Opportunities for genomic selection of cheese-making traits in Montbéliarde cows

M.P. Sanchez, T. Tribout, S. Fritz, V. Wolf, C. Laithier, M. Brochard, D. Boichard

J. Dairy Sci. 105
https://doi.org/10.3168/jds.2021-21558

Opportunities for genomic selection of cheese-making traits in Montbéliarde cows

M. P. Sanchez,[1]* T. Tribout,[1] S. Fritz,[2] V. Wolf,[3] C. Laithier,[4] M. Brochard,[4,5] and D. Boichard[1]
[1]Université Paris Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France
[2]Allice, 75012 Paris, France
[3]ECEL Doubs, Territoire de Belfort, 25640 Roulans, France
[4]Institut de l'Elevage, 75012 Paris, France
[5]Umotest, 01250 Ceyzériat, France

ABSTRACT

As part of the From'MIR project, traits related to the composition and cheese-making properties (CMP) of milk were predicted from 6.6 million mid-infrared spectra taken from 410,622 Montbéliarde cows (19,862 with genotypes). Genome-wide association studies of imputed whole-genome sequences highlighted candidate SNPs that were then added to the EuroG10K BeadChip, which is routinely used in genomic selection. In the present study, we (1) assessed the reliability of single-step genomic BLUP breeding values (ssEBVs) for cheese yields, coagulation traits, and casein and calcium content generated from test-day records of the first 3 lactations, (2) estimated realized genetic trends for these traits over the last decade, and (3) simulated different cheese-making breeding objectives and estimated the responses for CMP as well as for other traits currently selected in the Montbéliarde breed. To estimate the reliability of ssEBVs, the available data were split into 2 independent training and validation sets that respectively contained cows with the oldest and the most recent lactation data. The training set included 155,961 cows (12,850 with genotypes) and was used to predict ssEBVs of 2,125 genotyped cows in the validation set. We first tested 4 models that included either lactation (LACT) or test-day (TD) records from the first (1) or the first 3 (3) lactations, giving equal weight to all 50K SNP effects. Mean reliabilities were 61%, 62%, 63%, and 64% for the LACT1, LACT3, TD1, and TD3 models, respectively. Using the most accurate model (TD3), we then compared the reliabilities of 3 scenarios with: SNPs from the Illumina BovineSNP50 BeadChip only, equally weighted (50K); 50K SNPs plus additional candidate SNPs, equally weighted (50K+); and 50K and candidate SNPs with additional weight given to 7 to 14 candidate SNPs, depending on the trait (CAND). The 50K+ and CAND scenarios led to similar mean reliabilities (67%) and both outperformed the 50K scenario (64%), whereas the CAND scenario generated the less biased ssEBVs. To assess genetic trends, SNP effects were estimated with a single-step GBLUP based on the TD3 model and the 50K scenario applied to the whole population (2.6 million performance records from 190,261 cows and 423,348 animals in the pedigree, of which 21,874 were genotyped) and then applied to 50K genotypes of 21,171 males and 311,761 females. We detected a positive genetic trend for all CMP during the last decade, probably due to selection for an increase in milk protein and fat content in Montbéliarde cows. Finally, we compared the selection responses to 3 different breeding objectives: the current Montbéliarde total merit index (TMI) and 2 alternative scenarios that gave a weight of 70% to TMI and the remaining 30% to either milk casein content (TMI-COMP) or a combination of 3 CMP (TMI-Cheese). The TMI-Cheese scenario yielded the best responses for all the CMP analyzed, whereas values in the TMI-COMP scenario were intermediate, with a slight effect on other traits currently included in TMI. Based on these results, a program of genomic evaluation for CMP predicted from mid-infrared spectra was designed and implemented for the Montbéliarde breed.

Key words: genomic selection, cheese-making traits, Montbéliarde

INTRODUCTION

More than 36% of total cow milk is processed into cheese products worldwide (International Dairy Federation, 2016) and this proportion has increased by 23% during the last decade. The milk processing industry thus stands to gain a great deal economically from the improvement of milk cheese-making properties (**CMP**). To assess CMP, various laboratory methods have been developed (reviewed in Bittante et al., 2012), but they

Received November 10, 2021.
Accepted February 22, 2022.
*Corresponding author: marie-pierre.sanchez@inrae.fr

are costly and time consuming, and thus difficult to implement on a large scale. Mid-infrared (**MIR**) spectrometry has been proposed as an alternative method for the prediction of various milk characteristics, including cheese yield or coagulation properties (De Marchi et al., 2014). Mid-infrared technology is cheap and routinely used. With this approach, CMP can be indirectly measured on a large scale and could therefore be incorporated into the breeding objective for dairy cattle.

The From'MIR project, initiated in 2015, aims to analyze CMP and milk composition traits predicted from MIR spectra in Montbéliarde cows from the Franche-Comté region; the milk from these cows is mainly used to produce a variety of high-quality cheeses. As part of the work of From'MIR, MIR prediction equations were developed (El Jabri et al., 2019) and a set of 8 CMP traits (3 laboratory cheese yields and 5 coagulation traits) was predicted with relatively high accuracy from 6.6 million MIR spectra taken from 410,622 Montbéliarde cows (19,862 with genotypes). In prior studies, we estimated genetic parameters and conducted GWAS on whole-genome sequences for CMP and milk composition traits (proteins, fatty acids, and minerals). These studies revealed medium-to-high heritabilities for CMP traits and high genetic correlations among CMP traits and between CMP and some milk composition traits (Sanchez et al., 2018a), as well as several genomic regions (QTL) with large effects on these traits (Sanchez et al., 2019). Candidate causative variants were selected from the sequence-based GWAS and incorporated into the EuroG10K BeadChip, which was developed by the EuroGenomics consortium (https://www.eurogenomics.com/actualites/the-eurog-md:-a-unique-genotyping-microarray-for-cattle-.html) and is routinely used in genomic selection.

The objectives of the present study were to (1) estimate the reliability of single-step GBLUP breeding values for cheese yields, coagulation traits, and casein and calcium content in milk using different data sets and models, (2) estimate realized genetic trends for these traits in recent years, and (3) simulate different breeding objectives to estimate responses for CMP traits as well as for other traits currently selected in the Montbéliarde breed.

MATERIALS AND METHODS

Reference Population: Animals, Genotypes, and Phenotypes

The original data set is described in detail in prior studies (Sanchez et al., 2018a, 2019). It consisted of 4,877,908 MIR spectra of milk samples from 311,613 Montbéliarde cows originating from 3,229 herds of the Franche-Comté region. To ensure that the data set was homogeneous in the present study, we retained cows whose age at first calving ranged between 22 and 44 mo and who had a complete first lactation and at least 3 test-day records. After filtering, the data set included 2,869,353 test-day records from 191,532 cows.

For this study, we did not perform any experiment on animals; thus, no ethical approval was required.

Of these cows, 19,564 were genotyped for the purpose of genomic selection using either the BovineSNP50 (50K, 6,498 cows) or the EuroG10K BeadChip (Illumina Inc., 13,066 cows). The latter chip contains generic SNPs as well as a research add-on for causal or predictive SNPs for traits of interest in cattle, including candidate causative SNPs identified for milk composition (Boichard et al., 2014; Sanchez et al., 2017) whose effects were later confirmed (Sanchez et al., 2018b). Missing genotypes were imputed using FImpute software (Sargolzaei et al., 2014) for the 47,794 autosomal SNPs [41,492 (50K) and 6,852 custom SNPs (50K+)] that passed all quality control filters (individual call rate >95%; SNP call rate >90%; minor allele frequency >1%; genotype frequencies in Hardy-Weinberg equilibrium with $P > 10^{-4}$).

The following traits were investigated in this study:

- 3 laboratory cheese yields: fresh curd yield ($\mathbf{CY_{FRESH}}$) = 100 × (g of curd/g of milk), curd yield in dry matter ($\mathbf{CY_{DM}}$) = 100 × (1 − g of DM whey/g of DM milk), and curd yield in protein and fat ($\mathbf{CY_{FAT\text{-}PROT}}$) = (PC + FC) × (g of milk/g of curd), where PC = protein content and FC = fat content;
- 5 coagulation traits for pressed cooked cheese (**PCC**) and soft cheese (**SC**): curd organization index ($\mathbf{K10/RCT_{PCC}}$ and $\mathbf{K10/RCT_{SC}}$), where RCT is rennet coagulation time to 0.5 firm index (**FI**) and K10 is time to obtain 10 FI, curd firmness at RCT ($\mathbf{a_{PCC}}$ and $\mathbf{a_{SC}}$), and at 2 times RCT ($\mathbf{a2_{SC}}$); and
- total casein ($\mathbf{\Sigma CN}$) and Ca contents in milk.

The MIR prediction accuracy, as estimated by the value of R^2 in a validation population, varied between 0.54 ($CY_{FAT\text{-}PROT}$) and 0.98 (ΣCN) depending on the trait (Table 1). All CMP and milk composition traits were predicted from MIR spectra using equations developed in 3 different projects: From'MIR for CMP traits (El Jabri et al., 2019), PhénoFinLait for caseins (Ferrand et al., 2012), and Optimir for Ca (Soyeurt et al., 2009; Gengler et al., 2016).

Table 1. Features of cheese-making properties and milk composition traits

Trait[1]	Description and unit	Mean	SD	R^2_{val}[2]	h^2 TD model[3]	h^2 LACT model[3]	σ_G[3]
CY_{FRESH}	100 × (g of curd/g of milk), %	37.6	7.5	0.82	0.38	0.70	3.2
CY_{DM}	100 × (g DM of curd/g DM of milk), %	66.9	5.0	0.89	0.39	0.72	2.2
$CY_{FAT\text{-}PROT}$	(g of milk fat + g of milk protein)/kg of curd, g/kg	187.6	21.7	0.54	0.37	0.70	8.2
a_{PCC}	Curd firmness at rennet coagulation time (RCT), firm index (FI)	19.1	2.6	0.76	0.42	0.72	1.2
$K10/RCT_{PCC}$	Time to obtain 10 FI from RCT (K10)/RCT, min/min	0.37	0.10	0.68	0.47	0.71	0.044
a_{SC}	Curd firmness at RCT, FI	19.4	2.7	0.76	0.45	0.73	1.3
$a2_{SC}$	Curd firmness at 2 times RCT, FI	23.2	2.1	0.69	0.48	0.72	1.0
$K10/RCT_{SC}$	Time to obtain 10 FI from RCT (K10)/RCT, min/min	0.36	0.11	0.72	0.47	0.72	0.049
ΣCN	Total caseins, g/100 g of protein	80.9	1.3	0.98	0.46	0.73	0.12
Ca	Calcium, mg/kg of milk	1,161	95	0.82	0.50	0.75	52.6

[1]For pressed cooked cheese (PCC) and soft cheese (SC).
[2]Accuracy of mid-infrared prediction equations (R^2 val).
[3]Heritability (h^2) and genetic standard deviation (σ_G) estimates for the test-day (TD; Sanchez et al., 2018a) and lactation (LACT) models.

Training and Validation Data Sets

To estimate the accuracy of genomic predictions, the data set was split into 2, so that training and validation sets were as independent as possible (Figure 1). Primiparous cows with genotypes were included in the training population (Training) if they had calved before May 2016 and if they had at least 3 TD records (13,272 cows). The validation population (**VAL**) was composed of genotyped primiparous cows who had calved after August 2016 and had at least 5 TD records (4,032 cows). All performance data of the Training cows that had been recorded after August 2016 were removed (Figure 1a). Out of the 642 bulls with genotyped daughters in the whole data set, 180 had daughters in both populations. Including these families in the analysis would result in increased accuracy but also increased bias. Therefore, we considered different bull families in each data set: we kept cows (genotyped or not) in the training set if at least 40% of the daughters of a bull were Training cows, and we removed all half-sisters of these cows in the VAL set. For all other bulls (with less than 40% of daughters in Training), we kept the genotyped daughters in the VAL set and removed all half-sisters (genotyped or not) of these cows from the Training set. We selected 155,961 cows (12,850 with genotypes and 143,111 without genotypes) from 3,278 bulls (including 496 bulls with genotyped daughters) for use in Training and 2,125 genotyped cows from 146 other bulls for use in the VAL set (Figure 1b).

Single-Step Genomic Evaluation

Single-step estimated breeding values (**ssEBVs**) were calculated using the hybrid single-step model proposed by Fernando et al. (2016) and implemented in HSSGBLUP software (Tribout et al., 2020). Briefly, this method combines information on phenotypes, pedigrees, and genotypes to predict genomic breeding values for all the animals of the pedigree (genotyped or not) and directly provides marker effect estimates. Coherence between pedigree and genotypes was obtained by fitting a **J** vector (Hsu et al., 2017) for each unknown parent group.

Models and Scenarios

The ssEBVs were estimated using 2 categories of models that respectively considered the mean performance over the course of a lactation (**LACT**) or the individual test-day records (**TD**) of each cow. Both LACT and TD models were applied to the first lactation (**L1**) as well as to the first 3 lactations (**L3**). In total, 4 different models were tested: **LACT1** (1

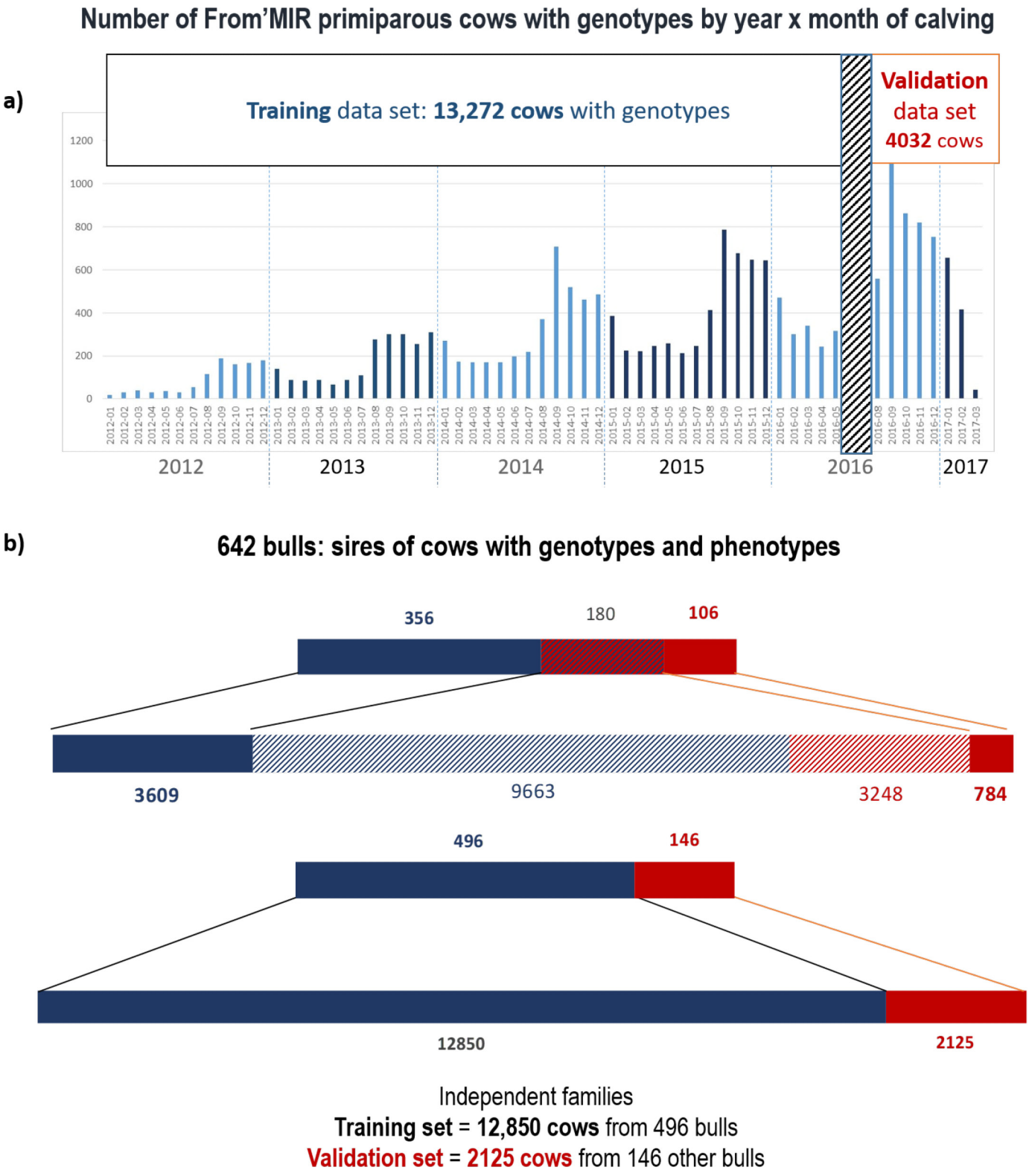

Figure 1. Selection of cows for training and validation sets.

observation per cow in L1), **LACT3** (up to 3 observations per cow, 1.8 on average in L1–L3), **TD1** (with an average of 7.1 observations per cow in L1), and **TD3** (with an average of 13.3 observations per cow in L1–L3). The effects and number of observations included in each model are described in Table 2. The contemporary group was the herd × year combination for LACT and the herd × test-day for TD. The other fixed effects were the year × month of calving and the age of the cow at first calving. In the TD models, the lactation curve was fitted by the effect of DIM (one class every 10 d, within parity). In the LACT models, the effect of the number of TD records was included as a proxy of average DIM. Variances of the random effects of each model were estimated using Wombat software (Meyer, 2007) with a pedigree relationship matrix. More details can be found in Sanchez et al. (2018a).

In addition to the polygenic models without genotyping information (**LACT1-BLUP** and **TD1-BLUP**), the LACT1, LACT3, TD1, and TD3 models were first tested using the 50K scenario (**LACT1–50K**, **LACT3–50K**, **TD1–50K**, and **TD3–50K,** respectively), and included all polymorphic SNPs of the BovineSNP50 chip (i.e., 41,942 SNPs). From these, we selected the model that generated the most accurate ssEBVs (TD3) and tested 2 additional scenarios in which we added SNPs from the custom part of the EuroG10K chip (i.e., considering a total of 47,794 SNPs). In the **TD3–50K+** scenario, we used constant effect variances for all 47,794 SNPs, whereas in the **TD3-CAND** scenario, we established specific variances for 7 to 14 candidate SNPs, depending on the trait (Table 3). These SNPs had been identified in previous studies as the best candidates for QTL with large effects on milk composition and CMP traits (Sanchez et al., 2017, 2018b, 2019). In the TD3-CAND scenario, we calculated the proportion of genetic variance for each SNP and each trait as follows: $\%\sigma_a^2 = 100\left(\frac{2p(1-p)\alpha^2}{\sigma_a^2}\right)$, with σ_a^2 being the additive genetic variance, p and $(1-p)$ being the allelic frequencies, and α being the allelic substitution effect estimated in GWAS. The cumulative effects of the candidate SNPs explained from 17.5% to 58.4% of the total additive genetic variance of the traits, with the remaining part (41.6% to 82.5%) being equally distributed among all other SNPs (47,780 to 47,787 depending on the trait).

Table 2. Features of models and training populations tested

Name	LACT1	LACT3	TD1	TD3
Phenotype	Mean phenotype over lactation (LACT) 1	Mean phenotype over lactations 1 to 3	Test-day (TD) records in lactation 1	Test-day records in lactations 1 to 3
Model[1]	$\mathbf{y = Xb + Za + e}$	$\mathbf{y = Xb + Za + Zp + e}$	$\mathbf{y = Xb + Za + Zp + e}$	$\mathbf{y = Xb + Za + Zp + e}$
Fixed effects **b**	Herd × year	Herd × year	Herd × TD	Herd × TD
	Month × year calving	Month × year calving	Month × year calving	Month × year calving
	Age at first calving	Age at calving	Age at first calving	Age at calving
	No. of TD records	No. of TD records × parity	DIM	DIM × parity
Training population				
No. of cows with phenotypes[2]	155,479	155,919	155,961	155,961
No. of TD records	155,479	287,800	1,115,000	2,070,000
No. of animals in pedigree	413,787	422,857	414,919	423,348
No. of animals in pedigree with genotypes	21,764	21,874	21,747	21,874

[1]**a**, **p**, and **e** are animal, permanent environment, and residual random effects with variances σ_a^2, σ_p^2, and σ_e^2 estimated with Wombat software (Meyer, 2007), respectively; **X** and **Z** are incidence matrices for fixed and random effects, respectively.
[2]The number of cows with phenotypes differed slightly among models due to the filters applied to the corresponding data set.

Reliability and Bias

The reliability and bias of each model and scenario were then estimated using the validation population.

Table 3. Candidate SNPs of the 18 QTL detected for milk composition and cheese-making traits[1]

				% of additive genetic variance associated with each trait × SNP combination											
QTL	BTA	SNP ID	MAF	CY_{FRESH}[2]	CY_{DM}	$CY_{FAT\text{-}PROT}$	a_{PCC}	K10/ RCT_{PCC}	a_{SC}	$a2_{SC}$	K10/ RCT_{SC}	ΣCN	Ca	No. of traits affected	Functional annotation
1	1	rs132965371	0.45			0.75	1.1				0.48			3	Intron *SLC37A1*
2	2	rs110123105	0.79	0.55					0.94	0.58			0.51	4	Intergenic
3	2	rs133944376	0.37				0.39					1.3		2	Missense *ALPL*
4	5	rs211210569	0.12	1.3	1.8	1.1								3	Intron *MGST1*
5	5	rs525880746	0.04	2.8	2.7	1.2		3.9	0.41	3.5	2.6	1.9	5.4	9	Upstream *GRAMD4*
6	6	rs136800773	0.73	2.5		3.0	4.4		2.3		2.0	4.3	0.73	7	Downstream *SLC34A2*
7	6	rs43703016	0.40	7.7	7.2	8.7	2.3	42.7	4.3	35.8	42.1	1.5	1.3	10	Missense *CSN3*
8	7	rs29024414	0.08			1.5	3.0		2.9		1.0	2.2	3.1	6	Upstream *ENSBTAG00000004032*
9	11	rs110066229	0.46	8.3	10.4	23.0	10.4	2.1	3.0	2.8	3.3	16.2		9	Missense *PAEP*
10	14	rs109035586	0.33	12.7	16.4	2.8		3.4		3.2	2.3	2.5	4.2	8	Upstream *GPT*
11	16	rs109033026	0.08	0.45	0.45		0.26	0.58	0.58	0.47	0.57	0.52		8	Intergenic
12	17	rs448501071	0.06					1.9		1.2	1.0	0.58		4	Upstream *BRI3BP*
13	19	rs136067046	0.32	1.4	1.7									2	Upstream *FASN*
14	20	rs385744846	0.07	0.41		1.3	0.56		2.4		0.87	2.3	2.2	7	Intron *ANKH*
15	22	rs41642478	0.26	1.9	2.2	0.73	0.36	2.0	0.90	1.3	1.6	0.78		9	Intron *FAM19A4*
16	24	rs109478290	0.21	0.46	0.48			0.55	0.62	0.41	0.48			6	Intron *LMAN1*
17	26	rs41255693	0.46	0.50	0.59									2	Missense *SCD*
18	27	rs208675276	0.41	1.8	2.6	1.7								3	5′UTR *GPAT4*
		Total (% σ_a^2)		42.9	46.5	45.9	22.8	57.1	18.3	49.3	58.4	34.2	17.5		
		No. of QTL		14	11	11	9	8	10	9	12	11	7		

[1]MAF: minor allele frequency; UTR: untranslated region.

[2]For each SNP and trait, the % of genetic variance was calculated as follows: $\%\sigma_a^2 = 100\left(\frac{2p(1-p)\alpha^2}{\sigma_a^2}\right)$, with σ_a^2 being the additive genetic variance, p and $(1-p)$, allelic frequencies, and α, the allelic substitution effect estimated from GWAS; see description of traits in Table 1.

The ssEBVs were estimated by applying the Training genomic prediction equations to the genotypes of the 2,125 validation cows. Using GENEKIT software (Ducrocq, 1998), we adjusted the phenotypes of primiparous VAL cows for nongenetic effects estimated with the TD1 model (described in Table 2) applied to the complete Training + VAL data set, but ignoring marker information. To compare the accuracies of the different models, we calculated, for each trait, the correlation coefficients (**r**) between means of individual performances (corrected for nongenetic effects using the TD1 model) and ssEBVs estimated with the LACT1–50K ($\mathbf{r}_{\mathbf{LACT1–50K}}$), LACT3–50K ($\mathbf{r}_{\mathbf{LACT3–50K}}$), TD1–50K ($\mathbf{r}_{\mathbf{TD1–50K}}$), or TD3–50K ($\mathbf{r}_{\mathbf{TD3–50K}}$) model. Similarly, for the TD3 model, the 3 scenarios were compared by calculating correlations between mean individual performances (corrected for nongenetic effects using the TD1 BLUP model) and ssEBVs estimated with the 50K ($\mathbf{r}_{\mathbf{TD3–50K}}$), 50K+ ($\mathbf{r}_{\mathbf{TD3–50K+}}$), or CAND ($\mathbf{r}_{\mathbf{TD3\text{-}CAND}}$) scenario. Reliabilities were calculated by dividing the squared correlations ($r^2_{\text{LACT1–50K}}$, $r^2_{\text{LACT3–50K}}$, $r^2_{\text{TD1–50K}}$, $r^2_{\text{TD3–50K}}$, $r^2_{\text{TD3–50K+}}$, or $r^2_{\text{TD3-CAND}}$) by the heritability of the corresponding trait estimated using Wombat software (Meyer, 2007) with the LACT1 model; these are reported in Table 1. Bias was estimated as the deviation from 1 of the slope of the regression of corrected performances on ssEBVs.

Estimation of Genetic Trends

To estimate genetic trends in CMP and milk composition traits in Montbéliarde cattle, we applied the TD3–50K model to the whole From'MIR data set. This was equivalent to the data set used when the TD3–50K model was applied to the training population (423,348 animals in the pedigree, of which 21,874 with genotypes for 41,942 SNPs), except that all phenotypes were kept (i.e., 2.6 million TD records from 190,261 cows).

We then applied the genomic prediction equations obtained from the whole data set to all genotyped Montbéliarde animals belonging to the Umotest breeding company, including 17,859 males born between 2005 and 2018 and 256,861 females born between 2008 and 2018. Genetic trends were assessed separately for males and females by averaging ssEBVs per year of birth.

Breeding Simulation with Different Breeding Goals

Finally, we conducted a breeding simulation to investigate the effect of incorporating cheese-making traits into the breeding goal of the population. We compared 3 different breeding goals: the total merit index (**TMI**), which is the current breeding goal in French Montbéliarde cattle, and 2 alternative indexes (**TMI-COMP** and **TMI-Cheese**), defined as follows:

1) TMI = 0.45 Milk + 0.145 Udder-Health + 0.18 REPROD + 0.05 LONG + 0.05 Speed + 0.125 MORPH, with weights expressed in genetic standard deviations. This is a combination of
   - the production index Milk = 14 PY + 2 FY + 4.2 PC + 1.3 FC, which combines breeding values for milk protein yield (**PY**), fat yield (**FY**), protein content (**PC**), and fat content (**FC**);
   - the udder health index Udder-Health = 0.6 Cell + 0.4 MAST, which combines breeding values for somatic cell score (Cell) and clinical mastitis (MAST);
   - the fertility index REPROD = 0.5 FERC + 0.25 FERH + 0.25 CALV-AI, which combines breeding values for conception rate of cows (FERC) and heifers (FERH) and the interval between calving and first artificial insemination (CALV-AI);
   - functional longevity (LONG);
   - speed of milking (Speed);
   - the overall morphology index MORPH = 0.4 Udder + 0.3 Body + 0.15 Legs + 0.10 Rump + 0.05 Muscularity, with Udder, Body, Legs, Rump, and Muscularity being the aggregate index for udder conformation, body capacity, feet and legs, rump morphology, and muscularity, respectively;
2) TMI-COMP = 0.7 TMI + 0.3 ΣCN;
3) TMI-Cheese = 0.7 TMI + 0.1 (CY_{DM} + a_{PCC} − $K10/RCT_{PCC}$).

We simulated selection by truncation on bulls based on the breeding strategy applied by the Umotest breeding company. For the traits investigated in the present study, we considered the ssEBVs of bulls estimated with the TD3–50K model with the whole data set, whereas for all other traits routinely included in genomic evaluations, we considered official genomic indexes. To ensure a sufficient number of bulls per year of birth, only the 14,389 bulls born between 2009 and 2017 were considered (1,022 to 2,210 bulls per year of birth). Within each year of birth, we sorted bulls by decreasing TMI, TMI-COMP, or TMI-Cheese value, calculated using the formulas described above. In a given year, we selected the 80 bulls with the highest TMI, TMI-COMP, or TMI-Cheese indexes among all the candidates (1,600 candidates on average). We then calculated the responses to selection for the ssEBVs estimated in the present study (10 milk CMP and

composition traits) as well as for the 8 official indexes described earlier: TMI, Udder-Health, REPROD, PY, FY, PC, FC, and milk yield (**MY**).

For each breeding goal scenario, we estimated the annual responses to selection by calculating the selection differential of each ssEBV or index (i.e., the difference between the mean of the 80 best bulls and the mean of all candidates). The yearly responses, expressed in genetic standard deviations, were then averaged over the whole period (2009–2017).

RESULTS

Reliability of Single-Step Genomic Evaluation

Estimates of heritability coefficients and genetic standard deviations are in Table 1. The heritability coefficients for the lactation traits were used to estimate the reliability of the evaluations.

To estimate the reliability of single-step genomic evaluation for milk cheese-making traits, we removed these phenotypes from the validation population and calculated breeding values (ssEBVs). We then assessed the reliabilities of each model × trait combination in the validation population.

As described in the Materials and Methods, we first compared 6 models, without genotyping data (LACT1-BLUP and TD1-BLUP) or applied to the 50k genotypes (LACT1–50K, LACT3–50K, TD1–50K, and TD3–50K). The correlations, calculated between ssEBVs and adjusted performance, are presented in Supplemental Table S1 (https://doi.org/10.6084/m9.figshare.19477925). All genomic models largely outperformed the polygenic model. Mean correlation values of all traits for LACT1 and TD1 models were, respectively, 0.37 and 0.38 for polygenic models and 0.66 and 0.67 for genomic models. In the following part, only genomic models will be compared.

Regardless of the model tested and the trait analyzed, all reliabilities were higher than 59% (Figure 2). We found slightly higher reliabilities for milk coagulation and composition traits, but overall, reliabilities were similar for all traits. For a given category of model, based on performance data averaged over the course of lactation (LACT) or individual test-day records (TD), reliabilities were slightly higher for models that included data from the second and third lactations together with those from the first (+1 point on average). In all cases, the TD models (TD1–50K and TD3–50K) gave more reliable ssEBVs than the LACT models (LACT1–50K and LACT3–50K) by about 2 points on average. Across all traits, average reliabilities were 61%, 62%, 63%, and 64% for the LACT1–50K, LACT3–50K, TD1–50K, and TD3–50K models, respectively.

Because the TD3 model gave the most reliable results for all traits, we applied this model to the training population to test 3 scenarios that differed in the density or weighting (or both) of SNPs (TD3–50K, TD3–50K+, and TD3-CAND). We then compared the reliabilities and biases of ssEBVS calculated with the

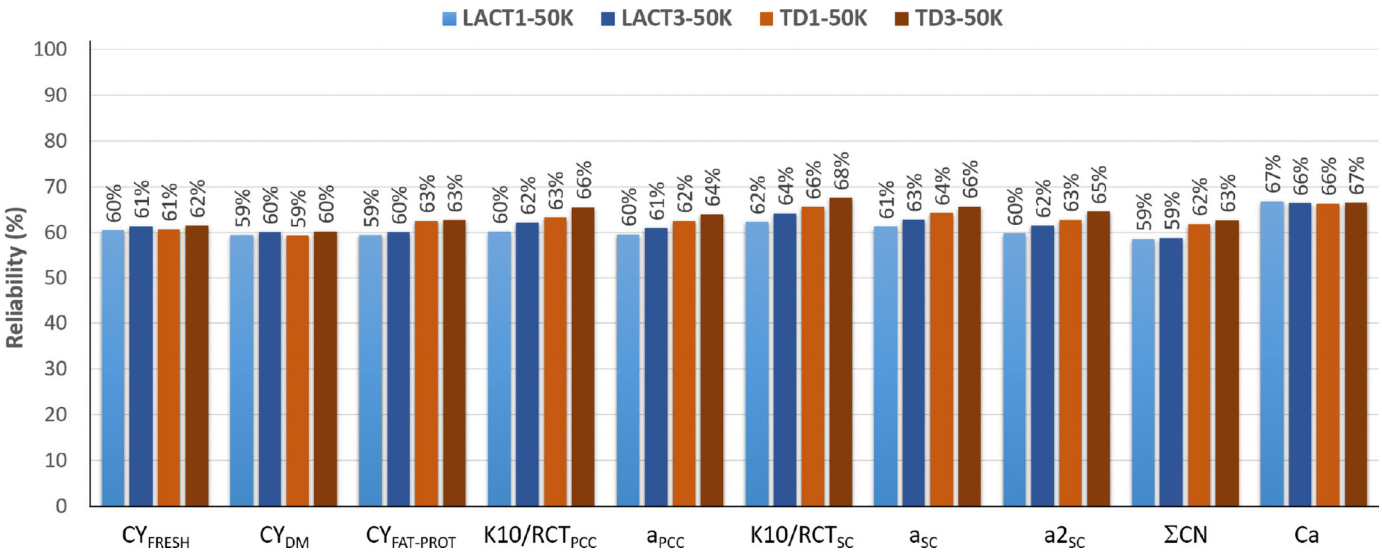

Figure 2. Reliabilities estimated in the validation population (n = 2,150) with the 4 tested models. LACT = lactation; TD = test-day. CY_{FRESH} = fresh curd yield; CY_{DM} = curd yield in dry matter; $CY_{FAT\text{-}PROT}$ = curd yield in protein and fat; $K10/RCT_{PCC}$ curd organization index for pressed cooked cheese, where RCT is rennet coagulation time to 0.5 firm index (FI) and K10 is time to obtain 10 FI; aPCC = curd firmness at RCT for pressed cooked cheese; $K10/RCT_{SC}$ = curd organization index for soft cheese; a_{SC} = curd firmness at RCT for soft cheese; $a2_{SC}$ = curd firmness at 2 times RCT; ΣCN = total casein; Ca = Ca content.

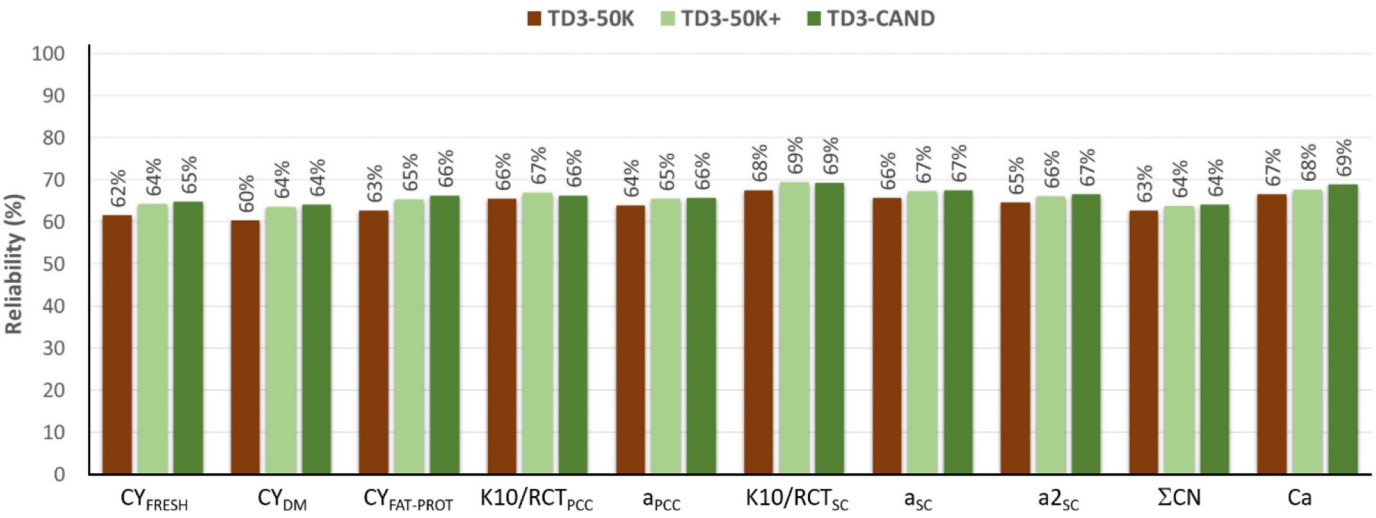

Figure 3. Reliabilities estimated in the validation population (n = 2,150) with the 3 tested scenarios. TD = test day; CAND = weighted candidate SNPs.

3 scenarios in the validation population. The inclusion of SNPs from the custom part of the EuroG10K chip (the TD3–50K+ scenario: 5,852 additional SNPs compared with the TD3–50K scenario) further improved the reliability of ssEBVs (Figure 3). Across all traits, the average gain was about 2 points: from +2.7 to +3.4 points for cheese yields and from +1.1 to +1.9 points for coagulation parameters and casein and calcium content. The TD3-CAND scenario, which assigned higher weights to candidate causal SNPs, provided only a marginal gain of 0.25 points on average, with differences depending on the trait (from +0.2 to +1.1 points for a_{PCC}, ΣCN, a_{SC}, $a2_{SC}$, CY_{DM}, CY_{FRESH}, $CY_{FAT\text{-}PROT}$, and Ca versus from −0.1 and −0.7 points for $K10/RCT_{SC}$ and $K10/RCT_{PCC}$, respectively). Instead, when we evaluated bias in the scenarios, based on the slope of the regression line of corrected performance on ssEBVs, we found different results (Figure 4). For all traits and all 3 scenarios, slopes were less than 1, indicating bias in the estimation of ssEBVs. The TD3-CAND scenario generated the least biased ssEBVs, whereas the other 2 scenarios gave equivalent results, with the TD3–50K

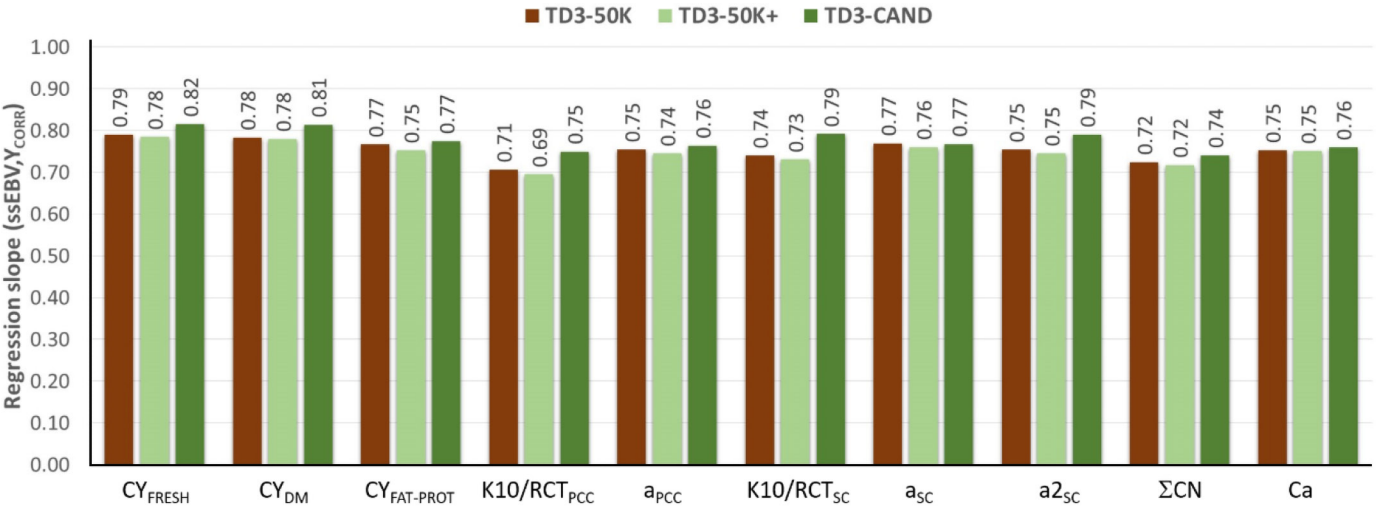

Figure 4. Biases estimated in the validation population (n = 2,150) with the 3 tested scenarios. TD = test day; CAND = weighted candidate SNPs.

scenario slightly outperforming the TD3–50K+ scenario.

Genetic Trends in Cheese-Making Traits

Genetic trends in CMP traits were estimated by calculating the mean genetic values (ssEBVs) of males and females per year of birth, and are shown in Figure 5. Over the period 2005–2018, we observed regular improvement in ssEBVs for all of the cheese-making criteria and milk composition traits analyzed in this study (i.e., an increase in ssEBVs for all traits except for 2 traits for which reduction is desirable: $CY_{FAT\text{-}PROT}$, which is inversely proportional to curd weight, and K10/RCT, which is inversely proportional to the curd organization speed). In the Comté PDO region, the average genetic level of Montbéliarde bulls and cows with respect to cheese yields and coagulation parameters is therefore higher today than in 2005. When expressed as genetic standard deviations (σ_g), the traits with the most marked genetic improvement over this period were coagulation parameters and total caseins in milk (between 0.5 and 0.6 σ_g). The genetic trend was instead more moderate for cheese yields (between 0.2 and 0.4 σ_g) and calcium content in milk (0.1 σ_g).

Effect of Different Breeding Strategies on Cheese-Making Traits and Other Traits

Annual responses to selection, averaged over the 2009–2017 period, were calculated with the current breeding objective (TMI) and 2 alternative scenarios that gave a weight of 70% to TMI and 30% to either casein content in milk (TMI-COMP) or to a combination of 3 cheese-making traits (TMI-Cheese). As expected, the addition of new criteria to the breeding goal had an effect on the breeding values of all investigated traits (Figure 6). For almost all traits, the TMI-COMP scenario yielded values that were intermediate between those of the TMI and TMI-Cheese scenarios. When we examined the traits incorporated in the current breeding goal (i.e., noncheese-making traits), we found that both alternative scenarios resulted in lower responses than TMI for PY, FY, MY, and Udder-Health, a similar response for REPROD, and higher responses for PC and FC. However, the reductions in gain per year were relatively limited (0.1 σ_g for PY, FY, and Udder-Health and 0.2 σ_g for Milk) considering the high weight given to casein content or cheese-making traits in the alternative scenarios. Instead, the additional emphasis on selection on cheese-making traits in bulls in TMI-COMP, and a fortiori in TMI-Cheese, led to much higher response levels for all cheese-making traits included in this study (i.e., all cheese yields and coagulation parameters measured on pressed cooked and SC). For CY_{DM}, for example, the annual gain was almost doubled with the TMI-COMP scenario (+0.68 σ_g) and almost tripled with the TMI-Cheese scenario (+0.95 σ_g) compared with the TMI scenario (+0.38 σ_g). Gains for other cheese-making traits were similar. Milk composition traits were also significantly improved, to a similar extent, in both alternative scenarios. Total casein content increased annually by 0.52, 0.86, and 0.85 σ_g, whereas calcium content increased by 0.38, 0.50, and 0.52 σ_g with the TMI, TMI-COMP, and TMI-Cheese scenarios, respectively.

DISCUSSION

The results presented in this study are very encouraging for the implementation of single-step genomic evaluation of CMP in the Montbéliarde breed. Breeding values with high reliability were obtained and favorable genetic trends in CMP traits were identified over the last 13 yr. However, we showed that amending the current breeding goal to include a combination of 3 CMPs could result in more rapid improvement to all CMP traits, with only a limited effect on other traits currently selected in the Montbéliarde breed.

By testing different models and scenarios, we showed that all genomic models strongly outperformed the polygenic model (with a reliability over 0.59 vs. less than 0.25) and we were able to identify which genomic model gave the best results. The model that generated the most reliable breeding values was the test-day model that considered all the observations recorded once a month during the first 3 lactations of the cows. Compared with the other models tested (lactation models with performance data averaged over the lactation period or the test-day model for the first lactation only), the best-performing model integrated the largest amount of information and was more accurate in modeling phenotypes. Furthermore, by incorporating SNPs from the custom part of the EuroG10K BeadChip, we included those that were particularly concentrated in the genomic regions affecting many important traits (QTL) in cattle, which probably explained the gain in reliability obtained with these scenarios (50K+ and CAND scenarios) compared with the scenario based only on 50K SNPs. Indeed, in addition to having the candidate variants identified by the PhenoFinlait and From′MIR projects, the custom part of the EuroG10K chip is also particularly enriched in SNPs from QTL regions associated with milk production traits (Boichard et al., 2018). Among these QTL regions are, for example, the regions of the casein or *PAEP* genes, which have been widely investigated in recent decades for their strong effects on milk composition and cheese-making traits

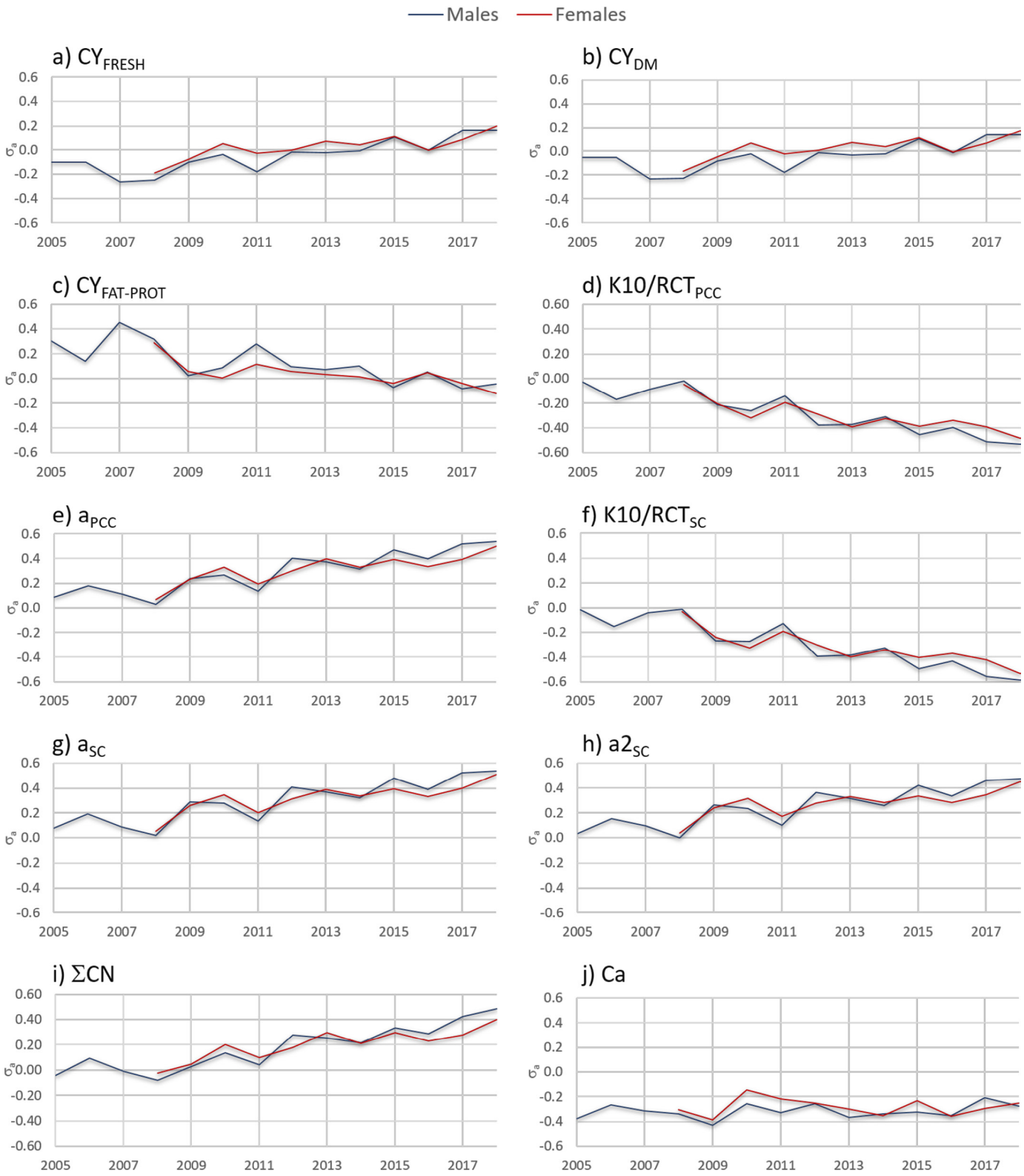

Figure 5. Genetic trends in genetic standard deviation (σ_a) estimated for bulls born between 2005 and 2018 and cows born between 2008 and 2018 for milk cheese-making and composition traits. CY_{FRESH} = fresh curd yield; CY_{DM} = curd yield in dry matter; $CY_{FAT\text{-}PROT}$ = curd yield in protein and fat; $K10/RCT_{PCC}$ curd organization index for pressed cooked cheese, where RCT is rennet coagulation time to 0.5 firm index (FI) and K10 is time to obtain 10 FI; aPCC = curd firmness at RCT for pressed cooked cheese; $K10/RCT_{SC}$ = curd organization index for soft cheese; a_{SC} = curd firmness at RCT for soft cheese; $a2_{SC}$ = curd firmness at 2 times RCT; ΣCN = total casein; Ca = Ca content.

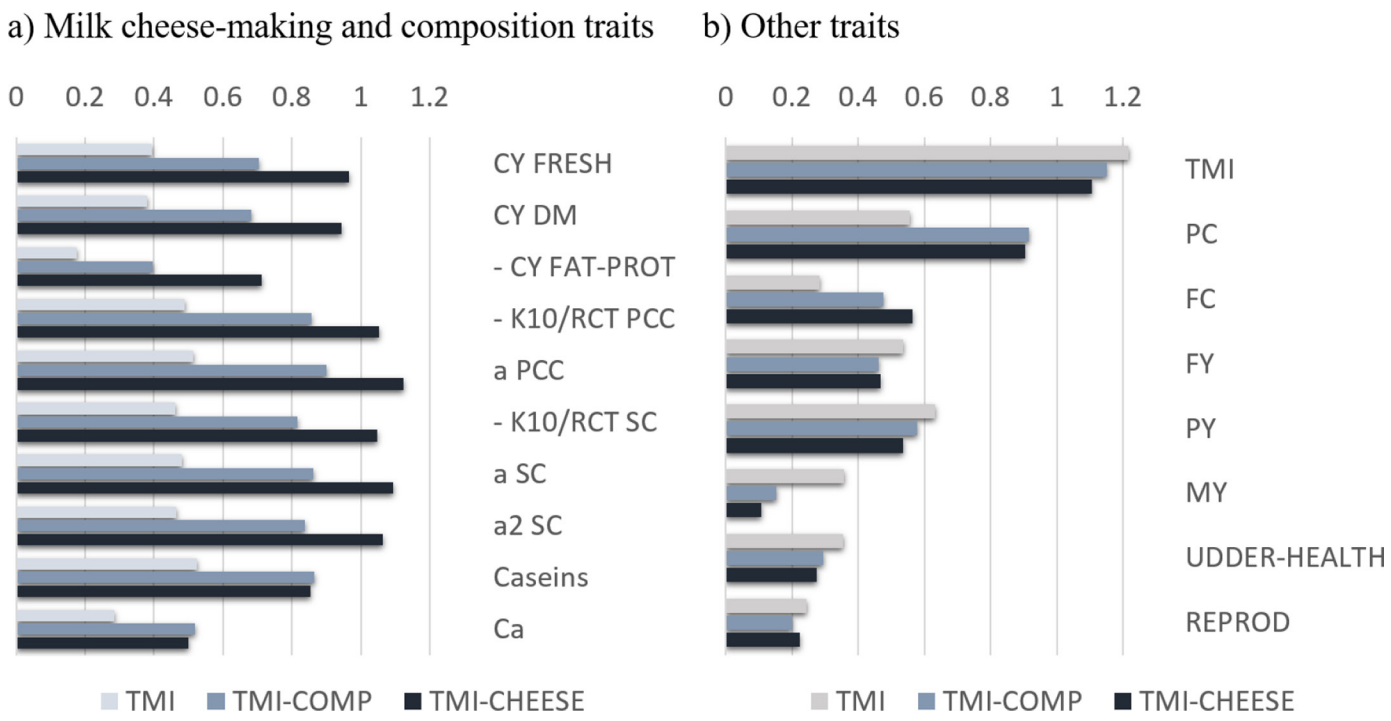

Figure 6. Responses to selection for actual total merit index without cheese-making properties (TMI), 70% TMI + 30% casein content (TMI-COMP), and 70% TMI + 30% cheese-making properties (TMI-Cheese) on milk cheese-making and composition traits and other traits (in genetic standard deviations). CY_{FRESH} = fresh curd yield; CY_{DM} = curd yield in dry matter; $CY_{FAT\text{-}PROT}$ = curd yield in protein and fat; $K10/RCT_{PCC}$ curd organization index for pressed cooked cheese, where RCT is rennet coagulation time to 0.5 firm index (FI) and K10 is time to obtain 10 FI; aPCC = curd firmness at RCT for pressed cooked cheese; $K10/RCT_{SC}$ = curd organization index for soft cheese; a_{SC} = curd firmness at RCT for soft cheese; $a2_{SC}$ = curd firmness at 2 times RCT. PC = protein content; FC = fat content; FY = fat yield; PY = protein yield; MY = milk yield; REPROD = reproduction.

in many dairy cattle breeds (Martin et al., 2002; Ganai et al., 2009). Scenarios with custom SNPs are mostly beneficial to cheese yields, which are highly genetically correlated with milk fat content (Sanchez et al., 2018a) and thus particularly affected by the *DGAT1* gene region (Sanchez et al., 2019). This region contains numerous SNPs that are featured in the research add-on of the EuroG10K chip (153 SNPs in the 50K+ scenario versus 25 SNPs in the 50K scenario). As a consequence, the average absolute values of SNP effects estimated in this region were higher in the 50K scenario than in the 50K+ or CAND scenarios (e.g., 0.052, 0.021, and 0.019 points for cheese yield in DM, respectively); this was due to the fact that, in the 50K+ and CAND scenarios, the effect of the chromosomal region was distributed over a larger number of SNPs, each with smaller individual effects. As expected, in the CAND scenario, the candidate variant with 16% of the genetic variance of the trait had the largest estimated effect (0.75 points), whereas the other 152 SNPs in the region had a very small estimated effect. Regardless of the weights attributed to the SNPs, the relative SNP enrichment of this region in the 50K+ and CAND scenarios, and the consequent increase in LD with causal mutations, probably helped to better capture the effects of the causal mutations and thus increase the reliabilities of breeding values. The inclusion of weighted candidate mutations (CAND scenario) resulted in a somewhat similar result in terms of reliability, but with a different distribution of effects: the strong effect of the candidate causal mutation led to reductions in the effects of nearby SNPs. However, the slight increase in reliability in the CAND scenario was also accompanied by a reduction in bias of genetic values. This was likely a reflection of the fact that the estimated effect of a causal variant is expected to be more stable over generations than the estimated effects of neutral markers, which depend on recombinations and distance to the reference population and are thus more prone to inflation.

The reliabilities of the breeding values we obtained (66% on average for all traits with the best model) were high in spite of a medium-size reference population (3,278 sires with a total of 155,961 daughters having MIR information, including 12,850 cows with genotypes in the training population). This result is explained by both the heritability of the traits and the number of repeated records per cow, resulting in highly informative phenotypes per cow, explain these reliabili-

ties. They are consistent with the theoretical accuracy expected given the heritability estimates of these traits and the size of the reference population included in the present study (Goddard and Hayes, 2009). With this level of reliability, comparable to that obtained for other currently selected traits (e.g., milk production traits), it becomes possible to implement a program of genomic evaluation of cheese-making traits predicted from MIR spectra. It is important to note here that our division of the reference population into 2 independent subpopulations, although necessary to give rigorous reliability estimates, had the effect of greatly reducing the size of the population used to estimate SNP effects (12,850 cows out of the 19,564 cows in the reference population). The inclusion of all cows with phenotypes and genotypes, including the new ones generated since the time of the study (around +40,000/yr), will further improve the prediction accuracy of genomic values.

Another point to consider is that CMP measurements are indirect MIR predictions and all our results (reliability but also genetic trend; next paragraph) assume that these predictions are robust over the whole data set (i.e., over herds and years, beyond fat and protein contents). Even if it is probably the case for the best-predicted traits, our reliability results may be optimistic and a periodic validation of the CMP prediction equations will be useful to avoid any potential drift between real CMP values and predictions.

The favorable genetic trends for cheese yields and coagulation parameters we observed over the last 13 yr are likely due to strong genetic correlations between milk protein and fat content and cheese-making traits (Sanchez et al., 2018a), the latter having probably been selected indirectly together with the former, which is included in the TMI in the Montbéliarde breed. We found a particularly strong response for total casein content, which represents about 80% of total proteins in milk. This result suggests that selection on protein content, which has a relatively important economic weight in the Montbéliarde TMI, has led to an improvement of CMP and, in particular, coagulation parameters, which present the strongest genetic correlations with protein content (Sanchez et al., 2018a). These results are consistent with the increase detected in the frequencies of some milk protein variants that are associated with better cheese-making abilities in French dairy cattle breeds (Sanchez et al., 2020).

To enhance this positive trend, we show here that a more significant genetic gain in cheese-making traits could be possible with only a limited effect on the other traits currently considered in the breeding objective for Montbéliarde cattle. This could be accomplished by amending the breeding objective to include 3 traits (cheese yield in DM and 2 coagulation traits for PCC), which were identified as the most relevant by a group of experts, composed of farmers, cheesemakers, and ripeners in the Franche-Comté region.

Despite the potential economic benefit of improving CMP, to the best of our knowledge, very few countries currently include CMP in their breeding objective. The United States calculates a so-called cheese index (Dollars Cheese Merit Index, CM$), but like the current TMI in Montbéliarde, it simply gives a greater weight to protein content (VanRaden, 2004). In Belgium, the ProFARMilk project (2011–2017) investigated the technological properties of milk, as predicted by MIR spectrometry, with respect to its ability to be processed into cheese, yogurt, and butter (Colinet et al., 2013). This project led to the implementation of a tool for monitoring milk processing abilities in the framework of milk recording in Wallonia, with the initial goal of discarding noncoagulating milk and eventually enabling the selection of these criteria (AWE, 2018). Italy, which processes more than 80% of its milk into cheese, has for many years used indicators of cheese aptitude (e.g., casein content or lactodynamic behavior of milk measured with a Formagraph) in the milk payment scale in the Parmigiano-Reggiano PDO region (Malacarne et al., 2004). In addition, some coagulation parameters predicted by MIR spectrometry have been routinely recorded since 2011 (De Marchi et al., 2012; Pretto et al., 2012). Finally, in the Veneto region, an index (Cheese Aptitude Index) that combines coagulation speed (RCT) and firmness (a30), with equivalent weights, has been published periodically since January 2012 for Holstein bulls (http://www.intermizoo.com/research/cheese-aptitude-index).

In France, traits associated with technological properties for cheese-making are not included in the current breeding objective of any dairy cattle breed. This study demonstrates, though, that their inclusion could have an obvious economic benefit for the dairy sector (e.g., by increasing cheese yields). However, it is worth noting that we have attributed a relatively high weight to cheese-making traits in the scenarios tested in the present study, to emphasize and illustrate the potential effect of selection. In practice, lower weight is given to CMP, ensuring positive gains for both MY and CMP. Selection on cheese-making traits will require further study to determine the optimal economic weights for these traits with respect to all the traits of the breeding goal. In addition, before selecting for cheese yields and milk coagulation traits, it will be necessary to assess the effects of such selection on the sensory quality of ripened cheeses. For this purpose, we have implemented a pilot genomic evaluation program in the Montbéliarde

breed that makes indexes available for these traits, which will make it possible to control genetic trends in CMP with respect to the sensory quality of cheeses.

CONCLUSIONS

This study investigated the potential for selecting milk technological traits, as predicted from MIR spectra, in the Montbéliarde dairy cattle breed. We showed it was possible to obtain reliable breeding values for cheese yields and coagulation parameters and identified the model that produced the most reliable and least biased results. The application of genomic prediction equations to all genotyped bulls and cows revealed favorable genetic trends in cheese-making traits over the last 13 yr. Finally, we found that incorporating cheese-making traits into the breeding goal could lead to more rapid improvement in these traits, with only a limited effect on other traits currently selected in this breed. Further investigation is needed to determine how best to integrate cheese-making traits into the breeding objective and to study the effect of selecting these traits on the quality of ripened cheeses. Regardless, all of these results are very encouraging for the implementation of single-step genomic evaluation of cheese-making parameters in the Montbéliarde breed, which could subsequently be extended to other dairy cattle breeds.

ACKNOWLEDGMENTS

This study was funded by the French Ministry of Agriculture, Agro-food, and Forests; the French Dairy Interbranch Organization (CNIEL); the Regional Union of Protected Designation cheeses of Franche-Comté (URFAC); and the Regional Council of Bourgogne-Franche-Comté, as part of the project FROM'MIR. The authors have not stated any conflicts of interest.

REFERENCES

AWE. 2018. La transformation fermière du lait. Accessed Nov. 10, 2021. http://www.awenet.be/awe/UserFiles/file/asbl/lait/Folder%20transformation%20du%20lait%202018.pdf.

Bittante, G., M. Penasa, and A. Cecchinato. 2012. Invited review: Genetics and modeling of milk coagulation properties. J. Dairy Sci. 95:6843–6870. https://doi.org/10.3168/jds.2012-5507.

Boichard, D., M. Boussaha, A. Capitan, D. Rocha, C. Hozé, M. P. Sanchez, T. Tribout, R. Letaief, P. Croiseau, C. Grohs, W. Li, C. Harland, C. Charlier, M. S. Lund, G. Sahana, M. Georges, S. Barbier, W. Coppieters, S. Fritz, and B. Guldbrandtsen. 2018. Experience from large scale use of the EuroGenomics custom SNP chip in cattle. In 11th WCGALP. Auckland, New Zealand. Molecular Genetics 4:675.

Boichard, D., A. Govignon-Gion, H. Larroque, C. Maroteau, I. Palhiere, G. Tosser-Klop, R. Rupp, M. P. Sanchez, M. Brochard, Y. Amigues, M. Y. Boscher, and H. Leveziel. 2014. Genetic determinism of milk composition in fatty acids and proteins in ruminants, and selection potential. INRA Prod. Anim. 27:283–298. https://doi.org/10.20870/productions-animales.2014.27.4.3074.

Colinet, F. G., T. Troch, O. Abbas, V. Baeten, F. Dehareng, E. Froidmont, H. Soyeurt, P. Dardenne, M. Sindic, and N. Gengler. 2013. Potentiel d'utilisation de la spectrométrie moyen infrarouge pour prédire le rendement fromager du lait et étudier sa variabilité génétique. In Renc. Rech. Rum. Vol. 20, Paris, France.

De Marchi, M., M. Penasa, F. Tiezzi, V. Toffanin, and M. Cassandro. 2012. Prediction of milk coagulation properties by Fourier Transform Mid-Infrared Spectroscopy (FTMIR) for genetic purposes, herd management and dairy profitability. Pages 47–53 in International Strategies and New Developments in Milk Analysis. VI ICAR Reference Laboratory Network Meeting. Vol. ICAR Technical Series No. 16, Cork, Ireland.

De Marchi, M., V. Toffanin, M. Cassandro, and M. Penasa. 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. J. Dairy Sci. 97:1171–1186. https://doi.org/10.3168/jds.2013-6799.

Ducrocq, V. 1998. Genekit, BLUP software. June 2011 version. INRA GABI.

El Jabri, M., M. P. Sanchez, P. Trossat, C. Laithier, V. Wolf, P. Grosperrin, E. Beuvier, O. Rolet-Répécaud, S. Gavoye, Y. Gauzere, O. Belysheva, E. Notz, D. Boichard, and A. Delacroix-Buchet. 2019. Comparison of Bayesian and partial least squares regression methods for mid-infrared prediction of cheese-making properties in Montbéliarde cows. J. Dairy Sci. 102:6943–6958. https://doi.org/10.3168/jds.2019-16320.

Fernando, R., H. Cheng, B. Golden, and D. Garrick. 2016. Computational strategies for alternative single-step Bayesian regression models with large numbers of genotyped and non-genotyped animals. Genet. Sel. Evol. 48:96. https://doi.org/10.1186/s12711-016-0273-2.

Ferrand, M., G. Miranda, S. Guisnel, H. Larroque, O. Leray, F. Lahalle, M. Brochard, and P. Martin. 2012. Determination of protein composition in milk by mid-infrared spectrometry. Pages 41–45 in Proc. International Strategies and New Developments in Milk Analysis. VI ICAR Reference Laboratory Network Meeting.

Ganai, N. A., H. Bovenhuis, J. A. van Arendonk, and M. H. Visker. 2009. Novel polymorphisms in the bovine beta-lactoglobulin gene and their effects on beta-lactoglobulin protein concentration in milk. Anim. Genet. 40:127–133. https://doi.org/10.1111/j.1365-2052.2008.01806.x.

Gengler, N., H. Soyeurt, F. Dehareng, C. Bastin, F. Colinet, H. Hammami, M.-L. Vanrobays, A. Lainé, S. Vanderick, C. Grelet, A. Vanlierde, E. Froidmont, and P. Dardenne. 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. J. Dairy Sci. 99:4071–4079. https://doi.org/10.3168/jds.2015-10140.

Goddard, M., and B. Hayes. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet. 10:381–391. https://doi.org/10.1038/nrg2575.

Hsu, W. L., D. J. Garrick, and R. L. Fernando. 2017. The accuracy and bias of single-step genomic prediction for populations under selection. G3 (Bethesda) 7:2685–2694. https://doi.org/10.1534/g3.117.043596.

International Dairy Federation. 2016. Dairy's global impact. Accessed Nov. 10, 2021. https://fil-idf.org/dairys-global-impact.

Malacarne, M., P. Formaggioni, P. Franceschi, and A. Summer. 2004. Seasonal variations of milk quality in Parmigiano-Reggiano cheese manufacture on a period of 10 years. Sci. Tecn. Latt. Cas. 55:63–77.

Martin, P., M. Szymanowska, L. Zwierzchowski, and C. Leroux. 2002. The impact of genetic polymorphisms on the protein composition of ruminant milks. Reprod. Nutr. Dev. 42:433–459. https://doi.org/10.1051/rnd:2002036.

Meyer, K. 2007. WOMBAT - A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J. Zhejiang Univ. Sci. B 8:815–821. https://doi.org/10.1631/jzus.2007.B0815.

Pretto, D., N. Lopez-Villalobos, M. Penasa, and M. Cassandro. 2012. Genetic response for milk production traits, somatic cell score, acidity and coagulation properties in Italian Holstein-Friesian population under current and alternative selection indices and breed-

ing objectives. Livest. Sci. 150:59–66. https://doi.org/10.1016/j.livsci.2012.08.001.

Sanchez, M. P., M. El Jabri, S. Minéry, V. Wolf, E. Beuvier, C. Laithier, A. Delacroix-Buchet, M. Brochard, and D. Boichard. 2018a. Genetic parameters for cheese-making properties and milk composition predicted from mid-infrared spectra in a large dataset of Montbéliarde cows. J. Dairy Sci. 101:10048–10061. https://doi.org/10.3168/jds.2018-14878.

Sanchez, M. P., S. Fritz, C. Patry, A. Delacroix-Buchet, and D. Boichard. 2020. Frequencies of milk protein variants and haplotypes estimated from genotypes of more than 1 million bulls and cows of 12 French cattle breeds. J. Dairy Sci. 103:9124–9141. https://doi.org/10.3168/jds.2020-18492.

Sanchez, M. P., A. Govignon-Gion, P. Croiseau, S. Fritz, C. Hozé, G. Miranda, P. Martin, A. Barbat-Leterrier, R. Letaïef, D. Rocha, M. Brochard, M. Boussaha, and D. Boichard. 2017. Within-breed and multi-breed GWAS on imputed whole-genome sequence variants reveal candidate mutations affecting milk protein composition in dairy cattle. Genet. Sel. Evol. 49:68. https://doi.org/10.1186/s12711-017-0344-z.

Sanchez, M. P., Y. Ramayo-Caldas, V. Wolf, C. Laithier, M. El Jabri, A. Michenet, M. Boussaha, S. Taussat, S. Fritz, A. Delacroix-Buchet, M. Brochard, and D. Boichard. 2019. Sequence-based GWAS, network and pathway analyses reveal genes co-associated with milk cheese-making properties and milk composition in Montbéliarde cows. Genet. Sel. Evol. 51:34. https://doi.org/10.1186/s12711-019-0473-7.

Sanchez, M. P., V. Wolf, M. El Jabri, E. Beuvier, O. Rolet-Répécaud, Y. Gaüzère, S. Minéry, M. Brochard, A. Michenet, S. Taussat, A. Barbat-Leterrier, A. Delacroix-Buchet, C. Laithier, S. Fritz, and D. Boichard. 2018b. Short communication: Confirmation of candidate causative variants on milk composition and cheesemaking properties in Montbéliarde cows. J. Dairy Sci. 101:10076–10081. https://doi.org/10.3168/jds.2018-14986.

Sargolzaei, M., J. Chesnais, and F. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives. BMC Genomics 15:478. https://doi.org/10.1186/1471-2164-15-478.

Soyeurt, H., D. Bruwier, J. Romnee, N. Gengler, C. Bertozzi, D. Veselko, and P. Dardenne. 2009. Potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. J. Dairy Sci. 92:2444–2454. https://doi.org/10.3168/jds.2008-1734.

Tribout, T., V. Ducrocq, and D. Boichard. 2020. HSSGBLUP: A single-step SNP BLUP genomic evaluation software adapted to large livestock populations. In ICQG, Brisbane, Australia.

VanRaden, P. M. 2004. Invited review: Selection on net merit to improve lifetime profit. J. Dairy Sci. 87:3125–3131. https://doi.org/10.3168/jds.S0022-0302(04)73447-5.

ORCIDS

M. P. Sanchez https://orcid.org/0000-0002-1371-5342
T. Tribout https://orcid.org/0000-0003-2500-7375
S. Fritz https://orcid.org/0000-0002-3378-1048
D. Boichard https://orcid.org/0000-0003-0361-2961