]). Due to recombination, ROH length tends to decrease over generations [29, 30], which allows to estimate the age of inbreeding events along the genome and to trace back population history [29, 31, 32]. Shorter ROH tend to result from older inbreeding (older common ancestors) whereas longer ROH tend to result from more recent inbreeding (more recent common ancestors) [29, 31, 32]. However, different criteria have been proposed to define ROH-based inbreeding estimates and there is no consensus about which one should be preferred (e.g. the number of heterozygous loci allowed in one ROH, the minimal length, etc.). Consequently, ROH-based inbreeding estimates might differ, depending on the criteria used [28].

Pedigree data are useful to evaluate genetic diversity for individuals that are not genotyped, whereas measures based on molecular data (such as ROH) allow the study of the realized proportion of genome that is autozygous and the pattern of inbreeding along the genome. In spite of their advantage over pedigree-based inbreeding estimates, genomic-based estimates began to be used to manage genetic diversity in selection programs only recently [33–35].

In this context, the goal of our study was to understand the impact of GS on genetic diversity in three French dairy cattle breeds: Montbéliarde and Normande, two national breeds, and Holstein, an international breed. We measured the relative changes in rates of genetic merit and in three different estimates of genetic diversity: (i) pedigree-based inbreeding and kinship, (ii) ROH-based inbreeding coefficients, and (iii) mean length of ROH; and these estimates were taken before and after the

Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 3 of 13

implementation of GS in these three breeds in France. Based on our findings, we compare pedigree and ROH-based inbreeding rates within breeds and explore the changes in breeding schemes that could explain the differences observed between national and international breeds.

Methods

Animals, pedigree and genotyping data

We used three datasets that consisted of all progeny-tested bulls born between 2005 and 2010 and all marketed bulls born between 2012 and 2015 from three French dairy cattle breeds: 1307 Montbéliarde bulls and 1016 Normande bulls (national breeds), and 4694 Holstein bulls (international breed) (see Table 1 and Fig. 1). We chose these categories of bulls in order to obtain cohorts of comparable sizes and to compare bulls that are selected for artificial insemination by breeding companies (either to be marketed or to be tested). Key data about these breeds in France are available in Additional file 1: Table S1.

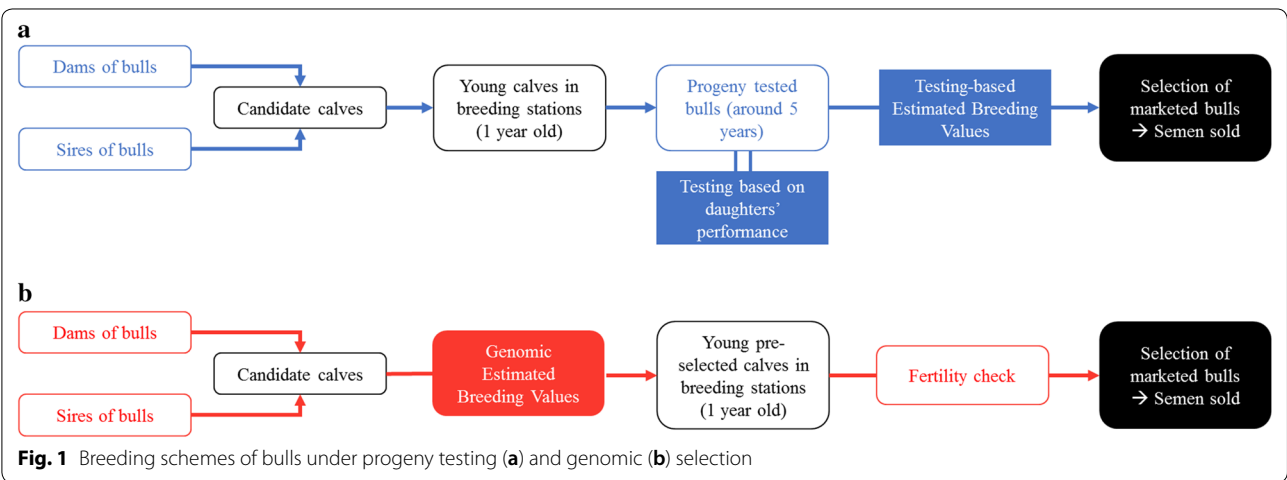Genealogies of all bulls in the three datasets were traced back as far as possible. The mean numbers of generations traced back for bulls in the datasets were 15.4 ± 1.7, 15.8 ± 1.6 and 15.4 ± 1.9 for Montbéliarde, Normande and Holstein, respectively (longest paths computed with R [36] package pedigree [37]).

All the bulls in the datasets were genotyped for 50 K SNPs (Illumina Infinium® BovineSNP50 BeadChip). Quality control filtering was performed by removing SNPs that were genotyped in less than 95% individuals and SNPs with a minor allele frequency (MAF) lower than 1%. No individual had less than 95% genotyped SNPs. After quality control filtering, 39,992, 40,135 and 41,377 SNPs remained in the datasets for Montbéliarde, Normande and Holstein, with on average one SNP every 62.5 ± 66.5 kb, 62.2 ± 66.4 kb and 60.4 ± 63.8 kb, respectively.

The total genetic merit index ISU (standing for *Index Synthèse Unique*) is an individual index that combines production, functional and type traits, which are weighted depending on the breed and subsequent breeding goals. The formula to obtain total genetic merit index ISU is built in order to obtain a standard error of the index of 20 points within each breed [38]. Within each breed, all genetic merit estimates were calculated relative to the same arbitrary base.

Table 1 Number of animals in each dataset

Breed	Birth period	Number of individuals born each year				Total number of individuals
		Min	Mean (± sd)	Max	Median	
Montbéliarde	2005–2010 Progeny-tested bulls	82	138 (± 34)	176	141	826
	2012–2015 Marketed bulls	81	120 (± 29)	149	125	480
Normande	2005–2010 Progeny-tested bulls	82	111 (± 21)	135	112	666
	2012–2015 Marketed bulls	73	87 (± 11)	100	87	347
Holstein	2005–2010 Progeny-tested bulls	368	573 (± 157)	697	606	3440
	2012–2015 Marketed bulls	294	312 (± 12)	322	316	1248

Fig. 1 Breeding schemes of bulls under progeny testing (**a**) and genomic (**b**) selection

Doublet *et al. Genet Sel Evol* (2019) 51:52

Page 4 of 13

Measures of genetic diversity

Pedigree-based measures of genetic diversity

Pedigree-based inbreeding coefficients F_{ped} and kinship coefficients f_{ped} were computed with the Pedig software [39] for all the animals in the three datasets. The mean of all kinship coefficients f_{ped} for bulls born between 2005 and 2010 and the mean of all kinship coefficients f_{ped} for bulls born between 2012 and 2015 were compared with a two-sample t test. We also computed pedigree-based coefficients by taking only the last five generations for each bull into account, with the Pedig software [39] for all animals in the three datasets.

Measures of genetic diversity based on runs of homozygosity

ROH reflect autozygous segments of the genome. A ROH was defined as a homozygous segment of at least 15 SNPs or 1000 kb long, with at least one SNP per 75 kb. Two consecutive SNPs could not be included in the same ROH if they were more than 150 kb apart. ROH were detected using the PLINK 1.9 "homozyg" function [40, 41] (command line: plink --cow --bfile genotyping_data_filename --homozyg --homozyg-kb 1000 --homozyg-snp 15 --homozyg-window-snp 15 --homozyg-density 75 --homozyg-gap 150 --out output_filename).

ROH-based inbreeding estimates, $F_{ROH,i}$, were computed as the proportion of genome included in ROH [10, 24] as follows:

$$F_{ROH,i} = \frac{\Sigma L_{ROH,i}}{L_{auto}},$$

where $\Sigma L_{ROH,i}$ is the total length of ROH for individual i, and L_{auto} the length of the autosomal genome covered by SNPs after withholding gaps longer than 150 kb between two SNPs, corresponding to the length of the autosomal genome on which ROH can be detected. These parameters allowed the detection of ROH on 77.0, 77.0 and 78.6% of the autosomal genome of Montbéliarde, Normande and Holstein bulls, respectively (see Additional file 2: Table S2).

For each individual, the mean ROH length, i.e. $L_{ROH,mean,i}$, was computed as:

$$L_{ROH,mean,i} = \frac{\Sigma L_{ROH,i}}{N_{ROH,i}},$$

where $\Sigma L_{ROH,i}$ is the total length of ROH for individual i in kb, and $N_{ROH,i}$ the total number of ROH for individual i.

Pearson's product-moment correlations between pedigree- and ROH-based inbreeding coefficients were computed for each breed in order to check if it was relevant to compare both inbreeding estimates.

To compare the variability of pedigree and ROH-based inbreeding rates, we calculated the coefficient of variation

of these two parameters within each breed as the ratio between the overall standard deviation of inbreeding and its overall mean for each dataset.

We obtained approximate inbreeding rates per generation for a selection type by multiplying the mean inbreeding rate by the mean parental generation interval for this period (in years) (see Additional file 3: Table S3).

Demographic parameters

Generation interval I_1 was computed with an adapted version of Pedig software [39]. I_1 was defined as the mean difference between the birthdate of an individual and its parents, in months.

For each breed, the number of bulls born per year was divided by the number of bulls born in 2005 in the three datasets, giving S_i:

$$S_i = \frac{N_i}{N_{2005}},$$

where N_i is the number of bulls born in year i.

The effective number of bulls, $Ne_{breed, selection}$, for a given breed and a given type of selection was computed as:

$$Ne_{breed, selection} = \frac{1}{\sum_{i=1}^{n_{breed, selection}} \left(\frac{o_i}{\sum_{j=1}^{n_{breed, selection}} o_j} \right)^2},$$

where $n_{breed, selection}$ is the number of bulls for a given breed and a given type of selection and o_i the number of offspring of bull i [42, 43]. We also computed the effective/census number of bulls ratio, $Ne_{breed, selection} / n_{breed, selection}$. The two types of selection were progeny testing selection (PTS) (individuals born between 2005 and 2010), and genomic selection (GS) (individuals born between 2012 and 2014).

95% confidence intervals for $Ne_{breed, selection}$ and $Ne_{breed, selection} / n_{breed, selection}$ were generated by random resampling with replacement of the datasets (bootstrap with 1000 iterations). For a given breed, $Ne_{breed, PTS}$ and $Ne_{breed, GS}$, or $Ne_{breed, PTS} / n_{breed, PTS}$ and $Ne_{breed, GS} / n_{breed, GS}$ were statistically different if confidence intervals did not overlap.

Impact of genomic selection on genetic diversity and genetic gain

In order to assess the impact of GS on genetic diversity and genetic gain, we used the following linear model, using the R function lm [36]:

$$Y_i = \begin{cases} a_1 + b_1.x_i + \varepsilon_i, & 2005 \leq x_i \leq 2010 \\ a_2 + (b_1 + \delta).x_i + \varepsilon_i, & 2012 \leq x_i \leq 2015 \end{cases}$$

Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 5 of 13

where Y_i is the variable of interest for bull i (F_{ped}, f_{ped}, $F_{ROH,i}$, $L_{ROH,mean,i}$ for genetic diversity and ISU for genetic gain), x_i is the birth year of bull i, and b_1 the associated coefficient of regression if bull i was born between 2005 and 2010 (PTS) or ($b_1 + \delta$) if born between 2012 and 2015 (GS). The impact of GS was measured with the δ coefficient. Its significance was tested with an analysis of variance.

The quality of linear regressions was checked by looking at the square root of the standardized residuals of each linear regression (see Additional file 4: Figures S1–S7).

The relative change, RC, of the slopes of regression before and after GS was computed as:

$$RC = \frac{\delta}{|b_1|}.$$

For a better interpretation of relative change, please note that the sign of RC quantifies the direction of the change of the slopes (if $RC > 0$, the new slope will be higher than the previous slope and if $RC < 0$, it will be lower) and its absolute value quantifies the magnitude of the change of the slopes.

Results
Genetic merit
The total genetic merit index ISU increased at a significantly higher rate in Montbéliarde, Normande and Holstein under GS (between 2012 and 2015) than under PTS (between 2005 and 2010), with respective relative changes of 0.50, 0.71 and 0.33 (see Table 2 and Fig. 2a). The slopes of ISU during GS are equivalent to increases of ~ 4.4, 5.0 and 7.4% per year relative to the mean genetic merit ISU in 2012 of 111, 118 and 139 points for Montbéliarde, Normande and Holstein, respectively.

Genetic diversity
Pedigree-based measures of genetic diversity
Pedigree-based inbreeding rates increased significantly in Normande and Holstein under GS schemes (between 2012 and 2015) compared to PTS (between 2005 and 2010), i.e. from 0.059 to 0.088% per year to 0.19 and 0.49% per year for Normande and Holstein, respectively (from 0.34 to 0.46% per generation to 0.70 and 1.39% per generation, respectively) with respective relative changes of 2.20 and 4.50 (see Table 2). This was not the case in Montbéliarde, for which the inbreeding rate per year between 2005 and 2010 was 0.090% (0.52% per generation) (see Table 2). In 2005, the overall pedigree-based inbreeding level in Holstein did not differ significantly from that in Montbéliarde and Normande (see Fig. 2b). However, in 2015, the pedigree-based inbreeding level was significantly higher in Holstein than in Montbéliarde (p-value $< 1e-10$) and Normande (p-value $< 1e-10$) (see Fig. 2b). The slopes

Table 2 Trends of the different inbreeding criteria estimating genetic gain and genetic diversity per year

Parameter	Breed	b_1 ($\pm$ SE)	b_2	δ ($\pm$ SE)	p-value of δ	RC
ISU	MON	3.25 ($\pm$ 0.28)	4.89	1.64 ($\pm$ 0.72)	0.023	0.50
	NOR	3.42 ($\pm$ 0.31)	5.85	2.43 ($\pm$ 0.79)	2.1e−03	0.71
	HOL	7.77 ($\pm$ 0.16)	10.31	2.54 ($\pm$ 0.41)	6.9e−10	0.33
F_{ped} (in %)[a]	MON	0.090 ($\pm$ 0.029)	0.11	0.019 ($\pm$ 0.076)	0.81	0.21
	NOR	0.059 ($\pm$ 0.024)	0.19	0.13 ($\pm$ 0.061)	0.035	2.20
	HOL	0.088 ($\pm$ 0.013)	0.49	0.40 ($\pm$ 0.033)	<1e−10	4.50
f_{ped} (in %)[a]	MON	− 0.14 ($\pm$ 0.006)	− 0.013	0.13 ($\pm$ 0.018)	<1e−10	0.91
	NOR	− 0.11 ($\pm$ 0.008)	0.072	0.19 ($\pm$ 0.022)	<1e−10	1.64
	HOL	0.20 ($\pm$ 0.002)	0.22	0.019 ($\pm$ 0.005)	1.5e−04	0.092
F_{ROH} (in %)[a]	MON	0.17 ($\pm$ 0.033)	0.076	− 0.095 ($\pm$ 0.085)	0.26	− 0.56
	NOR	0.12 ($\pm$ 0.033)	0.14	0.019 ($\pm$ 0.082)	0.82	0.16
	HOL	0.080 ($\pm$ 0.018)	0.55	0.47 ($\pm$ 0.045)	<1e−10	5.93
Mean ROH[b] length	MON	4.75 ($\pm$ 1.64)	0.09	− 4.66 ($\pm$ 4.22)	0.27	− 0.98
	NOR	1.32 ($\pm$ 1.74)	1.2	− 0.149 ($\pm$ 4.36)	0.97	− 0.11
	HOL	1.53 ($\pm$ 0.965)	13	11.9 ($\pm$ 2.43)	1.1e−06	7.77

b_1 is the slope of regression of each parameter depending on birth year for progeny-tested bulls born between 2005 and 2010 (progeny testing selection), and δ the difference between the slopes of regression of each parameter depending on birth year for progeny testing selection and for marketed bulls born between 2012 and 2015 (genomic selection), for Montbéliarde (MON), Normande (NOR) and Holstein (HOL). The relative change (RC) is equal to $\frac{\delta}{|b_1|}$. b_2 is the slope of the parameter for marketed bulls born between 2012 and 2015 (genomic selection), equal to $b_1 + \delta$. The p-value of δ corresponds to the significance of the non-nullity of δ

ISU, total merit index; F_{ped}, pedigree-based inbreeding; f_{ped}, pedigree-based kinship; F_{ROH}, inbreeding based on Runs of Homozygosity (ROH)

[a] Slopes and standard errors are displayed in %. *p*-values and relative changes are not

[b] Runs of homozygosity

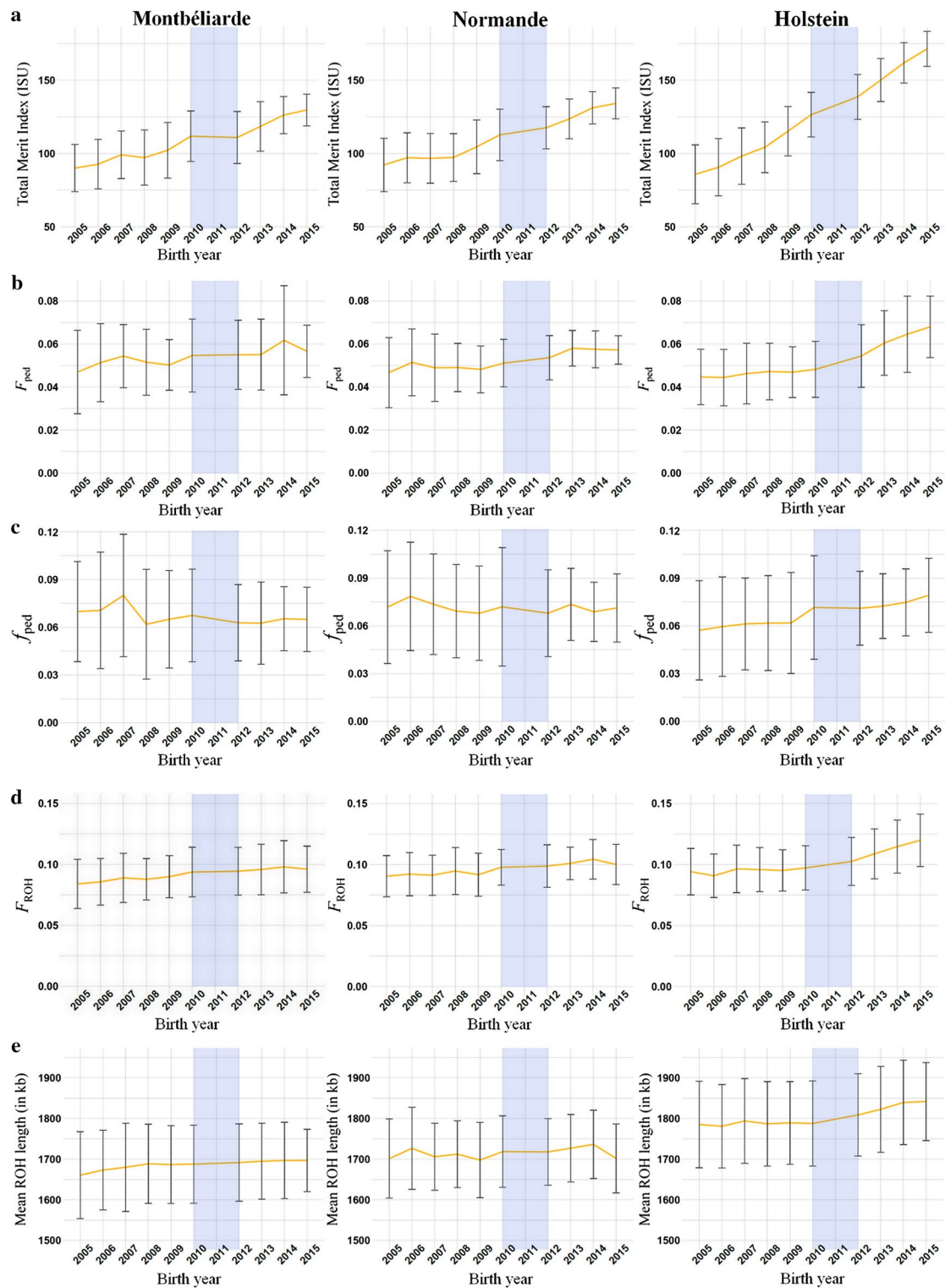Doublet *et al. Genet Sel Evol* (2019) 51:52

Page 6 of 13

Fig. 2 Trends of the different parameters estimating genetic gain and genetic diversity. The yellow line represents the mean for each birth year and the error bars the standard deviation. Total merit indices (ISU) (**a**), inbreeding (**b**) and kinship (**c**) based on pedigree data, inbreeding based on runs of homozygosity (ROH) (**d**) and mean ROH length (in kb) (**e**) were retrieved or computed for progeny-tested bulls born between 2005 and 2010 and for marketed bulls born between 2012 and 2015, for Montbéliarde, Normande and Holstein

of pedigree-based inbreeding coefficients during GS are equivalent to increases of ~ 2.0, 3.5 and 8.9% per year of the mean pedigree-based inbreeding coefficients in 2012 for Montbéliarde, Normande and Holstein, respectively.

For inbreeding coefficients based on pedigree and calculated over the last five generations only, the slope of regression was negative during PTS for the three breeds. Under GS, it decreased significantly for Montbéliarde and Normande, reaching -0.32% and -0.23% per year, respectively (-1.22% and -0.86% per generation, respectively), whereas it increased significantly for Holstein, reaching 0.27% per year (0.78% per generation) (Fig. 3) (see Additional file 5: Tables S4, S5).

Pedigree-based kinship increased at a significantly higher rate in Montbéliarde, Normande and Holstein under GS (between 2012 and 2015) than under PTS (between 2005 and 2010), with respective relative changes of 0.91, 1.64 and 0.092 (see Table 2). The slopes of pedigree-based kinship coefficients during GS are equivalent to increases of about -0.20, 1.1 and 3.1% per year of the mean pedigree-based kinship coefficients in 2012 for Montbéliarde, Normande and Holstein, respectively (see Fig. 2c). In Montbéliarde and Normande, pedigree-based kinship coefficients seemed to plateau during both PTS and GS, although we observed a drop in the average pedigree-based kinship in Montbéliarde in 2008 (see Fig. 2c). The mean value of pedigree-based kinship coefficients was significantly lower for Montbéliarde born between 2012 and 2015 than for Montbéliarde bulls born between 2005 and 2010 (p-value < 1e−10), as well as for Normande bulls (p-value < 1e−10) (see Fig. 2c). By contrast, this value was significantly higher for Holstein bulls born between 2012 and 2015 than for those born between 2005 and 2010 (p-value < 1e−10) (see Fig. 2c).

Measures of genetic diversity based on runs of homozygosity (ROH)

The rate of ROH-based inbreeding increased significantly in Holstein under GS (between 2012 and 2015) compared to under PTS (between 2005 and 2010), i.e. from 0.080% per year to 0.55% per year for Holstein (from 0.39% to 1.36% per generation) with a relative change of 5.93 (see Table 2). This was not the case for Montbéliarde and Normande, for which inbreeding rates of 0.17 and 0.12% per year were found between 2005 and 2010, respectively (0.96 and 0.71% per generation) (see Table 2). In 2005, the overall ROH-based inbreeding level in Montbéliarde was significantly lower than that in Normande and Holstein (p-values = 0.0031 and 2.1e−08, respectively) (see Fig. 2d). However, in 2015, the level of ROH-based inbreeding was significantly higher in Holstein than in

Montbéliarde and Normande (p-values < 1e−10, in both cases) (see Fig. 2d). The slopes of ROH-based inbreeding coefficients during GS are equivalent to increases of ~ 0.8, 1.4 and 5.4% per year of the mean ROH-based inbreeding coefficients in 2012 for Montbéliarde, Normande and Holstein, respectively.

Mean ROH length increased at a significantly higher rate in Holstein under GS (between 2012 and 2015) than under PTS (between 2005 and 2010), with a relative change of 7.77 (see Table 2). This was not the case in Montbéliarde and in Normande, for which the mean ROH length plateaued under GS (see Fig. 2e). From 2005 to 2015, the overall mean ROH length for Montbéliarde was significantly shorter than that for Normande, which in turn was significantly shorter than that for Holstein (p-values < 1e−10 in all cases, with average mean ROH lengths of 1.68, 1.71 and 1.80 Mb, respectively) (see Fig. 2c). The slopes of mean ROH lengths during GS showed increases of ~ 0.005, 0.07 and 0.7% per year of the mean ROH lengths in 2012 for Montbéliarde, Normande and Holstein, respectively.

Correlations between and variability of pedigree- and ROH-based inbreeding coefficients

Pearson's product-moment correlation between pedigree- and ROH-based inbreeding coefficients were equal to 59% (95% confidence interval: [55%, 62%]), 50% [45%, 55%] and 59% [57%, 61%] for Montbéliarde, Normande and Holstein, respectively.

Coefficients of variation of pedigree-based inbreeding were equal to 0.33, 0.25 and 0.31 for Montbéliarde, Normande and Holstein, respectively. Coefficients of variation of ROH-based inbreeding were equal to 0.22, 0.18 and 0.21 for Montbéliarde, Normande and Holstein, respectively.

Demographic parameters
Generation intervals

The mean difference I_1 between the birthdate of an individual and its parents reached a plateau at ~ 70 months in Montbéliarde and in Normande, and 59 months in Holstein between 2005 and 2009. I_1 was reduced by a factor 1.7 in Montbéliarde, by 1.9 in Normande and by a factor 2 in Holstein between animals born between 2005 and 2009 (under PTS) and animals born in 2015 (under GS). To date, this interval has not reached a new plateau for the three breeds (see Fig. 4).

Number and use of bulls

A histogram of the proportion of the total number of offspring per bull for each breed and selection type is available in Additional file 6: Figure S8 and Additional file 7:
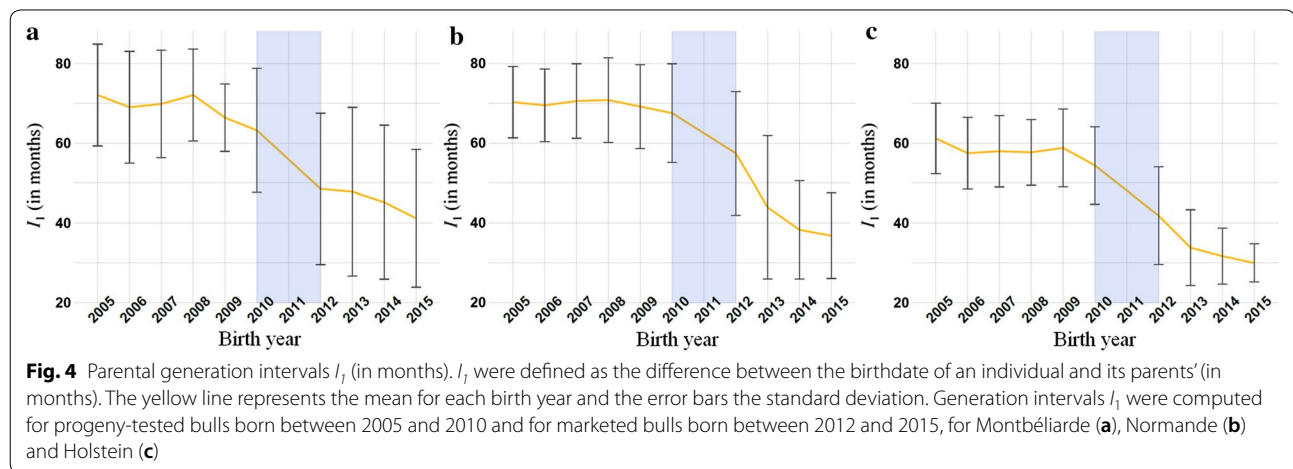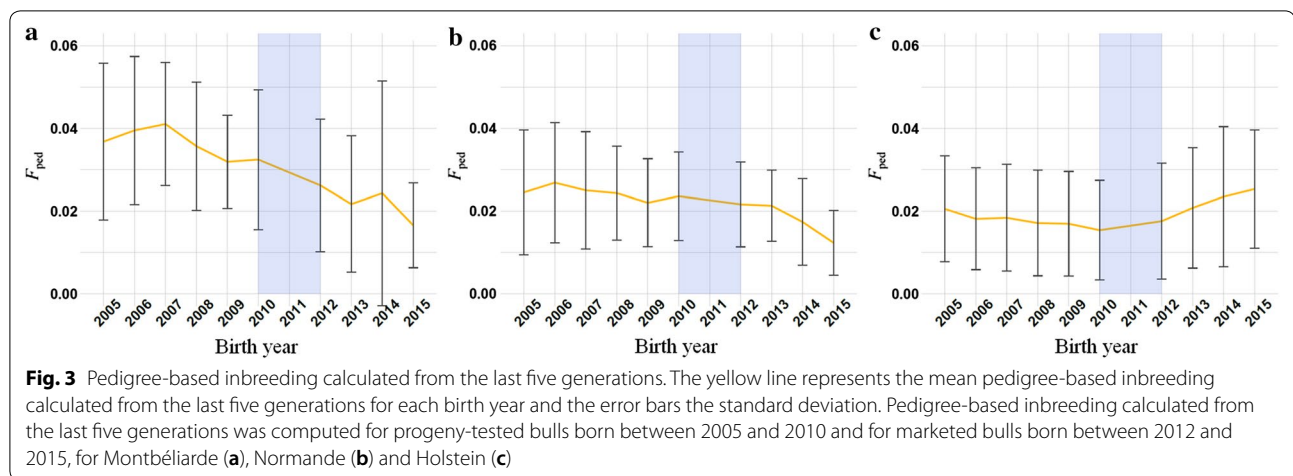
Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 8 of 13

Fig. 3 Pedigree-based inbreeding calculated from the last five generations. The yellow line represents the mean pedigree-based inbreeding calculated from the last five generations for each birth year and the error bars the standard deviation. Pedigree-based inbreeding calculated from the last five generations was computed for progeny-tested bulls born between 2005 and 2010 and for marketed bulls born between 2012 and 2015, for Montbéliarde (**a**), Normande (**b**) and Holstein (**c**)

Fig. 4 Parental generation intervals I_1 (in months). I_1 were defined as the difference between the birthdate of an individual and its parents' (in months). The yellow line represents the mean for each birth year and the error bars the standard deviation. Generation intervals I_1 were computed for progeny-tested bulls born between 2005 and 2010 and for marketed bulls born between 2012 and 2015, for Montbéliarde (**a**), Normande (**b**) and Holstein (**c**)
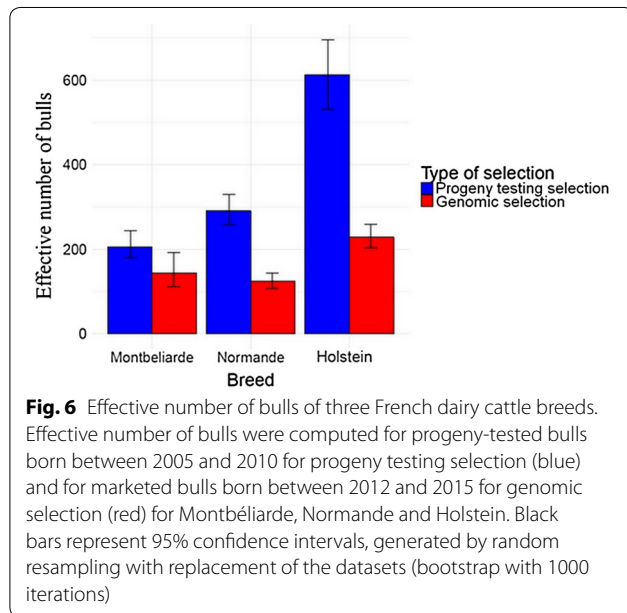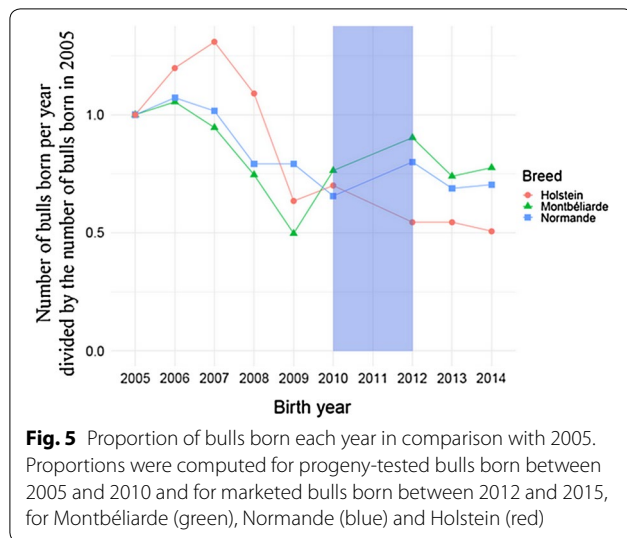
Table S6. The ratio between number of marketed bulls born in 2014 and number of progeny-tested bulls born in 2005 was equal to 77.6% (128 compared to 165) in Montbéliarde, 70.4% (88 compared to 125) in Normande, and 50.7% (294 compared to 580) in Holstein (see Fig. 5).

The effective number of bulls used was significantly larger between 2005 and 2010 (under PTS) than between 2012 and 2014 (under GS) in Normande and Holstein, but not in Montbéliarde (see Fig. 6). The ratio of the effective and the actual number of bulls ratio under PTS, $\frac{Ne_{\text{breed, PTS}}}{n_{\text{breed, PTS}}}$, was significantly lower than the same ratio under GS, $\frac{Ne_{\text{breed, GS}}}{n_{\text{breed, GS}}}$, in Montbéliarde, Normande and Holstein, with $\frac{Ne_{\text{breed, GS}}}{n_{\text{breed, GS}}}$ of 0.009, 0.021 and 0.003%, respectively (see Fig. 7).

Discussion

In dairy cattle breeds, GS was expected to increase the annual genetic gain by up to 70 to 108% depending on the scenarios used for the breeding schemes (with different proportions of young genotypes bulls and proven bulls used) [1, 2, 44] but only a few studies have examined its impact on genetic diversity. These studies were preliminary and mainly predictive since they were conducted before GS had really started. Genomic evaluations were implemented in 2009 in French dairy cattle breeds [4] and they have drastically changed dairy cattle breeding schemes for the three major French dairy cattle breeds. Thus, after a decade of GS, we can evaluate the impact of new breeding schemes on both the genetic gain and diversity in these three breeds in France. The Holstein, Montbéliarde and Normande breeds in France vary in population size with 628,000, 319,000, 2,402,000 cows, respectively, as of January 2018 [45]. These breeds differ also in terms of breeding program: the Holstein breed is widespread and genetic material is exchanged globally (*i.e.* it is an international breed), whereas Montbéliarde and Normande are French typical breeds managed at the country level (*i.e.* national breeds). Thus, the results drawn from our study can compare these two situations.

Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 9 of 13

Fig. 5 Proportion of bulls born each year in comparison with 2005. Proportions were computed for progeny-tested bulls born between 2005 and 2010 and for marketed bulls born between 2012 and 2015, for Montbéliarde (green), Normande (blue) and Holstein (red)

Fig. 7 Effective/census number of bulls ratios of three French dairy cattle breeds. Effective number of bulls/Number of bulls ratios $Ne_{breed, selection} / n_{breed, selection}$ were computed for progeny-tested bulls for individuals born between 2005 and 2010 for progeny testing selection (blue) and for marketed bulls born between 2012 and 2015 for genomic selection (red) for Montbéliarde, Normande and Holstein. Black bars represent 95% confidence intervals, generated by random resampling with replacement of the datasets (bootstrap with 1000 iterations)

Fig. 6 Effective number of bulls of three French dairy cattle breeds. Effective number of bulls were computed for progeny-tested bulls born between 2005 and 2010 for progeny testing selection (blue) and for marketed bulls born between 2012 and 2015 for genomic selection (red) for Montbéliarde, Normande and Holstein. Black bars represent 95% confidence intervals, generated by random resampling with replacement of the datasets (bootstrap with 1000 iterations)

To conduct this study, we focused on bulls of these three breeds born between 2005 and 2010 for PTS, and bulls born between 2012 and 2015 for GS.

An increase in annual genetic gain with GS has been observed in most dairy cattle populations. For instance, in the US Holstein cattle breed, the annual genetic gain for milk, fat and protein yield in cows born before or after 2010 has increased by 71, 111 and 81%, respectively, between 2008 and 2014 [46, 47]. This can be explained by a drastic decrease in the generation interval by ~ 50% on average for Holstein bulls [10, 46]. This effective decrease in generation interval is consistent with previous predictions [9] and is mainly due to the withdrawal of progeny testing prior to marketing of bulls.
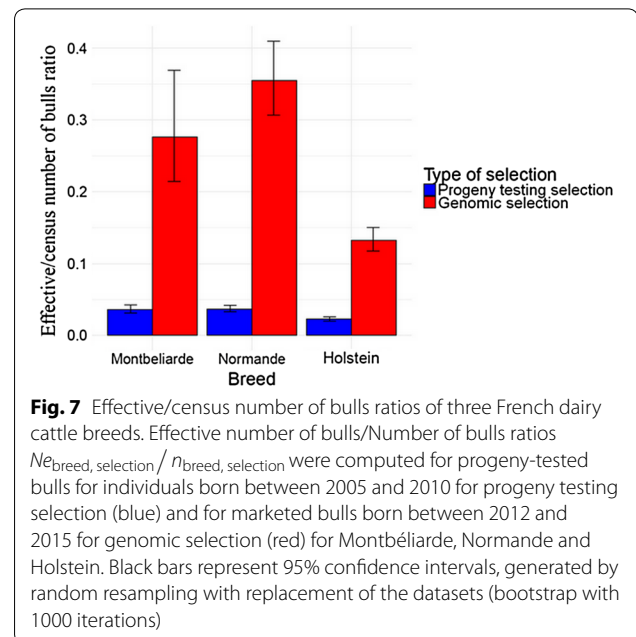
We obtained similar results for Montbéliarde, Normande and Holstein bulls. Their mean annual genetic gains (ΔG) increased by 50, 71 and 33%, respectively, after implementation of GS. At the same time, the bulls' parental generation intervals were reduced by a factor 1.7, 1.9 and 2 for the Montbéliarde, Normande and Holstein breeds, respectively, suggesting that, similar to the US Holstein situation, most of the extra genetic gain is due to reduced generation intervals. However, there are differences between breeds in terms of increase in annual genetic gain and decrease in generation interval, with the Holstein breed having less increase in gain relative to the reduction in generation interval. This may be due to different formulas used for total merit indices and different breeding goals.

Given this global increase in genetic gain for all the breeds studied here along with the decrease in generation interval, one might be concerned about the trend in genetic diversity. Indeed, on the one hand, an accelerated decrease in genetic diversity could be expected due to the decrease in generation interval. On the other hand, genomic evaluations can lead to a more diverse offer in terms of candidates for selection, and thus a moderation of inbreeding rate over time [5]. To study the possible change of inbreeding rate, we first looked at various estimates of inbreeding and diversity to recognize whether and why they differed. ROH-based inbreeding estimates aligned well with pedigree-based inbreeding estimates when averaged per breed. However, pedigree-based inbreeding estimates had a higher coefficient of

Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 10 of 13

variation between bulls within a breed (0.33, 0.25 and 0.31 for Montbéliarde, Normande and Holstein, respectively) than ROH-based inbreeding estimates (0.22, 0.18 and 0.21 for Montbéliarde, Normande and Holstein, respectively), for the three breeds. Pedigree depths were substantial and similar in the three breeds and there was limited variation among individual bulls; the mean number of generations traced for Montbéliarde, Normande and Holstein bulls was 15.4 ± 1.7, 15.8 ± 1.6 and 15.4 ± 1.9, respectively. Therefore, pedigree completeness was not a major cause of the higher coefficient of variation of pedigree-based inbreeding estimates in comparison with ROH-based inbreeding estimates. Pedigree-based estimations of individual inbreeding coefficients are based on expected IBD probabilities and because of Mendelian sampling these values might differ from ROH-based estimates, which reflects more the realized or effective inbreeding [48]. The correlations between pedigree- and ROH-based inbreeding estimates ranged from 50 to 59%, depending on the breed. These correlations were in the range of those previously reported for Jersey and Danish red cattle (52 and 54%, respectively) [49] but lower than those reported for Holstein (around 80%) [49] and simulated cattle (around 70%) [11]. Errors and depth of pedigree are very plausible reasons for these differences between reported and presented correlations. Moreover, such differences in these correlations are expected because pedigree-based inbreeding corresponds to the expected value of inbreeding while ROH-based inbreeding corresponds to its realized value. In groups of highly selected bulls, linkage disequilibrium and homozygosity can be strong, which leads to large discrepancies between expected (based on pedigree) and observed (based on ROH) inbreeding and subsequently lower correlations.

In all three breeds, the ratio between the effective and actual number of bulls (Ne/N) increased after implementation of GS. This reflects changes in breeding schemes in all three breeds, which should allow for a decrease in inbreeding rate and more generally for a slowdown in the annual loss of genetic diversity. However, we distinguished two contrasting types of behavior of inbreeding based on ROH. First, for the two national breeds, Montbéliarde and Normande, we did not detect any significant change in the annual inbreeding rate (ΔF). Second, for the Holstein international breed there was a significant increase in ΔF, which has increased from 0.080% to 0.55% per year based on ROH, and from 0.088% to 0.49% per year based on pedigree. This result does not align with the fact that all three breeds experienced a significant increase in annual genetic gain, with the lowest increase observed in Holstein while the inbreeding rate

in Holstein was highest. This result suggests that changes in breeding schemes reflected by an increase of Ne/N of bulls in Holstein did not allow for a slowdown in the annual loss of genetic diversity.

For Holstein, ΔF per generation also increased from 0.39% to 1.36% based on ROH, and from around 0.46% to 1.39% based on pedigree. Thus, the decrease in generation interval did not explain all the increase in ΔF per year for Holstein, since ΔF per generation increased too. The observed inbreeding rate in Holstein is higher than the 1% per generation acceptable rate of inbreeding according to FAO guidelines [50]. Since the accumulation of inbreeding through ROH is not linear, and the observed level and rate of inbreeding could be underestimated when based on ROH, the inbreeding problem is likely even more severe than suggested by these figures.

We used a linear regression to estimate the rate of ROH-based inbreeding because it gave better results than a logarithmic regression in terms of residuals of the model (see Additional file 4: Figures S1–S7). When estimated from pedigree data, the slope of inbreeding increased significantly for both Normande and Holstein bulls, reaching 0.19% and 0.49% per year, respectively. The inbreeding rate reached 0.70% and 1.39% per generation, respectively, so the decrease in generation interval did not explain all the increase in inbreeding rate per year. Thus, the Normande breed remained under the recommended FAO threshold whereas Holstein exceeded this when estimating inbreeding on pedigree data as well as on ROH data. Why there are such differences between these two breeds should be further investigated, and in particular, the differences in breeding schemes, such as the use of foreign bulls or the setting-up of quotas for the number of inseminations allowed per bull.

From 2005 to 2015, the overall mean ROH length for Montbéliarde (1.68 Mb) was significantly shorter than that for Normande (1.71 Mb), which in turn was significantly shorter than that for Holstein (1.80 Mb). The length of ROH depends on two main criteria, for given detection parameters and genotyping density: (i) the effective population size and the level of linkage disequilibrium [51]; and (ii) the number of generations that separates an individual from the ancestor from which the ROH originated, with longer ROH reflecting more recent inbreeding [28, 52]. We investigate these two possibilities to explain the differences in mean ROH lengths between breeds.

Estimated effective population sizes (Ne) of French dairy cattle breeds were equal to 79 for Montbéliarde, 87 for Normande and 96 for Holstein in 2017 [53]. Based on these results, effective/census population sizes ratios

Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 11 of 13

(Ne/N) of 0.009, 0.021 and 0.003% for these breeds, respectively, were calculated [53]. The effective number of bulls born under GS for the period between 2012 and 2015 was estimated to be 144 for Montbéliarde, 124 for Normande and 229 for Holstein, and effective/census number of bull ratios were 28, 35 and 13%, respectively. In both cases, the ratios between effective and census effective population sizes were lowest for Holstein and higher for Normande and Montbéliarde. We can hypothesize that these differences are probably linked to the small numbers of bulls that have strong contributions and the low ratios between effective and census number of bulls, particularly in Holstein [54]. Moreover, whereas the distribution of the number of offspring per bull became less skewed for the three breeds under GS than under PTS, this change seemed to be less pronounced in Holstein (see Additional file 6: Figure S8, Additional file 7: Table S6). This was consistent with a relatively low ratio between effective and census number of bulls in Holstein under GS in comparison with Montbéliarde and Normande, and with the hypothesis of a remaining effect of a few bulls with strong contributions. There can be two effects explaining the greater observed ROH length for Holstein. The first is the small effective population size relative to census population size resulting in a highly structured genome in this breed with long haploblocks [55] where recombination is unlikely to have occurred leading to long ROH. The second is recent inbreeding, likely to have occurred because of the massive use of few elite bulls. This was suggested by kinship estimated from pedigree being higher for the period between 2012 and 2015 than for the period between 2005 and 2010 for Holstein and not for national breeds, which can be explained by the intensive use of a few individuals as grandfathers of bulls, mostly coming from the USA [56]. We chose not to compute kinship from GRM because of their high sensitivity to the estimation of allelic frequencies (data not shown) [21]. The hypothesis that a few individuals were used as grandfathers of bulls in Holstein was consistent with the increase by a factor 8.77 (with a relative change of 7.77) of the annual rate of mean ROH length in Holstein breed after the beginning of GS, whereas it remained unchanged in the Montbéliarde and Normande breeds. This increase might be counterintuitive. Indeed, GS may lead to relatively shorter ROH than PTS [10]. Therefore, the increase in mean ROH length in Holstein might be underestimated and was corroborated by the increase in the number of long ROH per individual in Holstein with the beginning of GS, whereas it was not the case in Montbéliarde and Normande (see Additional file 8:

Table S7, Additional file 9: Figures S9–S11]). Therefore, the increase in the number of long ROH per individual in Holstein might be responsible for the increased mean ROH length within the breed. Shorter ROH tend to result from older inbreeding (older common ancestors) whereas longer ROH tend to result from more recent inbreeding (more recent common ancestors) [29, 31, 32]. Thus, the increase in the mean ROH length and in the number of long ROH in Holstein suggests an increase in the number of ROH segments arising from recent inbreeding. This introduction of recent inbreeding was confirmed by the observed increase in pedigree-based inbreeding calculated from the last five generations that started when GS was implemented, with an inbreeding rate ranging from -0.083% per year for PTS to 0.27% per year for GS (-0.40% and 0.78% per generation, respectively). In conclusion, in Holstein, this increase in the mean ROH length might reflect the substantial introduction of recent inbreeding, but also a reduction in the number of selected bulls.

Conclusions

After the introduction of genomic selection, the annual genetic gain increased based on estimated breeding values for bulls of three large French dairy cattle breeds: Montbéliarde, Normande and Holstein. At the same time, the inbreeding rate per year and per generation increased in Holstein bulls, an international breed, but not in Montbéliarde and Normande bulls, two national breeds. Small effective population sizes resulted in highly structured genomes with long haploblocks, particularly in Holstein, as a consequence of the low effective/census number of bulls ratios. A joint increase in genetic gain and conservation of genetic diversity is possible by (i) applying optimal contribution selection; (ii) carrying out on-farm mating plans; and (iii) ensuring that these tools are flexible enough to be used effectively in the field. Our study highlights the fact that changes in breeding schemes can impact genetic diversity, and thus, it is important to continue monitoring genetic diversity and to study the future impact of different breeding schemes strategies on genetic diversity. The recent implementation of genomic evaluations in small regional breeds, for which effective population size can be limited, should be studied carefully in order to ensure sustainability of breeding schemes in the future, in terms of both genetic gain and diversity. The spread of new technologies in breeding schemes, such as female reproductive technologies and embryo transfer that might increase global selection intensity through the dams of bulls, represents a risk in terms of genetic diversity that should be anticipated in the light of our conclusions.

Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 12 of 13

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12711-019-0495-1.

Additional file 1: Table S1. Description of the three studied dairy cattle breeds: Montbéliarde, Normande and Holstein.

Additional file 2: Table S2. Length of the autosomal genome in three French dairy cattle breeds. Length of autosomal genome on which ROH can be detected corresponds to the length of the autosomal genome covered by SNPs after withholding gaps longer than 150 kb between two SNPs. The proportion of autosomal genome on which ROH can be detected is the length of autosomal genome on which ROH can be detected divided by the total length of the autosomal genome (from *Bos taurus* assembly 3.1.1).

Additional file 3: Table S3. Parental generation intervals. Mean parental generation intervals for bulls born between 2005 and for bulls born between 2010 and 2015 in months and in years, for Montbéliarde, Normande and Holstein.

Additional file 4: Figure S1. Square root of standardized residuals for the linear regression of pedigree-based inbreeding depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S2.** Square root of standardized residuals for the linear regression of pedigree-based inbreeding for the last five generations depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S3.** Square root of standardized residuals for the linear regression of pedigree-based kinship depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S4.** Square root of standardized residuals for the linear regression of F_{ROH}, inbreeding based on ROH depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S5.** Square root of standardized residuals for the logarithmic regression of F_{ROH}, inbreeding based on ROH depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right) ($F_{ROH, transformed} = \log10(1 - F_{ROH})$). **Figure S6.** Square root of standardized residuals for the linear regression of the mean length of ROH depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right). **Figure S7.** Square root of standardized residuals for the linear regression of total merit index (ISU) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right).

Additional file 5: Table S4. Pedigree-based inbreeding calculated from the last five generations in percent per year. **Table S5.** Estimation of the slopes of pedigree-based inbreeding calculated from the last five generations in percent per generation.

Additional file 6: Figure S8. Proportion of the total number of offspring per bull for each breed and selection type. Progeny testing selection corresponds to bulls born between 2005 and 2010 and genomic selection to bulls born between 2012 and 2014.

Additional file 7: Table S6. Number of bulls and offspring for each birth period in the dataset.

Additional file 8: Table S7. ROH categories lengths thresholds.

Additional file 9: Figure S9. Confidence intervals of the number of ROH per ROH length category per selection type for Montbéliarde bulls. **Figure S10.** Confidence intervals of the number of ROH per ROH length category per selection type for Normande bulls. **Figure S11.** Confidence intervals of the number of ROH per ROH length category per selection type for Holstein bulls.

Authors' contributions
PC and GR conceived and designed the research. CH provided data. ACD performed the statistical analysis and drafted the manuscript. PC, DL and GR supervised the statistical analyses. ACD, PC, DL, AM, CH, SF and GR contributed to the interpretation of the results. ACD, PC, DL, AM, CH, CDB, SF and GR contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Author details
[1] GABI, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France. [2] ALLICE, Paris, France. [3] Institut de l'Elevage, Paris, France.

References
1. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: genomic selection in dairy cattle: progress and challenges. J Dairy Sci. 2009;92:433–43.
2. de Roos APW, Schrooten C, Veerkamp RF, van Arendonk JAM. Effects of genomic selection on genetic improvement, inbreeding, and merit of young versus proven bulls. J Dairy Sci. 2011;94:1559–67.
3. Wang J. Pedigrees or markers: which are better in estimating relatedness and inbreeding coefficient? Theor Popul Biol. 2016;107:4–13.
4. Boichard D, Guillaume F, Baur A, Croiseau P, Rossignol MN, Boscher MY, et al. Genomic selection in French dairy cattle. Anim Prod Sci. 2012;52:115–20.
5. Daetwyler HD, Villanueva B, Bijma P, Woolliams JA. Inbreeding in genome-wide selection. J Anim Breed Genet. 2007;124:369–76.
6. Meuwissen THE. Maximizing the response of selection with a predefined rate of inbreeding. J Anim Sci. 1997;75:934–40.
7. Engelsma KA, Veerkamp RF, Calus MPL, Windig JJ. Consequences for diversity when prioritizing animals for conservation with pedigree or genomic information: prioritization with pedigree or genomic data. J Anim Breed Genet. 2011;128:473–81.
8. Eynard SE, Windig JJ, Hiemstra SJ, Calus MPL. Whole-genome sequence data uncover loss of genetic diversity due to selection. Genet Sel Evol. 2016;48:33.
9. Schaeffer LR. Strategy for applying genome-wide selection in dairy cattle. J Anim Breed Genet. 2006;123:218–23.
10. Doekes HP, Veerkamp RF, Bijma P, Hiemstra SJ, Windig JJ. Trends in genome-wide and region-specific genetic diversity in the Dutch–Flemish Holstein–Friesian breeding program from 1986 to 2015. Genet Sel Evol. 2018;50:15.
11. Forutan M, Ansari Mahyari S, Baes C, Melzer N, Schenkel FS, Sargolzaei M. Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. BMC Genomics. 2018;19:98.
12. Dickerson GE, Hazel LN. Effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. J Agric Res. 1944;69:459–76.
13. Barrett RDH, Schluter D. Adaptation from standing genetic variation. Trends Ecol Evol. 2008;23:38–44.
14. Leroy G. Inbreeding depression in livestock species: review and meta-analysis. Anim Genet. 2014;45:618–28.
15. Pryce JE, Haile-Mariam M, Goddard ME, Hayes BJ. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genet Sel Evol. 2014;46:71.
16. Reed DH, Frankham R. Correlation between fitness and genetic diversity. Conserv Biol. 2003;17:230–7.

Doublet *et al. Genet Sel Evol*   (2019) 51:52

Page 13 of 13

17. Meuwissen THE, Luo Z. Computing inbreeding coefficients in large populations. Genet Sel Evol. 1992;24:305–13.
18. Danchin-Burge C, François L, Laloë D, Leroy G, Verrier E. Using the information collected for genetic evaluation to assess the French ruminant and equine within-breeds genetic variability. Interbull Bull. 2013;47:111–9.
19. Boichard D, Maignel L, Verrier E. The value of using probabilities of gene origin to measure genetic variability in a population. Genet Sel Evol. 1997;29:5–23.
20. Oliehoek PA, Bijma P. Effects of pedigree errors on the efficiency of conservation decisions. Genet Sel Evol. 2009;41:9.
21. Eynard SE, Windig JJ, Leroy G, van Binsbergen R, Calus MP. The effect of rare alleles on estimated genomic relationships from whole genome sequence data. BMC Genet. 2015;16:24.
22. Eynard SE, Windig JJ, Leroy G, Verrier E, Hiemstra SJ, Van Binsbergen R, et al. The use of whole genome sequence data to estimate genetic relationships including rare alleles information. In: Proceedings of the 10th world congress of genetics applied to livestock production: 17–22 August 2014; Vancouver. 2014.
23. de Cara MÁR, Villanueva B, Toro MÁ, Fernández J. Using genomic tools to maintain diversity and fitness in conservation programmes. Mol Ecol. 2013;22:6091–9.
24. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L, et al. Runs of homozygosity in European populations. Am J Hum Genet. 2008;83:359–72.
25. Curik I, Ferenčaković M, Sölkner J. Inbreeding and runs of homozygosity: a possible solution to an old problem. Livest Sci. 2014;166:26–34.
26. Ferenčaković M, Hamzić E, Gredler B, Solberg TR, Klemetsdal G, Curik I, et al. Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. J Anim Breed Genet. 2013;130:286–93.
27. Keller MC, Visscher PM, Goddard ME. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics. 2011;189:237–49.
28. Howrigan DP, Simonson MA, Keller MC. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. BMC Genomics. 2011;12:460.
29. Purfield DC, Berry DP, McParland S, Bradley DG. Runs of homozygosity and population history in cattle. BMC Genet. 2012;13:70.
30. Ferenčaković M, Sölkner J, Curik I. Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. Genet Sel Evol. 2013;45:42.
31. Ceballos FC, Joshi PK, Clark DW, Ramsay M, Wilson JF. Runs of homozygosity: windows into population history and trait architecture. Nat Rev Genet. 2018;19:220–34.
32. Kirin M, McQuillan R, Franklin CS, Campbell H, McKeigue PM, Wilson JF. Genomic runs of homozygosity record population history and consanguinity. PLoS One. 2010;5:e13996.
33. Bosse M, Megens HJ, Madsen O, Crooijmans RPMA, Ryder OA, Austerlitz F, et al. Using genome-wide measures of coancestry to maintain diversity and fitness in endangered and domestic pig populations. Genome Res. 2015;25:970–81.
34. Gómez-Romano F, Villanueva B, Fernández J, Woolliams JA, Pong-Wong R. The use of genomic coancestry matrices in the optimisation of contributions to maintain genetic diversity at specific regions of the genome. Genet Sel Evol. 2016;48:2.
35. Sonesson AK, Woolliams JA, Meuwissen TH. Genomic selection requires genomic control of inbreeding. Genet Sel Evol. 2012;44:27.
36. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2018. https://www.R-project.org/.
37. Coster A. Pedigree: Pedigree functions. 2012. https://CRAN.R-project.org/package=pedigree.
38. Launay A, Rocland M. Genetic evaluation—methods and definitions. http://idele.fr/no_cache/recherche/publication/idelesolr/recommends /genetic-evaluation-methods-and-definitions.html. Accessed 17 Dec 2018.
39. Boichard D. Pedig: a fortran package for pedigree analysis suited for large populations. 2002. https://prodinra.inra.fr/?locale=fr#!ConsultNotice:7255. Accessed 3 Oct 2018.
40. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.
41. Purcell S, Chang C. PLINK 1.9. http://www.cog-genomics.org/plink/1.9/. Accessed 3 Oct 2018.
42. Tarpy DR, Nielsen DI. Sampling error, effective paternity, and estimating the genetic structure of Honey bee colonies (Hymenoptera: Apidae). Ann Entomol Soc Am. 2002;95:513–28.
43. Starr CK. Sperm competition, kinship, and sociality in the Aculeate Hymenoptera. In: Smith RL, editor. Sperm competition and the evolution of animal mating systems. New York: Academic Press; 1984. p. 427–64.
44. Colleau JJ, Fritz S, Guillaume F, Baur A, Dupassieux D, Boscher MY, et al. Simulation des potentialités de la sélection génomique chez les bovins laitiers. INRA Prod Anim. 2015;28:251–8.
45. Groupe Economie du Bétail. Chiffres-clés Bovins 2018—Productions Bovines lait et viande. http://idele.fr/filieres/publication/idelesolr/recommends/chiffres-cles-bovins-2018.html. Accessed 22 Nov 2018.
46. García-Ruiz A, Cole JB, VanRaden PM, Wiggans GR, Ruiz-López FJ, Van Tassell CP. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. Proc Natl Acad Sci USA. 2016;113:E3995–4004.
47. Howard JT, Pryce JE, Baes C, Maltecca C. Invited review: inbreeding in the genomics era: Inbreeding, inbreeding depression, and management of genomic variability. J Dairy Sci. 2017;100:6009–24.
48. Hill WG, Weir BS. Variation in actual relationship as a consequence of Mendelian sampling and linkage. Genet Res (Camb). 2011;93:47–64.
49. Zhang Q, Calus MP, Guldbrandtsen B, Lund MS, Sahana G. Estimation of inbreeding using pedigree, 50 k SNP chip genotypes and full sequence data in three cattle breeds. BMC Genet. 2015;16:88.
50. FAO. Secondary guidelines for development of national farm animal genetic resources management plans: Management of small populations at risk. Rome: FAO and UNEP; 2000. p. 219.
51. Waples RS, England PR. Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. Genetics. 2011;189:633–44.
52. Gibson J, Morton NE, Collins A. Extended tracts of homozygosity in outbred human populations. Hum Mol Genet. 2006;15:789–95.
53. Danchin-Burge C. Indicateurs de variabilité génétique—races bovines—Edition 2018. http://idele.fr/no_cache/recherche/publication/idelesolr/recommends/indicateurs-de-variabilite-genetique-races-bovines-edition-2018-1.html. Accessed 22 Nov 2018.
54. Boichard D, Ducrocq V, Fritz S, Colleau JJ. Where is dairy cattle breeding going? A vision of the future. Bull Int Bull Eval Serv. 2010;41:63–8.
55. Jónás D, Ducrocq V, Croiseau P. Short communication: the combined use of linkage disequilibrium–based haploblocks and allele frequency–based haplotype selection methods enhances genomic evaluation accuracy in dairy cattle. J Dairy Sci. 2017;100:2905–8.
56. Le Mézec P, Danchin-Burge C. Davantage de diversité avec la génomique ? plutôt non… http://idele.fr/contact/publication/idelesolr/recommends/les-programmes-de-selection-et-de-diffusion-de-taureaux-dia-a-lere-de-la-genomique-et-leurs-effets.html. Accessed 21 Jan 2019.

Publisher's Note

Additional file 1 Table S1: Description of the three studied dairy cattle breeds: Montbéliarde, Normande and Holstein

Breed	Montbéliarde	Normande	Holstein	Reference
Number of cows in France (on 2018, January 1)	628	319	2,402,000	Groupe Economie du Bétail, 2018
Number of milk-recorded cows in France (in 2017)	430,186	194,524	1,572,614	Thomas et al., 2018
Total number of AI performed by pure bred bulls in France (in 2017)	723,782	460,32	3,581,695	Moureaux and Le Mezec, 2018
Number of first AI performed by pure bred bulls in France (in 2017)	442,235	262,594	1,982,418	Moureaux and Le Mezec, 2018
Total number of pure bred bulls (in 2017)	1,061	725	4,38	Moureaux and Le Mezec, 2018
Mean milk yield per cow per lactation in France in 2017 (in kg)	6,933	6,59	9,241	Thomas et al., 2018
Fat/protein content in milk in 2017 (in g/kg)	38.9/33.1	42.4/34.7	39.7/32.0	Thomas et al., 2018
Estimated number of effective ancestors for the females born between 2014 and 2017	18	21	20	Danchin-Burge, 2018
Mean inbreeding for the females born in France between 2014 and 2017 (in %)	4.8	5.0	4.4	Danchin-Burge, 2018

Danchin-Burge, C., 2018. Indicateurs de variabilité génétique - races bovines - Edition 2018.

Groupe Economie du Bétail, 2018. Chiffres-clés Bovins 2018 - Productions lait et viande.

Moureaux, S., Le Mezec, P., 2018. BGIAP 2017 Bilan Génétique de l'Insémination en races bovines laitières.

Thomas, G., Bourrigan, X., Astruc, J.-M., 2018. Résultats de Contrôle Laitier - France 2017.

Additional file 2 Table S2: Length of the autosomal genome in three French dairy cattle breeds

Length of autosomal genome on which ROH can be detected corresponds to the length of the autosomal genome covered by SNPs after withholding gaps longer than 150kb between two SNPs. The proportion of autosomal genome on which ROH can be detected is the length of autosomal genome on which ROH can be detected divided by the total length of the autosomal genome (from *Bos taurus* assembly 3.1.1).

Breed	Total length of the autosomal genome (in Mb)	Length of autosomal genome on which ROH can be detected (in Mb)	Proportion of autosomal genome on which ROH can be detected
Montbéliarde	2,493.12	1,919.31	77.0%
Normande	2,493.12	1,919.65	77.0%
Holstein	2,493.12	1,959.21	78.6%

Additional file 3 Table S3: Parental generation intervals

Breed	Mean generation interval between 2005 and 2010		Mean generation interval between 2012 and 2015	
	In months	In years	In months	In years
Montbéliarde	69.1	5.76	46.2	3.85
Normande	69.8	5.81	44.9	3.74
Holstein	58.0	4.84	34.3	2.86

Additional file 4: Residuals of regressions of genetic diversity and genetic gain parameters depending on birth year and selection type.

Figure S1: Square root of standardized residuals for the linear regression of pedigree-based inbreeding depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)
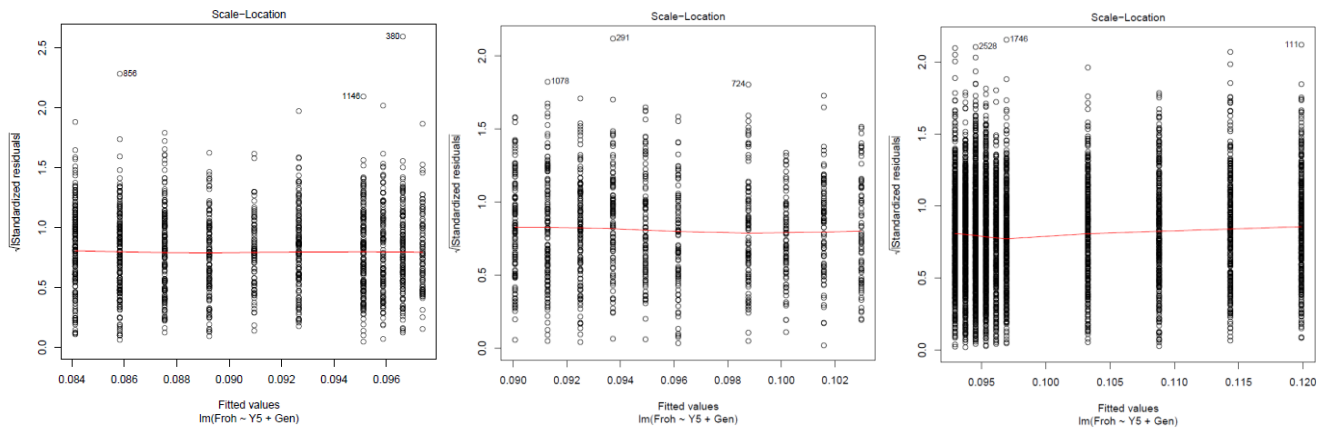

Figure S2: Square root of standardized residuals for the linear regression of pedigree-based inbreeding for the last five generations depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

Figure S3: Square root of standardized residuals for the linear regression of pedigree-based kinship depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)
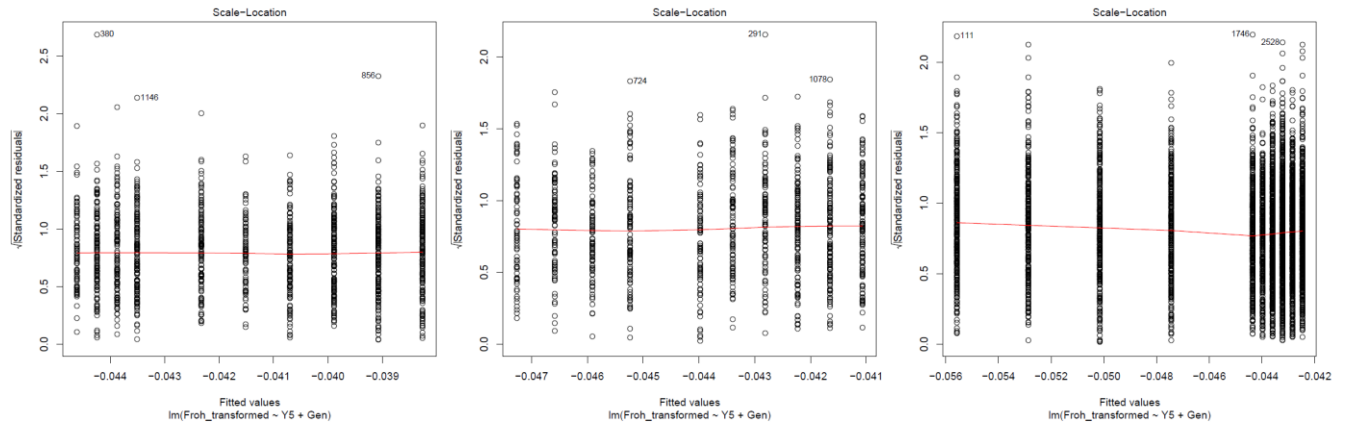

Figure S4: Square root of standardized residuals for the linear regression of F_{ROH}, inbreeding based on Runs of Homozygosity (ROH) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

Figure S5: Square root of standardized residuals for the logarithmic regression of F_{ROH}, inbreeding based on Runs of Homozygosity (ROH) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)
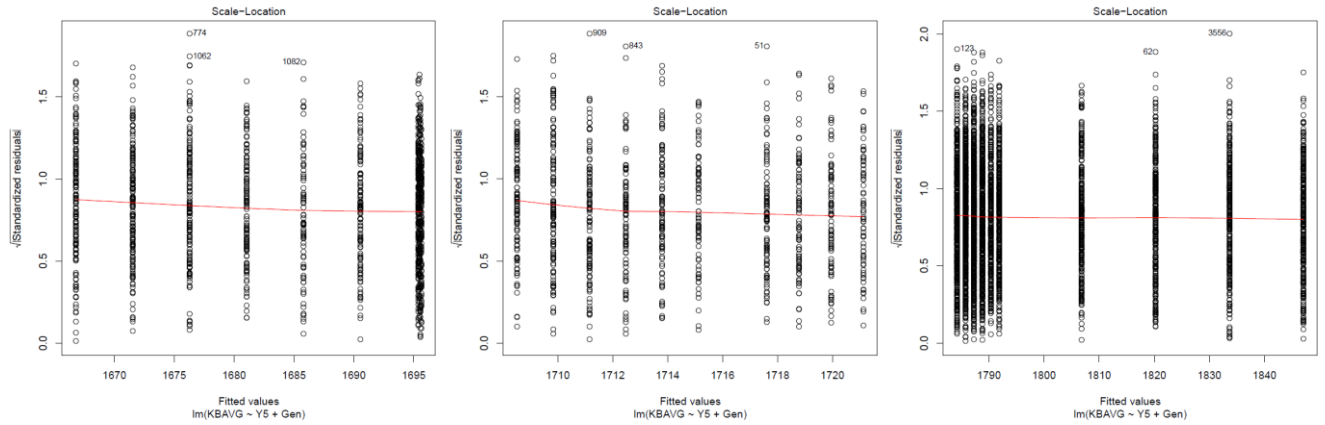
$F_{ROH, transformed} = \log 10(1 - F_{ROH})$

Figure S6: Square root of standardized residuals for the linear regression of the mean length of Runs of Homozygosity (ROH) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)
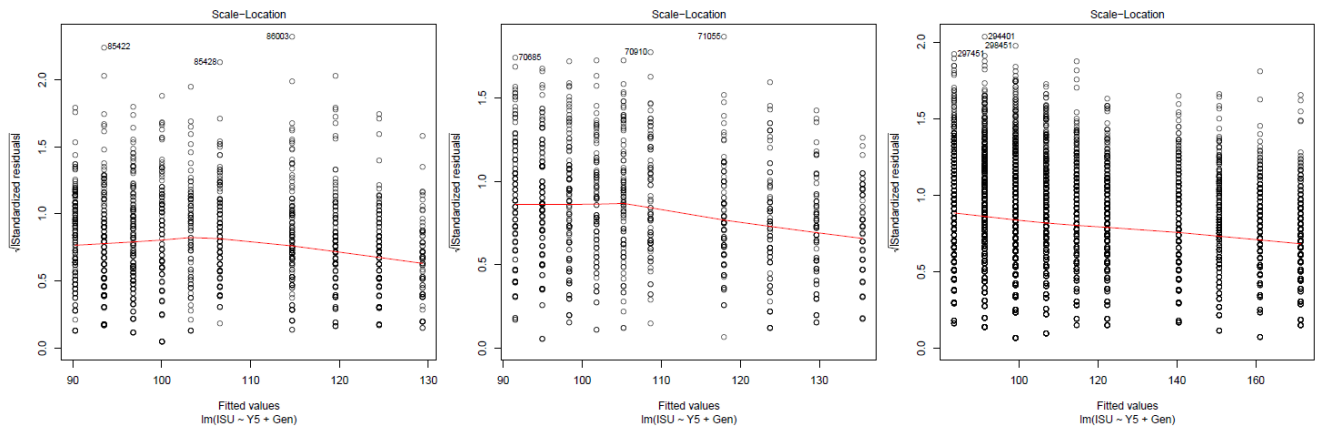

Figure S7: Square root of standardized residuals for the linear regression of Total Merit Index (ISU) depending on birth year (Y5) and selection type (Gen) for Montbéliarde, Normande and Holstein (from left to right)

Additional file 5: Pedigree-based inbreeding calculated from the last five generations per year and per generation.

Table S4: Pedigree-based inbreeding calculated from the last five generations in percent per year.

Breed	b_1 (± standard error)	p-value of b_1	δ (± standard error)	p-value of δ	b_2	Relative change
Montbéliarde	-0.16 (± 0.030)	< 2e-16	-0.16 (± 0.077)	0.039	-0.32	-1.00
Normande	-0.030 (± 0.024)	< 2e-16	-0.20 (± 0.059)	7.4e-04	-0.23	-6.65
Holstein	-0.083 (± 0.013)	9.6e-11	0.36 (± 0.032)	< 2.2e-16	0.27	4.30

b_1 is the slope of regression of pedigree-based inbreeding calculated from the last five generations depending on birth year for progeny tested bulls born between 2005 and 2010 (progeny testing selection), and δ the difference between the slopes of regression of pedigree-based inbreeding calculated from the last five generations depending on birth year for progeny testing selection and for marketed bulls born between 2012 and 2015 (genomic selection), for Montbéliarde, Normande and Holstein. The relative change is equal to $\frac{\delta}{|b_1|}$. b_2 is the slope of the pedigree-based inbreeding for the last five generations for marketed bulls born between 2012 and 2015 (genomic selection), equal to $b_1 + \delta$. The p-value of δ corresponds to the significance of the non-nullity of δ.

Table S5: Estimation of the slopes of pedigree-based inbreeding calculated from the last five generations in percent per generation.

Breed	b_1 per generation	δ per generation	b_2 per generation	Relative change per generation
Montbéliarde	-0.91%	-3.1e-03	**-1.22%**	-0.34
Normande	-0.18%	-6.9e-03	**-0.86%**	-3.94
Holstein	-0.40%	1.2e-02	**0.78%**	2.95

Additional file 6: Proportion of the total number of offspring per bull for each breed and selection type.
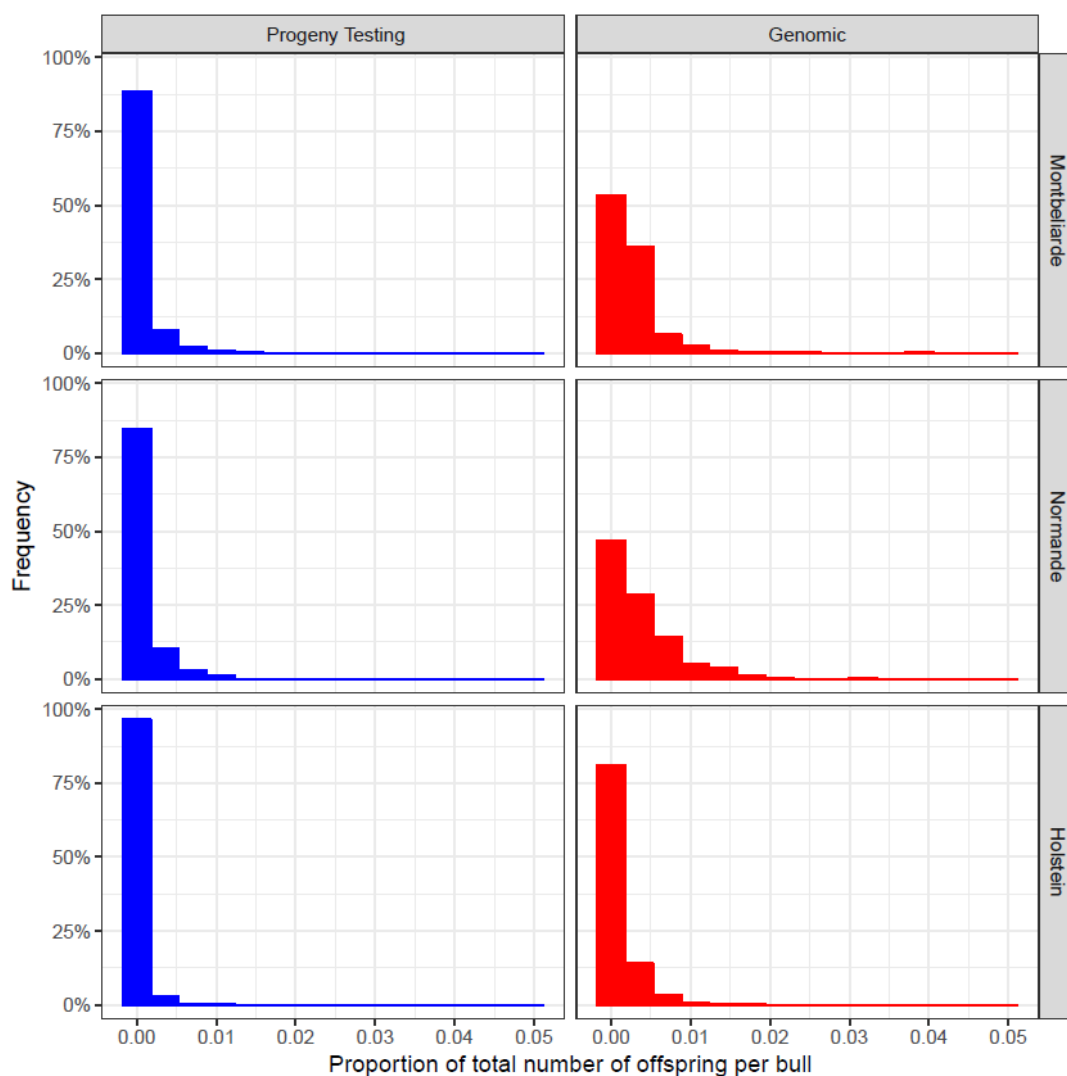

Figure S8: Proportion of the total number of offspring per bull for each breed and selection type

Progeny testing selection corresponds to bulls born between 2005 and 2010 and genomic selection to bulls born between 2012 and 2014.

For example, for Montbéliarde, around 50% of the bulls born under genomic selection (y-axis) are the fathers of less than 0.003 = 0.3% of all calves (x-axis) and around 30% of the bulls are the fathers of 0.3% to 0.6% of all calves.

Additional file 7: Number of bulls and offspring for each birth period in the dataset.

Table S6: Number of bulls and offspring for each birth period in the dataset.

Breed	Birth period	Total number of individuals	Total number of offspring
Montbéliarde	2005 – 2010 Progeny-tested bulls	826	1,102,266
	2012 – 2014 Marketed bulls	399	562,178
Normande	2005 – 2010 Progeny-tested bulls	666	542,939
	2012 – 2014 Marketed bulls	347	437,558
Holstein	2005 – 2010 Progeny-tested bulls	3440	4,573,893
	2012 – 2014 Marketed bulls	1248	2,390,778

Additional file 8: ROH categories lengths thresholds

For each breed, we selected bulls born in 2005, obtaining three cohorts: Montbéliarde bulls born in 2005, Normande bulls born in 2005 and Holstein bulls born in 2005. For each of these three cohorts, five ROH length categories were defined with R function kmeans (R Core Team, 2018) by applying it to the list of all ROH for all individuals of each cohort, for each breed separately. We obtained length thresholds for each category.

Table S7: ROH categories lengths thresholds

Category	Montbéliarde		Normande		Holstein	
	Minimum ROH length (kb)	Maximum ROH length (kb)	Minimum ROH length (kb)	Maximum ROH length (kb)	Minimum ROH length (kb)	Maximum ROH length (kb)
1	1000.079	1426.286	1000.051	1458.437	1000.079	1465.802
2	1426.639	2080.933	1459.128	2121.894	1466.58	2163.376
3	2084.642	3046.285	2127.988	3087.975	2164.072	3247.136
4	3055.704	4723.705	3091.578	4730.121	3249.867	5031.288
5	4749.088	8718.715	4761.585	8217.436	5040.034	9126.138

REFERENCES

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Additional file 9: Number of ROH in different length categories

For each breed, we selected bulls born in 2005, obtaining three cohorts: Montbéliarde bulls born in 2005, Normande bulls born in 2005 and Holstein bulls born in 2005. For each of these three cohorts, five ROH length categories were defined with R function kmeans (R Core Team, 2018) by applying it to the list of all ROH for all individuals of each cohort, for each breed separately. We obtained length thresholds for each category [see Additional file 8 Table S7]. Then, we counted how many ROH were in each category for each individual. This allowed computing the mean number of ROH per individual per category for all individuals born during progeny testing selection (between 2005 and 2010) or during genomic selection (between 2012 and 2015) for each breed separately. We then performed an anova to compare the number of ROH in each category between progeny testing and genomic selection and estimated confidence intervals of these numbers with R function emmeans (Russell, 2018).
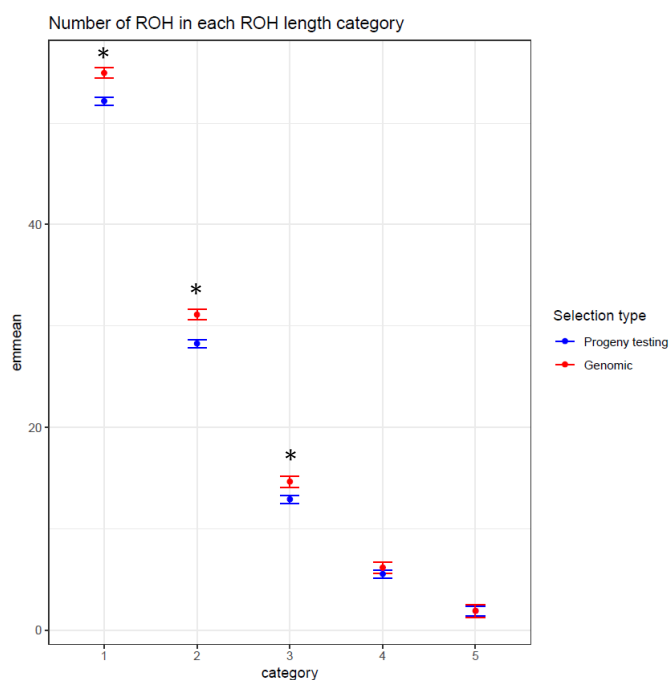

Figure S9: Confidence intervals of the number of ROH per ROH length category per selection type for Montbéliarde bulls. *: mean ROH numbers were significantly different between progeny testing and genomic selection for this category.
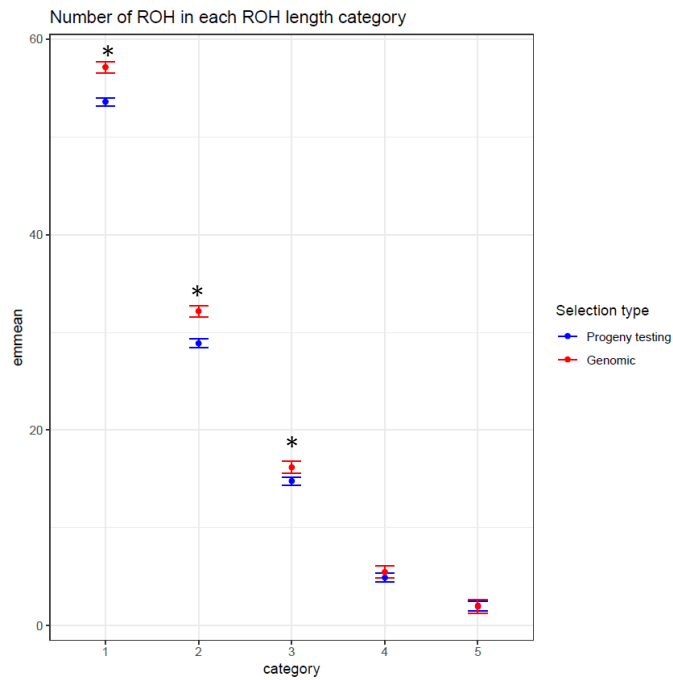
Figure S10: Confidence intervals of the number of ROH per ROH length category per selection type for Normande bulls. *: mean ROH numbers were significantly different between progeny testing and genomic selection for this category.
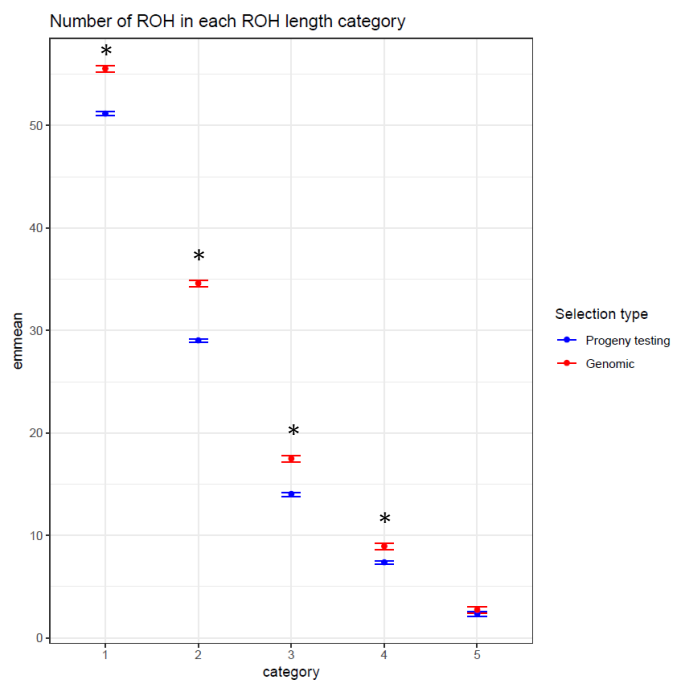

Figure S11: Confidence intervals of the number of ROH per ROH length category per selection type for Holstein bulls. *: mean ROH numbers were significantly different between progeny testing and genomic selection for this category.

In conclusion, small (categories 1 and 2) and medium length (category 3) ROH were significantly more numerous during genomic selection than during progeny testing selection for all three breeds. However, long ROH (category 4, between 3.25 and 5.03 Mb) were more significantly numerous during genomic selection than during progeny testing selection only for Holstein.

REFERENCES

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Russell Lenth (2018). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.3.0. https://CRAN.R-project.org/package=emmeans

III - Discussion

Les évaluations génomiques permettent d'augmenter le nombre de candidats à la sélection voire le nombre de taureaux diffusés, avec un renouvellement plus rapide du catalogue des taureaux diffusés. Cela devrait permettre de tirer parti des avantages des évaluations génomiques sur le plan du progrès génétique tout en ménageant la diversité génétique. Or, si le nombre de taureaux ou de candidats à la sélection n'augmente pas voire diminue, ou si les taureaux sélectionnés sont plus apparentés qu'auparavant, l'effet d'accélération de la perte de diversité génétique dû au raccourcissement important des intervalles de génération n'est plus compensé et la perte de diversité génétique s'accélère, comme nous l'avons observé dans la race Prim'Holstein.

La petite taille efficace des populations a entraîné une forte structuration de leurs génomes avec de l'apparition de longs blocs haplotypiques, en particulier chez la Prim'Holstein. C'est le résultat d'une intensité de sélection élevée (faible nombre de taureaux), de la consanguinité qui en résulte ainsi que de la dérive génétique, d'autant plus forte que la taille efficace de la population est petite. Une augmentation conjointe du progrès génétique et de la conservation de la diversité génétique est possible (i) en appliquant des méthodes de contribution optimale (OC) ; (iii) en réalisant des plans d'accouplement en ferme pertinents ; et (iii) en veillant à ce que ces outils soient suffisamment souples pour être utilisés efficacement sur le terrain.

Cette étude souligne le fait que les conséquences de la sélection génomique sur la diversité génétique dépendent de la manière d'utiliser les évaluations génomiques, et que la conception et les modifications des schémas de sélection, tout comme la manière de diffuser les taureaux sélectionnés, jouent un rôle majeur dans la gestion de la diversité génétique. La sélection génomique est donc un outil permettant l'augmentation du progrès génétique sans nécessairement mettre en danger la diversité génétique des races bovines laitières françaises.

Les entreprises de sélection sont responsables du maintien de la diversité génétique mais aussi du progrès génétique dans les races qu'elles sélectionnent. Ces résultats leur ont permis de savoir quel a été l'impact de leurs choix au moment de la mise en place de la sélection génomique, et si les modifications apportées aux schémas de sélection leur ont permis d'atteindre leurs objectifs en termes de progrès et de diversité génétiques.