Jourdain, J., Besnard, F., Barbat, A., Boussaha, M., Fritz, S., Chambrial, M., Fayolle, G., Gamarra Lacaze, G., Lacaze, S., Schibler, L., Dutheil, L., Foucras, G., Boichard, D., Capitan, A. (2026) Uncovering recessive fertility disorders in dairy cattle populations using large-scale genomic and phenotypic data, *Proceedings 13th World Congress on Genetics Applied to Livestock Production*
24077: https://doi.org/10.31274/wcgalp.24077 ©2026 Jourdain, et al. CC BY-NC-ND

Uncovering recessive fertility disorders in dairy cattle populations using large-scale genomic and phenotypic data

J. Jourdain[1,2,3], F. Besnard[1,2], A. Barbat[1], M. Boussaha[1], S. Fritz[1,2], M. Chambrial[4], G. Fayolle[5], G. Gamarra Lacaze[6], S. Lacaze[6], L. Schibler[7], L. Dutheil[8], G. Foucras[8], D. Boichard[1], and A. Capitan[1,*]

[1]Université Paris-Saclay, INRAE, AgroParisTech, GABI, Jouy-en-Josas, France; [2]ELIANCE, Paris, France; [3]SYNETICS France SAS, Noyal-sur-Vilaine, France; [4]ORIGENPLUS, L'Aigle, France; [5]UMOTEST, Ceyzeriat, France; [6]AURIVA-ELEVAGE , Denguin, France; [7]GENES DIFFUSION, Douai, France; [8]IHAP, Université de Toulouse, INRAE, ENVT, Toulouse, France; [*]Corresponding author: aurelien.capitan@inrae.fr

Introduction

Reproductive efficiency is a cornerstone of dairy herd sustainability, influencing the timing of lactation and the production of replacement and market animals. However, between 1970 and 2010, female fertility has experienced a steady decline in high-producing breeds, commonly attributed to long-term selection for milk yield combined with the low heritability (1–5%) and high environmental sensitivity of fertility traits (e.g., Ma *et al*., 2019). Although the integration of functional traits into selection indexes and genomic evaluations have helped to halt this trend over the past decade, restoring optimal reproductive performance remains a significant challenge. Genome-wide association studies have revealed a highly polygenic architecture for traits such as conception rate and calving interval, with numerous loci exerting modest effects and collectively explaining only a limited proportion of phenotypic variance (e.g., Ma *et al*., 2019; Liang *et al*., 2023). We therefore hypothesized that an increase in the frequency of hidden recessive variants may have substantially contributed to the historical decline in fertility. To test this, we adapted our Homozygous Haplotype Enrichment and Depletion (HHED) approach —developed to identify overlooked genetic defects by mining life-history data from large cohorts of genotyped animals (Besnard *et al*., 2024)— and applied it to the repeat breeder phenotype in three dairy cattle breeds.

Materials & Methods

HHED mapping. From the pool of genotypes generated for genomic selection, we selected individuals with three or more unsuccessful inseminations (repeat breeder group) and those that conceived at the first insemination (control group), each considered separately for heifers and first-parity cows. To minimize artifacts due to population stratification between groups, individuals were pruned and randomly removed to achieve a case-to-control ratio of 1:4 for heifers and 1:3 for cows within each paternal half-sib family. These family-balanced datasets included 181,020, 67,475 and 6005 heifers and 137,668, 61,656, and 5,204 first-lactation cows in the Montbéliarde, Normande, and Brown Swiss breeds, respectively. Each dataset was partitioned into a discovery population (80%) and an independent validation population (20%), while preserving the familial structure.

Phased Illumina BovineSNP50 genotypes, processed through the French genomic evaluation pipeline, were obtained for each individual, along with those of their sire and dam (or maternal grandsire). Using sliding windows of 20 markers (~1 Mb), the observed number of homozygotes (Nobs) was recorded for each haplotype present in the discovery population, while the expected number (Nexp) was calculated from ancestral genotypes as described in Besnard *et al*. (2024). As a prefiltering step, haplotypes were selected if they met the

following criteria: Nobs ≥ 10 in the repeat breeder group, enrichment in homozygotes relative to expectation (Nobs/Nexp ≥ 1) in that group, and depletion in the controls (Nobs/Nexp ≤ 1). These prefiltered haplotypes were evaluated using a 2×2 chi-square contingency test, comparing the number of observed homozygotes and non-homozygotes in both groups. The same test was performed with expected counts, as a control. Chi-square values were corrected for multiple testing using Bonferroni adjustment, accounting for 29,903 to 99,761 haplotypes tested depending on the analysis (divided by 20 to adjust for non-independence among overlapping haplotypes). Haplotypes with an adjusted p-value below 0.05 for the observed counts and above 0.05 for the expected counts were considered significant. In the validation population, only the lead haplotypes of significant loci identified in the discovery population were evaluated. Lead haplotype frequencies were estimated from 30,909 Brown Swiss, 252,606 Normande, and 838,283 Montbéliarde cattle genotyped between 2010 and 2025.

Search for candidate variants. Variants within a 5-Mb window around the most significant haplotype at each locus were extracted from a VCF file comprising 195 Brown Swiss, 160 Montbéliarde, 118 Normande and 1,396 additional sequenced animals from more than 70 breeds. Linkage disequilibrium between haplotypes and genotypes was estimated using the squared Pearson correlation coefficient (R^2). Breed-specific variants with $R^2 \geq 0.5$ were annotated using the Ensembl Variant Effect Predictor (VEP, release 110) and putatively deleterious variants (missense with a SIFT score ≤ 0.05, stop gain, frameshift, splice-site variants) were retained as candidates. Pathophysiological insights into the homozygous inactivation of orthologous and paralogous genes were obtained from a literature review.

Effects on traits. According to breed, yield deviations pre-adjusted for environmental effects were obtained from French national genetic evaluations (https://www.geneval.fr) for 41 to 49 production, morphology, health and reproduction traits. Linear mixed models fitted with the BLUPF90 software suite (Misztal *et al*., 2014) included haplotype status and year of recording as fixed effects and a random polygenic effect. Cohort sizes ranged from 20 to 475,497 animals, depending on haplotype status and trait. Differences between heterozygotes or homozygotes and non-carriers, were evaluated using t-tests with Benjamini–Hochberg correction to control the false discovery rate. In addition, multiple ovulation and embryo transfer (MOET) data on donor and recipient cow IDs, embryos and unfertilized oocytes collected, embryo quality, transfer details, and gestation outcomes were obtained from industry partners. Differences between haplotype status groups were assessed using t-tests and Fisher's exact tests (see Fig. 1d for cohort size details).

Results & Discussion

HHED analysis identified 13 recessive loci significantly associated with the repeat breeder phenotype in heifers and first-parity cows across breeds (Table 1). These loci displayed a spectrum from rare haplotypes with strong effects on Nobs/Nexp (e.g., B17.67H and N15.64H) to more frequent haplotypes with smaller effects (e.g., M23.25H and N6.37C).

To explore the molecular mechanisms underlying these recessive repeat breeder loci, we mined a large dataset of whole-genome sequences and identified candidate deleterious variants for 10 of 13 loci. Among affected genes, only *OVGP1* has a known role in female fertility, being essential for fertilization and early embryonic development in mammals (Yamatoya *et al*., 2024). Other genes including *TDRD12* (critical for spermatogenesis; Bao *et al*., 2024) and *CRISP1* (involved in sperm–egg interactions; Gonzalez *et al*., 2021) affect male fertility, and may also influence female fertility, although this remains speculative. Additional candidates impact biological functions essential for reproductive success: *ASCC2*

(DNA repair), *GFER* (mitochondrial function), *PIK3C3* (autophagy and follicle development), *LCORL* (growth), *DEPDC7* (cell proliferation and NF-κB signaling), and *CCDC61* and *CCDC57* (members of a protein family linked to ciliary and meiotic functions).

Table 1. Lead haplotypes from the HHED analysis and associated candidate variants

Haplo[1]	f[2]	Nobs/Nexp		Raw p_obs	Candidate variant	Type[3]	Gene
		Cases	Controls				
B17.67H	6.3	2.83	0.07	1.4×10^{-18}	Chr17 g.68932869G>A	Ms	*ASCC2*
M3.31H	5.5	2.92	0.54	1.6×10^{-147}	Chr3 g.31900060Del	Fs	*OVGP1*
M3.31C	5.5	2.30	0.48	1.4×10^{-60}			
M4.98H	1.4	2.55	0.25	3.1×10^{-12}	-	-	-
M18.43H	3.4	2.28	0.56	5.9×10^{-32}	Chr18 g.43231574Del	Fs	*TDRD12*
M18.43C	3.4	1.75	0.60	2.3×10^{-14}			
M18.54H	3.8	2.91	0.37	3.0×10^{-128}	Chr18 g.53505101G>T	Sp_a	*CCDC61*
M18.62C	7.8	1.22	0.81	4.8×10^{-12}	-	-	-
M23.25H	18.5	1.11	0.94	2.3×10^{-08}	Chr23 g.22527795C>G	Ms	*CRISP1*
M24.13H	15.5	1.07	0.89	4.1×10^{-08}	Chr24 g.13968541C>T	Sp_r	*PIK3C3*
N5.84H	7.1	2.54	0.52	1.4×10^{-92}	-	-	-
N5.84C	7.1	1.84	0.39	4.1×10^{-42}			
N6.37C	16.5	1.05	0.78	1.6×10^{-07}	Chr6 g.37401397G>T	Ms	*LCORL*
N15.64H	4.3	3.10	0.14	5.2×10^{-72}	Chr15 g.63682108T>A	Ms	*DEPDC7*
N15.64C	4.2	1.45	0.10	1.4×10^{-16}			
N19.50H	12.6	2.18	0.71	8.3×10^{-135}	Chr19 g.50711150C>T	Sg	*CCDC57*
N25.1H	6.8	2.24	0.63	6.9×10^{-42}	Chr25 g.1541663G>A	Ms	*GFER*
N25.1C	6.8	2.29	0.30	5.8×10^{-83}			

[1] Haplotype names are constructed as follows: first letter of the breed, chromosome number and position in Mb on the ARS-UCD1.2 assembly separated by ".", and first letter of parity (H = heifer, C = cow in first lactation).
[2] f: haplotype frequency in percent.
[3] Types of variant : Fs = frameshift, Ms = missense, Sp_a = splice acceptor, Sp_r = splice region, Sg = Stop gain

In addition, we studied the effects of the lead haplotypes on more than forty recorded traits, including five fertility traits: interval between first and last AI (AI1–n), non-return rates in heifers and cows (NRR_h and NRR_c), and conception rates in heifers and cows (CR_h and CR_c). With the exception of M18.62-*XXX* and N6.37-*LCORL*, which displayed significant effects (q-value ≤ 0.05) on numerous traits even in heterozygotes (Fig. 1a), most loci acted recessively. These loci predominantly affected fertility traits, and when they influenced other traits, the effects were generally smaller in absolute magnitude when expressed in genetic standard deviations (Fig. 1b,c). Finally, we analyzed MOET data for the most significant locus in each breed (B17.67-*ASCC2*, M3.31-*OVGP1*, and N19.50- *CCDC57*; Fig. 1d). The average number of "structures" (embryos or unfertilized oocytes) collected was similar between homozygous and wild-type donors. However, their nature, quality, and/or developmental success to calving when transferred to wild-type recipient dams differed significantly depending on the locus. Conversely, homozygous females used as recipient dams receiving wild-type embryos exhibited normal conception rates, suggesting that overall reproductive physiology is intact and that defects are limited to gamete and early embryonic development. These results indicate that homozygous dams can be used as recipients to mitigate their subfertility and allow both calving and lactation induction. Moreover, they provide a basis for functional studies to confirm the role of *OVGP1* in cattle and to investigate the contributions of *ASCC2* and *CCDC57* to female infertility in mammals using the bovine model.

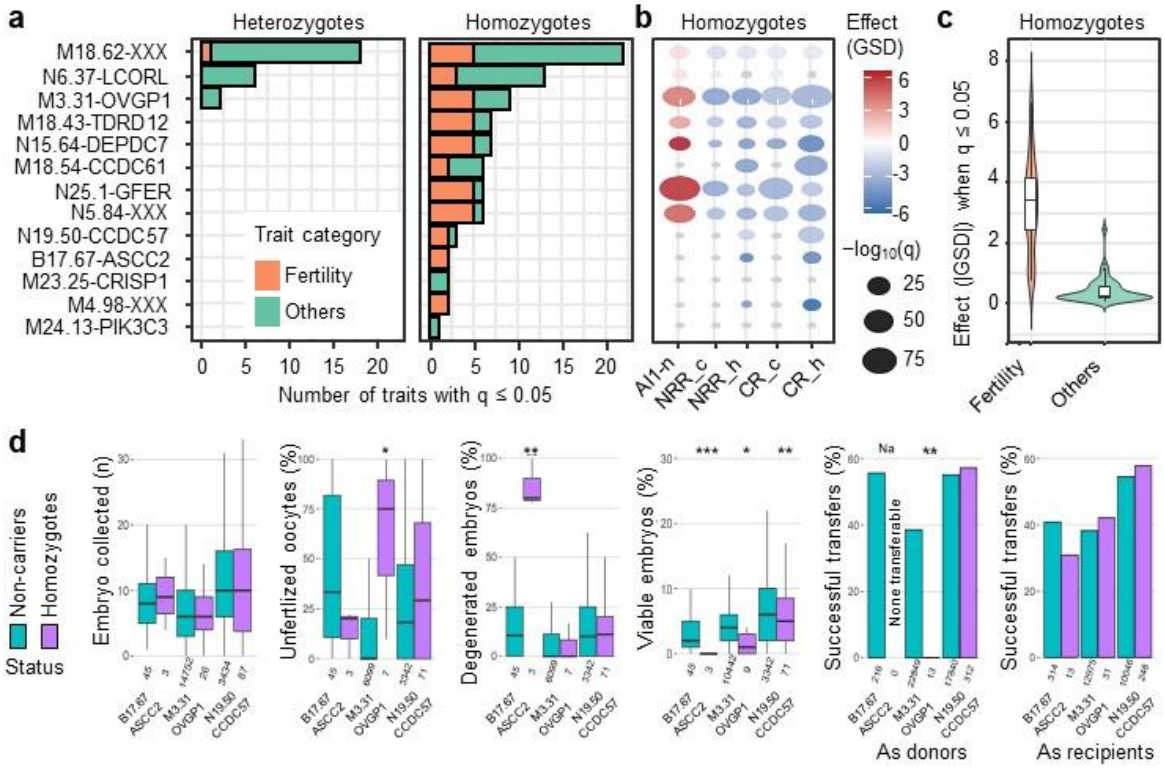

Figure1. Effects of haplotype statuses on recorded traits relative to wild type. a) Number of traits with significant effects per locus × haplotype status; b) fertility trait effects in homozygotes; c) effect sizes on fertility vs. other traits; d) MOET outcomes for three loci.

Conclusion

In conclusion, we identified 13 previously overlooked recessive loci that impair female fertility in dairy cattle. Accounting for these loci in breeding programs could tangibly improve reproductive performance, thereby ameliorating breeders' outcomes and reducing the environmental footprint of the dairy sector. These findings also open promising avenues to discover novel genes involved in human infertility, a major global health challenge.

Acknowledgements

This work was funded by APIS-GENE and the ANRT (Paris, France). We thank Eliance, its member companies, and Valogene for facilitating access to certain data.

References

Bao Z., Wang Y., Wang R., *et al*. (2024). J Genet Genomics 51:1322-1326. doi.org/10.1016/j.jgg.2024.06.014
Besnard F., Guintard A., Grohs C., *et al*. (2024). Genome Biol 25:248. doi.org/10.1186/s13059-024-03384-7
Gonzalez S.N., Sulzyk V., Weigel Muñoz M., *et al*. (2021). Front Cell Dev Biol 9:800351. doi.org/10.3389/fcell.2021.800351
Ma L., Cole J.B., Da Y., and VanRaden P.M. (2019). J Dairy Sci 102:3735–3743. doi.org/10.3168/jds.2018-15269
Misztal I., Tsuruta S., Lourenco D., *et al*. (2014). Manual for BLUPF90 family of programs. Available at: https://nce.ads.uga.edu/html/projects/programs/docs/blupf90_all8.pdf, visited on: 2026/01/07.
Liang Z., Prakapenka D., VanRaden P.M., *et al*. (2023) Int J Mol Sci 24:10496. doi.org/10.3390/ijms241310496
Yamatoya K., Kurosawa M., Hirose M., *et al*. Biol Reprod 110:465–475. doi.org/10.1093/biolre/ioad159