Rumigen

TOWARDS IMPROVEMENT OF **RUMI**NANT BREEDING
THROUGH **GEN**OMIC AND EPIGENOMIC APPROACHES

Book of Abstracts

The RUMIGEN project has received funding from European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and not the European Commission (EC). The EC is not liable for any use that may be made of them information contained herein.

Summary

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

Editorial by Eric Pailhoux, *Rumigen* project coordinator

It is my great pleasure to briefly introduce our RUMIGEN project. The aim of RUMIGEN was to develop or improve sustainable and socially acceptable breeding programs and management practices that would produce future generations of efficient and resilient animals while preserving genetic diversity. The project also sought to ensure that the proposed breeding scenarios were acceptable to stakeholders and citizens.

RUMIGEN aimed to address the issue of trade-offs by adapting dairy cows to heat stress. Statistical analyses were performed on large-scale data from commercial farms, including information on performance, pedigrees and genotypes, combined with meteorological data. Understanding the genomic regions that contribute to these trade-offs could inform the development of breeding strategies that optimize genetic improvement in terms of efficiency and resilience. The project also explored whether epigenetics could facilitate more efficient selection and precision agriculture, assessing the transmission of epigenetic marks to the next generation and studying how these marks reflect adaptation to the environment. Developing a high-throughput, inexpensive epigenetic tool (the RUMIGEN EpiChIP), was essential in order to address these questions with sufficient statistical power, and to gain an understanding of how complex traits are regulated, and which genetic and environmental stressors can modulate the epigenome. This knowledge should contribute to the design of more effective and sustainable practices, thereby improving animal welfare.

RUMIGEN has also developed new genomic tools that can distinguish between genetic diversity, genetic drift, genetic load and genomic polymorphisms, such as structural variants. These tools will enable the impact of genomic selection programs to be assessed and optimized, the genetic load and inbreeding depression to be identified, and breed-specific genomic polymorphisms to be characterized. Additionally, new whole genome sequences have been produced for various local cattle breeds. In addition, RUMIGEN has used genome editing to evaluate its potential as a new tool for improving genetic selection and diversity, particularly in local breeds, in both theoretical and practical terms for small ruminant species.

Ultimately, we hope that future breeding programs will be informed by the knowledge gained by RUMIGEN, while also taking into account the views of stakeholders and citizens on new technologies, as well as the latest information in the fields of genomics and epigenomics, as set out in the “Room of Acceptance” tool.

This booklet contains all the presentations given as part of the RUMIGEN project at various conferences. Additionally, all publications resulting from the project are listed alongside their associated abstracts. Enjoy reading!

The RUMIGEN project has received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

Presentations at conferences

ICAR INTERBULL MEETING 2022

(Montreal, Canada — May 30–June 3, 2022)

1. Towards new breeding tools in a context of climate change: first results of the RUMIGEN project on new phenotypes for heat tolerance traits

S. Mattalia[1], A. Vinet[2], J. Vandenplas[3], M.J. Carabaño[4], C. Diaz[4], M. Ramon[5], S. Aguerre[1], B.C.D. Cuyabano[2], D. Boichard[2], E. Pailhoux[2], H.A. Mulder[3]
1Institut de l'Élevage, UMT eBIS, UMR1313 GABI, Domaine de Vilvert, 78350 Jouy-en-Josas, France
2Université Paris-Saclay, INRAE, AgroParisTech, UMR1313 GABI, Domaine de Vilvert, 78350 Jouy-en-Josas, France
3Wageningen University & Research, Animal Breeding and Genomics, PO Box 338, 6700 AH Wageningen, The Netherlands
4Departamento de Mejora Genética Animal, INIA, 28040 Madrid, Spain
5Instituto Regional de Investigación y Desarrollo Agroalimentario y Forestal de Castilla-La Mancha (IRIAF-JCCM), 13300 Valdepeñas, Spain

RUMIGEN is a project financially supported by the EU that aims to develop breeding programs capable of managing the trade-offs between efficient production and resilience to extreme climate conditions. RUMIGEN is designed under a multi-disciplinary approach that mixes competencies in both genetics and social sciences. The genetic approach aims to enhance genomic selection using three levers: quantitative genetics, genome editing, and epigenetics. One of the objectives of RUMIGEN is to enlarge selection criteria and to provide genomic tools to select dairy cows tolerant to heat stress. Studies are dedicated to the definition of heat-tolerance traits based on production, reproduction and health records, as well as to the study of the trade-offs between these traits, and with those already selected. These analyses are based on performances recorded in commercial herds in France, Spain and the Netherlands (i.e., milk production traits and somatic cell scores recorded by Milk Recording Organizations, and fertility traits derived from AI events), in combination with meteorological data obtained from the corresponding Meteorology Agencies. Records are associated to meteorological information at the farm level, in order to measure the impact of heat stress. First results obtained for different breeds and in a large range of farming and climatic scenarios showed that the combination of both types of information is relevant to measure the decline of performances due to heat stress and to define heat stress indicators and new traits for future breeding tools.

WCGALP 2022

12th World Congress On Genetics Applied To Livestock Production
(Rotterdam, The Netherlands — July 3–8, 2022)

2. Genotype by temperature-humidity index interactions on milk production and udder health traits in the Montbeliarde cows

A. Vinet[1], S. Mattalia[2],[1] R. Vallée[2],[1] C. Bertrand[3], P. Bertuzzi[4], B. C. D. Cuyabano[1], D. Boichard[1]
1Université Paris Saclay, INRAE, AgroParisTech, UMR1313 GABI, Domaine de Vilvert, 78350 Jouy-en-Josas,

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

France
2Institut de l'Élevage, UMT eBIS, 75012 Paris, France
3INRAE, US 310 CTIG, Domaine de Vilvert, 78350 Jouy-en-Josas, France
4INRAE, US 1116 AGROCLIM, Domaine de Saint Paul, 84914 Avignon, France

This study, based on the joint analysis of on-farm performance and weather data, had two objectives: (1) to determine the response in production performances (milk, fat and protein yields, fat and protein contents) and udder health (somatic cell score) to temperature-humidity index (THI) variations in Montbeliarde cows; and (2) to estimate the interactions between genotype and THI, to enable the identification of the most adapted animals for facing the expected increases in temperature. Test-day records from first and second lactations from 2016 to 2020 were associated with the average THI during the three days before the test-day record. In total, 446,717 test-day records from 55,650 cows in first lactation and 457,516 test-day records from 58,229 cows in second lactation were analysed. The optimal THI was below 55 (i.e. ~ 12–13 °C) for all traits. Individual responses to THI were estimated by random regression models, which also included individual responses to days in milk. Regardless of the stage of lactation, genetic correlations along the THI gradient were above 0.80, which suggests that genotype-by-THI interactions were weak for production and udder health traits. Nevertheless, a variability in the individual slope of decay could be highlighted at high THI. The genetic correlation between production level at moderate THI and the slope at high THI was negative, while for somatic cell score, it was positive, indicating that heat stress amplifies the susceptibility to mastitis. The optimal THI for French Montbeliarde cows is below 55 for production and udder health traits. Genetic-by-THI interactions are weak in French Montbeliarde cows for production and udder health traits, but not all animals react in the same way to high temperatures. Even if there is little room for improvement, using a heat tolerance index in cattle selection would be relevant to anticipate the expected increases in temperature. Further investigations are needed to interpret this variability on production traits. However, the current selection for mastitis resistance seems appropriate to adapt cattle to rising temperatures.

French Academy of Veterinary Sciences Conference

(Paris, France — January 26, 2023)

3. Genome editing in farm animals: where we are in France and in Europe?*

J. Vilotte, E. Pailhoux
INRAE, Université Paris-Saclay, AgroParisTech, GABI / BREED, Jouy-en-Josas, France

New nucleases make it possible to modify the genome of many species in a targeted manner. With regard to livestock, the applications envisaged can be classified into three main categories: (i) projects aimed at fundamental research, particularly when model species (mouse, zebrafish, etc.) cannot be used (for horn development, seasonality, rumination, as examples); (ii) projects aimed at biomedical research to create animal models of human pathologies; and (iii) projects aimed at agronomy to bring a favorable trait, described in another species or breed, to a breed or species that does not have this trait. European legislation classifying animals obtained by this way as GMOs limits the development of these approaches and the funding of associated research. We will illustrate here, through the example of one of the rare European projects using these tools, their potential, their limits and the paradox induced by regulations that seem inappropriate, or at least less appropriate than those put in place in other regions of the world.

*Similar communications on genome-edited animals developed during the RUMIGEN project were also presented at the following events:

1. French Academy of Agriculture on June 13, 2023, in Paris, France
2. EcoGen episode 7 Webinar on November 28, 2025.

AIDA 2023

XX Jornadas Sobre Producción Animal (Zaragoza, Spain — May 16–17, 2023)

4. Comparative of nanopore and bisulfite sequencing in the detection of methylation marks on the dairy cattle epigenome

A. López-Catalina, O. González-Recio
INIA-CSIC, Madrid 28040, Spain.

The epigenome works as the intermediary between the environment and the genome. Gene expression can be altered by environmental changes such as certain stressors as temperature or metabolic status of the individual. Studying these epigenetics marks could help us to better understand gene expression and regulation, prenatal programming, and the effect of stressors on the phenotype, leading us to improve the accuracy of genetic evaluations. It is of special interest in farm animals and agrigenomics. However, epigenotyping a whole population is an expensive and difficult process. Epi chips or low pass sequencing could be a suitable strategy to determine epigenetic marks in target animals within a population. In this work, we compared DNA methylation from a blood sample from a heat-stressed cow using bisulfite sequencing and Nanopore Technology sequencing (MinION). We compared the methylated cytosines called with both techniques, using the software Bismark for bisulfite reads and Guppy, Nanopolish and Methylartist for the MinION reads. Bisulfite sequencing led to 584,135,436 sequences with a mean length of 151 bp, translating into a coverage of approximately 30X. ONT sequencing achieved 2,618,753 reads with a mean length of 10,061 bp, translating into a coverage of 7X. We studied the methylated sites detected by both techniques and obtained a correlation above 0.60 for the methylation status of these genomic regions. Nanopore sequencing allows us to detect methylated regions at a lower coverage. This information can be used to design tools that target specific methylated regions within a population. Further studies are being conducted to understand and point the exact differences between bisulphite and nanopore sequencing.

ICAR INTERBULL MEETING 2023

(Lyon, France — August 26–28, 2023)

5. Meta-Analysis for Heat Tolerance Traits in Holstein in France, the Netherlands and Spain

S. Aguerre[1], R. Vallée[1], A. Vinet[2], J. Promp[1], M.J. Carabaño[3], C. Diaz[3], M. Ramon[4], B.C.D.Cuyabano[2], D. Boichard[2], H.A. Mulder[5], J. Vandenplas[5], S. Mattalia[1]
1Institut de l'Elevage, UMT eBIS, UMR1313 GABI, Domaine de Vilvert, 78350 Jouy-en-Josas, France
2Université Paris Saclay, INRAE, AgroParisTech, UMR1313 GABI, Domaine de Vilvert, 78350 Jouy-en-Josas, France
3Departamento de Mejora Genética Animal, INIA-CSIC, 28040 Madrid, Spain

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

4Instituto Regional de Investigación y Desarrollo Agroalimentario y Forestal de Castilla-La Mancha (IRIAF-JCCM), 13300 Valdepeñas, Spain
5Wageningen University & Research, Animal Breeding and Genomics, PO Box 338, 6700 AH Wageningen, the Netherlands

In the context of climate change, livestock production systems face the challenge of ensuring that, although more and more exposed to heat-stress conditions, animals will be able to remain healthy while maintainingsatisfactoryproduction, responding to consumers' demand. As part of the European project RUMIGEN, France (INRAE, Idele), Spain (INIA, IRIAF), and the Netherlands (WUR) studied the response of different dairy traits to heat-stress conditions for Holstein cattle breed. Production(milk, fat and protein yields) and udder health (somatic cell score) traits were investigated under different meteorological conditions,represented by the temperature humidity index (THI) averaged over three days on the day of recordingand the twodays before. A THI=50 was considered as neutral condition (i.e. no heat/cold-stress) and used as reference for level comparisons. Specific rates of changes in production/health traits at a given THI were measured as the slopes (first derivative) of the estimated reaction-norm curves for each trait. Genetic evaluations were performed by each country using test-day records of cows in their first lactation between 2010 and 2020, to estimate both levels and slopes for each trait. Estimated breeding values (EBVs) and reliabilities were obtained for sires with at least 20 daughterswith test-day records.A meta-analysis was performed to estimate the genetic correlations between the three countries, using the Multiple Across Country Evaluationapproach. For each country, de-regressed proofs (DRPs) and effective record contributions (ERCs) were computed using a single trait model from the EBVs, reliabilities, and variance components estimated at a national level for specific THI values. The estimated genetic correlations obtainedwith slopes were weak and not credible, which could be explained by thelow heritability of the slopes and by the small proportion of performances recorded under heat-stress conditions. The DRPs on slopes, as they were calculated, were not able to capture the genetic (co)variability associated to these traits. Conversely, estimated genetic correlations for all level traits werehigh (between 0.81and 0.97) even if they were slightly lowerthan under thermo-neutral conditions. This shows a very good consistency of the three national genetic evaluations under heat-stressconditions. In conclusion, valuable predictions under heat-stressconditions could be obtained through international evaluations, that would result in Northern countries benefiting from the information that already exists in the Southern countries. However, the approach for low heritable traits such as slopes should be improved if there are not enough data at high THI.

6. Genetic correlations: a parameter or a latent phenotype in genetic evaluations?

SB.C.D. Cuyabano[1], P Croiseau[1], F Shokor[1],[2], MR Motta[3], S Aguerre[1],[4], andS Mattalia[1],[4]
1INRAE–GABI, Université Paris-Saclay, AgroParisTech, France
2Eliance, UMT eBIS, France
3University of Campinas, SP, Brazil
4Idele, UMT eBIS, France

Genetic correlations are relevant parameters in genetic evaluations, particularly when a breeding program aims to achieve genetic progress for multiple traits altogether. These correlations are usually

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information therein.

estimated from a base population as one ofthe many parameters that define the distribution used to predict breeding values for the selection candidates. In such a fashion, genetic correlationsare assumed to be identical for all selection candidates. However, with a preliminary study on the output predicted breeding valuesof sires with more than 500 daughters from the French Montbéliardepopulation, we observed that the genetic correlationamong daughters from different sires may differ substantially, i.e., different sires expresseddifferent genetic correlationsbetween traits through their daughters. Thus, if genetic correlationsare specific values inherent to each individual, they could be considered as a phenotype;in other words,genetic correlationsmay be the observableconsequence of a concealed regulatory trait guidingthe relationshipbetween observable traits. For antagonistic traits (e.g.production and fertility in dairy cattle), it is reasonable to believe that individuals on the extremes of the trade-off distribution are likely to present a low breeding value for this concealed regulatory trait. However, due to our inability to directly measure this potential regulatory trait, it can be considered a latent phenotype. Although a method to consider such hypothesisthat genetic correlations may be a latent phenotype is yet undefined, there is no doubt that such hypothesis has an impact on the medium to long-term perspectives of a breeding program, given its breeding goals. Hypothesizing that genetic correlations arelatent phenotypes, simulations can then be used to assess the genetic progress formultiple traits of interest in a breeding programover many generations, as well as to assess the trajectoryof genetic correlations between traits and the genetic progressof the latent regulatory phenotype driving such correlations.Such comprehension of the genetic progress forthe latent phenotype is of particular relevance, since a regulatory trait is likely to impact more than only two antagonistic traits, but manyof the traits selected for in a breeding program.

EAAP 2023

74th Annual Meeting of the European Federation of Animal Science
(Lyon, France — August 26–September 1, 2023)

7. Should we consider fertility when improving thermotolerance in dairy cattle?

M.J. Carabaño [1], C. Díaz [1], M. Ramon[2]
1INIA, Spain
2IRIAF, Spain

Breeding for adaptation to heat stress (HS) has been normally approached by the use of indicators derived from the decline in production traits (quantity and quality) under high temperatures in dairy species. However, indicators of impairment of other functions associated with HS, such as reproductive performance, which have been less studied, may be advantageous because the reproductive function is most affected by stress and because the antagonistic relationship between productive level and productive decline under HS. In this study, different indicators of heat tolerance through the use of artificial insemination (AI) outcome and meteorological data have been analysed and the relationship with traditional productivity indicators quantified. Data for conception rate (CR, measured as 0/1) in first lactation, first insemination from 913,493 Holstein cows in 7,858 herds in Spain were used to select the optimal phenotypic indicator for HT derived from fertility data. Ridge regression considering average temperatures for the 5,10,15 and 30 days pre and post AI were used to compare the alternative indicators and decide in which period the AI outcome is more sensitive

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

to the effects of heat. Substantial loss in CR was observed for all indicators beyond 20ºC of average temperature. Post AI indicators showed better goodness of fit than pre AI values. Genetic evaluations under a threshold random reaction norm model on CR using average temperature for 7 days post AI showed relevant variability in the slope of CR decay and an antagonistic relationship fertility level and slope of CR decay under HS. Correlations with heat tolerance derived from productive traits where close to null, indicating that selection for heat tolerance through maintaining productivity will not result in improving reproductive performance under HS and vice versa. This result points at the consideration of including both productive and reproductive declines when aiming at improving adaptation to high temperatures in dairy cattle. This study received funding from the European Union's Horizon 2020 research and innovation program under grant number 101000226 (Rumigen).

8. Genomic selection strategies and their potential to maintain rare alleles and de-novo mutations: a long-term assessment

M.F. Schrauf, Y.C.J. Wientjes, H.A. Mulder, J. Vandenplas
Wageningen University and Research, Animal Breeding and Genomics, P.O. Box 338, 6700AH Wageningen, Netherlands

Sustainable breeding programs need to balance short-term genetic improvement with the conservation of genetic diversity. While genomic selection has considerably increased the genetic gain for many breeding programs, consequences on diversity can be less desirable. This is particularly the case for rare alleles and de-novo mutations, as markers used in genomic selection are generally not strongly associated to rare alleles. Moreover, genomic selection allows for the selection of young individuals without records, thereby ignoring the effects of de-novo mutations. To study possible solutions, we simulated populations of 1000 individuals subject to 50 generations of selection. We evaluated four selection strategies to identify the ones which best conserve favourable rare variants and de-novo mutations in the presence of additive and non-additive gene action. The genomic selection strategies represent a variety of approaches to balance between genetic improvement and diversity management and were: truncation selection, which only focuses on short-term gain; optimal contribution selection, which balances that gain with a constraint in the relatedness of the selected individuals; allele-reweighted selection, which upscales the effect of rare alleles in the breeding values; and constrained allele loss selection, a novel strategy which balances short-term gain with a constraint on the reduction in frequency of rare alleles estimated to be favourable. Systematic differences between the strategies were not observed for traits with non-additive gene action. For the trait under additive gene action, allele-reweighted selection obtained a higher genetic gain than truncation selection while preserving a similar level of genetic variance. Meanwhile, constrained-allele-loss selection obtained a similar genetic gain than truncation selection while accumulating a higher number of favourable de-novo mutations. These and similar strategies may contribute to the sustainable long-term use of genomic selection. This project has received funding from the European Union’s Horizon 2020 Programme for Research & Innovation under grant agreement n°101000226.

9. Evaluation of heat stress effects on production traits and somatic cell score of Dutch Holstein cows

Jeremie Vandenplas[1], Mathijs van Pelt[2], Han Mulder[1]

1Wageningen University and Research, Animal Breeding and Genomics, P.O. Box 338, 6700AH Wageningen, Netherlands
2CRV, P.O. Box 454, 6800 AL, the Netherlands

Global climate changes are expected to have an impact on livestock production throughout the world. For example, heat waves should increase in frequency and intensity, resulting in temperate regions facing regularly hot periods. During these hot periods, dairy cows will suffer from heat stress, reducing their milk production, reproductive performance and welfare. Therefore, selecting heat-tolerant dairy animals is of interest to cope with future heat waves. Among several indicators for heat stress, one of them is the temperature-humidity index (THI) that combines the effects of temperature and humidity on animals. The first aim of this study is to evaluate heat stress effects on milk production traits and somatic cell score (SCS) of Dutch Holstein cows using a THI indicator. The second aim is to estimate genotype by THI interactions for the same traits to evaluate the magnitude of the individual variability in heat tolerance. Data included more than 1.5 million test-day records for milk, fat, and protein yields, and somatic cell score (SCS). Data were collected from around 500 thousand first and second-parity cows in 1,581 Dutch herds between 2010 and 2022. All test-day records were associated with 3-day average THI values computed from publicly available meteorological records. For both parities, decreases in milk yield are observed from a value of THI around 65 (i.e., between 18 and 23 degrees Celsius), and decreases in fat and protein yields are observed from a value of THI around 50. For SCS, increasing THI results in increasing SCS for the first parity (especially from a value of THI around 40), but does not impact SCS in the second parity. Individual responses to THI and to days-in-milk were estimated simultaneously using random regression models. Most genetic correlations along the THI gradient were higher than 0.90 at the 150th day-in-milk. Based on these results, genotype by THI interactions are weak for production traits and SCS of Dutch Holstein cows.This project has received funding from the European Union’s Horizon 2020 Programme for Research & Innovation under grant agreement n°101000226.

10. Impact of genomic selection on genetic diversity in five European local cattle breeds

J.J. Windig[1], T.H.E. Meuwissen[2], P. Croiseau[3], G. Restoux[3] and R. Bonifazi[1]
1Wageningen University & Research, Animal Breeding and Genomics, Droevendaalsesteeg 1, Radix, 6700 AH Wageningen, the Netherlands
2Norwegian University of Life Sciences, Department of Animal and Aquacultural Sciences, Oluf Thesensvei 6, 1433 Ås, Norway
3Université Paris-Saclay, INRAE, AgroParisTech, Joey-en-Josas, 78350, France

Genomic selection (GS) has revolutionized animal breeding and in general accelerated genetic gains in cattle breeding programs. However, in Holstein and Jersey cattle, the implementation of GS was accompanied by increases in inbreeding rates. Similarly to these popular breeds, GS has been implemented also for smaller populations and local breeds. In this study, we investigated inbreeding trends in five local cattle breeds from three European countries and evaluated the possible impact of implementing GS on breeds’ genetic diversity. The five breeds evaluated were: Abondance, Tarantaise, Vosgienne (from France), Norwegian Red (from Norway), and MRIJ (from the Netherlands). We estimated trends in inbreeding and kinship based on pedigree and genomic information (available as medium or high-density SNP genotypes), before and after the introduction of GS. The number of available genotyped animals for each breed was 16,478, 8,589, 4,474, 51,799, and 4,997, for Abondance, Tarantaise, Vosgienne, Norwegian Red, and MRIJ, respectively. Results show that inbreeding trends did not differ between pedigree-based or DNA-based estimates. In all five breeds, periods with higher and lower inbreeding rates occurred across time. However, no clear trend could be observed after the introduction of GS in any of the breeds. Inbreeding rates either slightly increased (MRIJ, Abondance), remained stable (Tarantaise, Vosgienne), or decreased (Norwegian Red). Such small deviations could be due to GS allowing for shorter generation intervals, screening for a larger number of individuals compared to traditional breeding schemes, and preselection of a smaller number of sires. Ultimately, our results suggest that the genetic management of these breeds is more important in determining inbreeding rates than the implementation of GS per se. This project has received funding from the European Union’s Horizon 2020 Programme for Research & Innovation under grant agreement n°101000226.

11. RUMIGEN: New breeding tools in a context of climate change: Impact of heat stress on the performances of dairy cows at the population level

S. Mattalia[1], A. Vinet[2], M.P.L. Calus[3], H.A. Mulder[3], M.J. Carabaño[4], C. Diaz[4], M. Ramon[5], S. Aguerre[1], J. Promp[1], R. Vallée[1], B.C.D. Cuyabano[2], D. Boichard[2], E. Pailhoux[6] and J. Vandenplas[3]
1Idele, Domaine de Vilvert, 78350 Jouy en Josas, France
2INRAE, GABI, Domaine de Vilvert, 78350 Jouy en Josas,France
3Wageningen University & Research, Animal Breeding and Genomics, P.O. Box 338, 6700 AH Wageningen, the Netherlands
4INIA, Departamento de Mejora Genética Animal, Ctra. de La Coruña km 7.5, 28040 Madrid, Spain
5IRIAF,13300, Valdepeñas, Spain
6INRAE, BREED, Domaine de Vilvert, 78350 Jouy en Josas, France

RUMIGEN is a project financially supported by the EU that aims to develop breeding programs able of managing the trade-offs between efficient production and resilience to extreme climate conditions. RUMIGEN is designed under a multi-disciplinary approach that mixes competencies in both genetics and social sciences. The genetic approach aims to enhance genomic selection using three levers: quantitative genetics, genome editing, and epigenetics. One of the objectives of RUMIGEN is to enlarge selection criteria and to provide genomic tools to select heat tolerant dairy cows. Studies are dedicated to the definition of heat-tolerance traits based on production, reproduction and health records, as well as to the study of the trade-offs between these traits, and with those already included in selection indexes. In France, Spain and the Netherlands, performances recorded in commercial herds (i.e., milk production traits, somatic cell scores and conception rate after first AI) were combined with meteorological data obtained from the nearest weather stations, and analysed in order to measure the impact of heat stress. First results obtained for different breeds and in a large range of farming and climatic scenarios showed that the combination of both types of information was relevant at the population level. Some differences between thresholds for optimal THI were observed between countries for some traits, which could be explained by different factors (farm management, exposure to outside temperatures, mitigation practices…). However, a decrease was observed for all performance traits with increasing temperatures, with consistent patterns of slopes between breeds and countries. Therefore this approach appears relevant to define novel traits related to heat tolerance.

12. A capture-based approach for DNA methylation analysis in cattle

V. Costes[1], A. López-Catalina[2], O. Gonzalez Recio[2], M. Ben Sassi[3], D.E. Rico[4], G. Foucras[5], C. Le Danvic[1], C. Hozé[1], M. Boussaha[6], D. Boichard[6], H. Jammes[6] and H. Kiefer[6]
1Eliance, 149 rue de Bercy, 75595 Paris cedex, France
2INIA-CSIC, Crta. de la Coruña km 7.5, 28040 Madrid, Spain
3Integragen, 5 rue Henri-Auguste Desbruères, 91030 Évry, France
4CRSAD, 1S0, Deschambault, QCG0A, Canada
5IHAP-ENVT, Université de Toulouse, 31076 Toulouse, France
6INRAE, Domaine de Vilvert, 78350 Jouy-en-Josas, France

DNA methylation is a potential source of phenotypic variation that may help refining genomic prediction in livestock, provided that affordable tools that can be used in routine analyses become available. Several methods have been developed for whole genome DNA methylation analysis, allowing important advances in the field of epigenetics. However, their cost is yet a limitation for their implementation in large populations. A cost-effective alternative is to target a panel of informative CpGs that show methylation variations across individuals. We assessed the potential of a capture-based approach to measure methylation at selected CpGs in the cattle genome. These CpGs were highlighted by reduced representation bisulfite sequencing (RRBS) conducted on 236 bull semen samples and 82 cow blood samples. Their methylation status in semen varied according to fertility, age, breed or early life nutrition; whereas it varied according to age, nutrition, physiological status and inflammatory status in blood samples. Furthermore, Oxford Nanopore Technology (ONT) was applied to 6 semen samples and 12 blood samples; and CpGs with differential methylation status according to fertility, somatic cell scores and heat stress were identified. The RRBS and ONT results were merged into a list of 128059 unique, non polymorphic CpGs. A custom design was conducted on the ARS-UCD1.2 genome assembly by Twist Bioscience, and 64237 probes targeting 97412 CpGs were synthesized. The technology was assessed on 71 semen, 91 blood and 30 mixed samples, for which RRBS and/or ONT data are available. Enzymatic conversion of DNA, capture, library construction and sequencing at a 100x depth on a Novaseq6000 were carried out by Integragen. Results about the repeatability and sensitivity of the method will be showed, as well as the consistency with RRBS and ONT data (H2020 grant 101000226, RUMIGEN).

13. Changes in genetic correlations over generations due to selection and random drift

BCD Cuyabano[1], S Aguerre[2], S Mattalia[2]
1Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France
2Idele, UMT eBIS, 75012 Paris, France

The evolution of genetic variance over generations is a topic that has been widely studied in quantitative genetics. In many breeding populations, both selection (resulting in the Bulmer effect) and limited population sizes (which builds up co-ancestry, resulting in random drift and inbreeding) result in the reduction of a trait's genetic variance. Until recently, most research related to this topic remained strictly theoretical, due to limitations in the availability of data over sufficient generations for many species. Moreover, studies regarding the consequences on the genetic correlations between traits under selection (as well as regarding their evolution over generations) are limited in the theoretical perspective, and to our knowledge, inexistent empirically. It is however, reasonable to expect that both the Bulmer effect and random drift should impact

genetic correlations between traits under selection. Furthermore, if changes are significant when compared to a base population (typically used for inferences of genetic parameters), they should be accounted for when performing genetic evaluations comprising the most recent generations of a breeding population. A preliminary study on synthetic data simulated to mimic that from a dairy cattle population with respect to production and fertility (h2=0.3 and h2=0.01, respectively, and genetic correlation -0.18) indicated that, as expected, when the population evolved through ten generations under random mating, although genetic variances decrease due to random drift, the genetic correlation remains unchanged. However, through ten generations under selection for either one of the traits individually, the genetic correlation was significantly attenuated. This attenuation was systematically progressive when selection was performed for production; when selection was performed for fertility, an immediate and relatively large attenuation is attained, then the genetic correlation is partially recovered and its trend follows the same as that of when selection was performed for production. Such pattern of changes in the genetic correlation is expected to be observed in real populations, a work that is currently undergoing.

14. Trade-off between fertility and production in French dairy cattle in the context of climate change

A. Vinet[1], S. Mattalia[2], R. Vallee[2], A. Barbat[1], C. Bertrand[3], B.C.D. Cuyabano[1] and D. Boichard[1]
1Université Paris-Saclay, INRAE, AgroParisTech, GABI, Jouy-en-Josas, 78350, France,
2Idele, UMT eBIS, Paris, 75012,France,
3INRAE, CTIG, Jouy-en-Josas, 78350

Climate change will induce harsher environments for European cattle production. One likely consequence is a growing antagonism between production and functional traits. To address this question, we investigated the evolution of trade-offs between fertility and production in French dairy cattle across a range from 15 to 75 of the temperature-humidity index (THI). Conception rate (CR) at first insemination (AI) and test-day protein yield (PY) from Holstein (HOL) and Montbeliarde (MON) cows were analyzed. Only first lactation performances, recorded between 2010 and 2020, were considered. CR and PY were modeled according to the average THI of the 8 days after the AI and of the 3 days before the test-day, respectively. In total, we analyzed 3,351,068 and 649,814 CR and 10,245,692 and 1,966,985 PY records from 3,368,605 and 656,164 HOL and MON cows, respectively. The evolution of trade-offs between CR and PY according to THI values was estimated with bivariate random regression models which included the fixed effects of herd x year, conventional vs sexed semen, weekday, age at AI, days in milk at AI for CR, and herd-test-day, age at calving, and days in milk for PY, and two random effects, the additive genetic effects of the sire for both traits and the permanent environment effect of the cow for PY. A sire model was chosen to allow very large analyses and thus ensure accurate estimates. The random effects were modeled with THI-dependent third-order Legendre polynomials. These models estimate genetic variances all along the THI trajectory as well as genetic correlations between traits across increasing THI conditions. The genetic correlations between PY and CR were moderately negative and evolved only slightly in the range of the observed THI. They increased with THI in HOL (around -0.1 for high THI), whereas they decreased in MON (around -0.15 for high THI). This study received funding from the European Union's Horizon 2020 research and innovation program under grant number 101000226 (Rumigen) and from APIS-GENE (CAICalor). The authors thank Meteo-France for the Safran database.

PAG 2024

Plant and Animal Genome Conference XXXI (San Diego, California, USA — January 12–17, 2024)

15. Data pertaining to creation and phenotyping of sheep with mutations within the PRLR gene*

S. G. Lillico

The Roslin Institute, University of Edinburgh, Scotland.

RUMIGEN is a European initiative involving 18 partners spread across 7 EU member states, plus Norway and the United Kingdom. Work packages within RUMIGEN span quantitative genetics, epigenetics, genome editing and societal (EU) acceptance of these technologies as contributors to ruminant breeding schemes. In order to better understand the genetic diversity of the European cattle population RUMIGEN aims to develop new tools including breed-specific genome assemblies. To better bridge our understanding of genotype / phenotype relationships, RUMIGEN aims to develop custom platforms to assess DNA methylation and sncRNA abundance. One of the main foci of RUMIGEN is to use these tools to develop breeding strategies that are able to adapt ruminants to a changing environment and the challenges this brings, while at the same time maintaining genetic diversity. This presentation will focus on one of these projects, in which we have applied genome editing to sheep with the aim of assessing the outcome of the well-known cattle slick phenotype in the context of thermal tolerance in another ruminant species.

* The SLICK genome-edited sheep project has also been presented at the following meetings:

1. Virtual talk, CAS Science Gala, University of Agriculture, Faisalabad — May 2024
2. In-person talk, Roslin workshop associated with the EU project "Green Innovation Strategies for Animal Health Management" — July 2024
3. In-person talk, Annual Meeting of the Gates-funded Centre for Tropical Livestock Genetics and Health, Nairobi — September 2024
4. In-person talk, Gujarat Biotechnology University, India (staff and ~100 MSc Animal Science students) — December 2024
5. In-person talk, Quality Meat Scotland Consortium, Edinburgh — May 2025
6. In-person talk, Transgenic Animal Research Conference, Tahoe, USA — August 2025
7. In-person talk, International Society for Biosafety Research, Ghent, Belgium — November 2025
8. Virtual talk, Agricultural students at Nottingham Trent University, UK — December 2025

ICQG 2024

7th International Conference on Quantitative Genetics (Vienna, Austria — July 25–26, 2024)

16. Altered Prior Mean of Allelic Effects: An Approach for Adequately Considering Gene-Edited Variants within Genomic Predictions

M. F. Schrauf, Y. C. J. Wientjes J. Vandenplas

Animal Breeding and Genomics, Wageningen University & Research, Wageningen, The Netherlands

Gene editing technologies offer a potential way to introduce novel genetic variation into livestock populations. However, the initial absence of phenotypic data from individuals carrying the edited alleles poses a significant challenge for estimating the genetic effects of these alleles in the population, which can then be lost due to

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

drift if the selection criteria is not adjusted. In this study, we explored Bayesian approaches aimed at incorporating prior knowledge on gene edited variants into genomic prediction models to facilitate their integration in selection schemes. Specifically, we focused on altering the prior means for the random allelic effects of the edited loci, a novel approach compared to previous efforts that primarily focused on altering only their prior variance. Concretely, we compared selection on conventional genomic estimated breeding values (GEBVs) with those obtained using altered prior variances, altered prior means, or a combination of both. The conventional prior considered for allele effects had mean zero and equal variance for all markers. In the altered priors, we only modified the prior effect of a single marker, in full linkage with the edited locus. For the altered prior mean we used the true effect of the edited allele, and for the altered prior variance we used the contribution of the edited locus to the genetic variance in the population in the first generation after gene edit. Our results show that selection on conventional GEBVs is at high risk of losing the gene edited alleles over successive generations, while selection on GEBVs with altered prior variance only improved the situation marginally. In contrast, selection on GEBVs with altered prior means effectively preserves the alleles and allows a fast increase of their frequency within the selected population. Notably, using these GEBVs based on altered prior means consistently leads to the edited allele becoming predominant in fewer than 10 generations. This study highlights the potential of altering the prior means of allelic effects to integrate prior knowledge on gene edited alleles, thereby ensuring the effective utilization of gene edited variants in genomic prediction schemes.

ISAG 2024

40th International Society for Animal Genetics Conference (Daejeon, South Korea, July 28–August 2, 2024)

17. Epigenetic Biomarkers Associated to Fetal Development during Maternal Lactation Using the RUMIGEN Methylation Array

A. López-Catalina, M. Gutiérrez-Rivas, O. González-Recio,
INIA-CSIC, Madrid 28040, Spain.

Incorporating epigenetic effects into breeding program remains challenging due to associated costs for detecting epigenetic marks, and to lack of clear applicability practices. Thanks to the development of the H2020 RUMIGEN methylation array, it is now possible to obtain information from DNA methylation marks throughout the genome of Bos Taurus cattle at an affordable cost. DNA methylation marks can modify expected SNP effects and influence phenotypic expression, potentially interacting with heritability estimates or the genetic merit of animals. One important environmental epigenetic effect is caused by a negative energy balance during embryo development when the gestated dairy cows are also lactating. To assess the impact of lactation on the methylation profiles of newborn dairy calves, we epi-genotyped 470 samples from calves born to lactating and nonlactating cows using the H2020 RUMIGEN methylation array, which includes 37,337 CpGs. The animals and their dams were genotyped using a 60K SNP chip. We integrated the epigenetic data with genomic information to identify CpGs associated with maternal lactation status using an epi-gwas approach. Further, the heritability of these CpGs across different genomic regions was estimated using gibbsf90+ software. We identified 155 CpGs significantly associated with maternal lactation status (Bonferroni-adjusted p-value <0.05). The heritability of the liability to methylation on these ranged from 0.30 to 0.95. These CpGs were linked to genes such as RCAN3, NOBOX, ALDH1A3, and WNT7A, involved in nervous system and embryo development, and pre-weaning survival. Additionally, we identified several HOX family genes, which are key regulators of embryonic development and body plan organization. These results pave the way to new strategies to incorporate epigenetic information in managements and breeding programs, facilitating our understanding of epigenetic effects in dairy cattle, and enabling more tailored and informed breeding decisions. This work was supported by European Union's Horizon 2020 Research and Innovation Program, 101000226.

EAAP 2024

75th Annual Meeting of the European Federation of Animal Science (Florence, Italy — August 26–30, 2024)

18. Sequence-based GWAS of heat tolerance traits in Holstein and Montbeliarde cattle

A. Vinet[1], S. Mattalia[2], R. Vallee[2], A. Barbat[1], C. Bertrand[3], C. Hoze[4], S. Taussat[4], M. P. Sanchez[1], M. Boussaha[1], B. Cuyabano[1], D. Boichard[1]
1 Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France
2 Idele, UMT eBis, 75012 Paris, France
3 INRAE, CTIG, 78350 Jouy-en-Josas, France
4 Eliance, UMT eBis, 78350 Jouy-en-Josas, France

Global warming will affect European cattle production systems, decreasing performance and welfare. To identify genomic regions involved in heat tolerance, breeding values (BV) along a temperature-humidity index (THI) gradient were estimated using a reaction-norm model for Holstein and Montbéliarde dairy cattle. Heat tolerance traits were defined as the slope of BV at THI 70 for test-day milk yield, conception rate at first insemination, and somatic cell score. Whole-genome sequence imputation (~13 M SNPs) from 4,564 Holstein and 1,737 Montbéliarde bulls was used to perform within-breed GWAS for the three traits. Few genomic regions were significantly associated with heat tolerance, and these differed between breeds and traits. Some regions corroborate previously reported findings, reinforcing their relevance. Selection for heat tolerance should therefore be considered.

19. Evolution of genetic parameters of production traits and conception rate in the Dutch Holstein population

J. Vandenplas[1], Y. Wientjes[1], P. Bijma[1], M. Schrauf[1], M. Van Pelt[2]
1 Wageningen University and Research, Animal Breeding and Genomics, P.O. Box 338, 6700AH Wageningen, Netherlands
2 CRV, P.O. Box 454, 6800AL Arnhem, Netherlands

Limited population size and selection can reduce the genetic variances of traits compared to base population genetic variances. These changes are a result of drift and the Bulmer effect, and can also impact the genetic correlations between traits. Here, we aim to evaluate the evolution of genetic parameters for production traits and conception rate (CR) in the Dutch Holstein population since 1990. The data included around 855,000 305 days-in-milk records for fat, protein and milk yields (MY), and around 676,000 records for CR at first artificial insemination for first-parity Holstein cows. Using a Gibbs sampling approach, changes in genetic (co)variances due to drift and selection were estimated by sex and per year of birth with bivariate models (one production trait and CR), in addition to base population (co)variances at five generations before 1990. For cows born in 1990, genetic variances were 22% lower for MY and 12% lower for CR than the base population genetic

variances. From 1990, genetic variances for cows increased on average by 0.3% (of the base population genetic variance) per year for MY, and decreased by 0.2% per year for CR. Corresponding genetic covariances were rather stable over time. Finally, genetic correlations between MY and CR changed from -0.45 in 1990 to -0.41 in 2016. Similar trends were observed for sires, but with steeper increases for genetic variances of production traits and correlations, especially from 2010. This project has received funding from the European Union's Horizon 2020 Programme for Research & Innovation under grant agreement n°101000226.

20. Linkage analysis based genomic relationships for optimum contribution selection

T. Meuwissen, X. Yu, P. Berg
Norwegian University of Life Sciences, BIOVIT, Oluf Thesens vei 6, 1433 Ås, Norway

In the pre-genomics era, inbreeding was based on pedigrees and estimated as the probability of Identity-by-Descent (IBD) at neutral loci unlinked to loci under selection. However, with genomic selection for genome-wide dense marker panels unlinked neutral loci hardly exist anymore. With genomic data, we can see what happens at the DNA level and a multitude of genomic measures of inbreeding have been developed. Some inbreeding measures are based on homozygosity, such as runs of homozygosity and excess of homozygosity, causing allele frequency changes towards 0.5 in optimum contribution selection (OCS). Other inbreeding measures are based on genetic drift, which favor the loss of rare alleles over a substantial increase of their frequency in OCS schemes. We argue that measures of inbreeding need to be based on IBD, such as the pedigree relationship matrix A, because diversity management based on IBD is neutral with respect to allele frequency changes. Genomic information can be included in the A matrix using linkage analysis (Gla). This linkage analysis approach was originally proposed by Fernando and Grossman for a single or few marker loci but is computationally extremely demanding for current dense SNP panels and large national data sets. Moreover, the Fernando and Grossman approach would calculate the Gla relationship matrix for all animals in the pedigree instead of only for genotyped animals such as the genomic relationship matrix (GRM) does. Depending on our needs, the latter may be computationally wasteful. Here we develop a linkage analysis approach to calculate the Gla relationship matrix using 50k SNP genotypes for 10,311 young genotyped cows in a pedigree of 393,129 Norwegian Red Cattle. We compared the estimated relationships to the GRM, the A matrix and the inbreeding coefficients to those obtained from Runs of Homozygosity.

21. Dairy cows' diversity of responses to heat stress: between resilience and adaptation

E. Vanbergue[1], L. Brun-Lafleur[1], P. Massabie[1], M. Prezelin[2], B. Vaillant[2], V. David[1]
1 Institut de l'élevage, 149 rue de Bercy, 75012 Paris, France
2 Ferme expérimentale Les Trinottières, La Futaie, 49140 Montreuil-sur-le-Loir, France

In a context of global warning, there is a need for implementing multi-criteria evaluation of dairy cows' physiological and behavioral adaptation to chronic and acute heat stress to 1/ provide references and facilitate decision-making process in herd management and 2/ specify the phenotyping strategy in genetic selection. To this end, thirty-six Holstein cows in late lactation/gestation, producing 22.5 (± 1.5) kg/d of milk, have been monitored during summer in Western France. Environmental conditions were described with Temperature Humidity Index (THI). Behavioral and physiological indicators recorded were ruminal temperature, time spent panting, resting, ruminating, and eating, milk production and composition, plasma non-esterified fatty acids, cortisol, and thyroxine, and colostrum immunoglobulins. K-means clustering algorithms identified 3 groups of cows accordingto their response to acute heat stress (average daily THI superior to 74 during 4 consecutive days). Group A was composed of cows with low milk losses (-0,5kg/d) and a slight increase of ruminal temperature (+0,24°C), group B was composed with cows with low milk losses (-0,5kg/d) and a higher increase of ruminal temperature (+0,36°C) and group C was composed of cows with higher milk losses (-2kg/d) and higher increase of ruminaltemperature (+0,36°C) and was characterized by higher milk fat and protein content during lactation. Group A, B, and C presented variable panting and/or eating/ruminating behavior adaptations. No difference in any of the other indicators measured was shown between groups. This project has received funding from the European Union’s Horizon 2020 Programme for Research & Innovation under grant agreement n°101000226.

22. Using Alternative Relationship Matrices for Genomic Prediction and Managing Genetic Diversity

X. Yu, P. Berg, T. Meuwissen
Norwegian University of Life Sciences, BIOVIT, Oluf Thesens vei 6, 1433 Ås, Norway

Modern breeding should make fast genetic progress to feed a growing world population, and also maintain the genetic diversity of the selective populations to meet future economic, environmental and societal demands. Optimum contribution selection (OCS) is considering both. It aims to maximize the breeding efforts, whilst constraining the rate of inbreeding. Before the genome era, only the numerator relationship matrix (NRM) from pedigree was available for inbreeding management. Genomic information has provided more measures of genetic relationships. The objectives of this study are to compare genetic improvements and diversity management in a long-term breeding program, using various relationship matrices as tools. Stochastic simulation software was developed to mimic a cattle population. Visible chip SNP, as well as invisible QTL and reference SNP were sampled to form a base population. NRM, the genomic relationship matrix (GRM), and IBD relationship matrix (IRM) were used to estimate breeding values, and/or as relationship constraints for optimum contribution selections. We conclude that using IRM for diversity management performed best in maintaining population diversity. IRM diversity management yielded least inbreeding, least loss of alleles across the genome and, most genetic and genic variance, and most potential for future breeding. The final allele reduced frequency of low minor allele frequency loci, making them less prone to fixation by genetic drift. The diversity maintained also provides more resources for later genetic progress. OCS schemes generate >25% more genetic progress than non-OCS schemes over 20 generations of selection in our simulations.

JOURNÉES 3R 2024

Rencontres Recherche Ruminants (Paris, France — December 4, 2024)

23. Etude de la tolérance à la chaleur des bovins par l’étude des interactions entre génotype et indice température-humidité sur la production et la santé de la mamelle en race Montbéliarde

A. Vinet[1], S. Mattalia[2,1], R. Vallée[2,1], C. Bertrand[3], P. Bertuzzi[4], B.C.D. Cuyabano[1], D. Boichard[1]
1Université Paris-Saclay, INRAE, AgroParisTech, GABI, Jouy-en-Josas, France
2Eliance (formerly UNCEIA), Paris, France
3Institut de l’Élevage (IDELE), Paris, France
4Organisme de Sélection Montbéliarde, Roulans, France

L'évolution du climat attendue dans les prochaines décennies dans les zones tempérées va modifier fortement les conditions d'élevage et entraîner une diminution du bien-être et des performances globales des animaux. Le premier objectif de cette étude est d'estimer la température critique marquant le début du stress thermique pour des vaches laitières de race Montbéliarde en production et de quantifier l'amplitude de la baisse des performances associée à ce stress thermique pour les caractères de production laitière et de comptage leucocytaire, transformé en score de cellules somatiques. Le second objectif est de quantifier l'ampleur des interactions génotype-environnement pour une gamme d'indices température-humidité (THI) afin de mettre en évidence une variabilité individuelle dans la tolérance à la chaleur. Pour cela, la base de données génétique française a été fusionnée avec la base de données Safran de Météo-France. Cette base de données fournit des mesures météorologiques quotidiennes estimées depuis 1958 sur une grille de 9892 carrés de 8x8km. Les performances à chaque contrôle laitier en première et deuxième lactations enregistrées entre 2017 et 2020 ont été associées à la moyenne des THI des 3 jours précédant le jour du contrôle. Au total, 301 078 performances de 39 469 vaches en première lactation et 292 648 performances de 37 593 vaches en deuxième lactation ont été analysées. Pour identifier la température critique les caractères ont été analysés à l'aide d'un modèle mixte comprenant des effets fixes (troupeau-année, stade de lactation, âge au vêlage, mois de vêlage, durée de gestation au moment du contrôle et THI) et des effets aléatoires (effet génétique additif et effet de l'environnement permanent des vaches). Cette analyse montre que le THI qui permet les meilleures performances de production et de santé de la mamelle varie selon le caractère mais qu'il est toujours inférieur à 55 (soit équivalent à une température inférieure ou égale à 13°C) : entre 50 et 55 pour la production laitière et entre 10 et 35 (soit équivalent à une température inférieure à 1°C) pour le taux protéique, le taux butyreux et le score de cellules somatiques. Pour un caractère donné, ce THI optimal est identique en première et deuxième lactation. La baisse de performance associée au stress thermique (THI=70, soit une température moyenne quotidienne de 23°C environ) va de 6% pour la production laitière en première et deuxième lactation à 12,5% pour le score de cellules somatiques en première lactation (7,2% en seconde lactation). Les moyennes des taux butyreux et protéiques diminuent quant à elles de 11% quel que soit le taux et le rang de lactation considéré. Les réponses individuelles au THI ont ensuite été estimées grâce à des modèles basés sur des régressions aléatoires. Ces modèles ont permis d'estimer les variances génétiques et les valeurs génétiques des animaux en fonction du THI ainsi que les corrélations génétiques entre un même caractère à différents THI. Pour un stade de lactation donné, les corrélations génétiques d'un même caractère à différents THI sont supérieures à 0,80 et la plupart sont supérieures à 0,90, ce qui suggère que les interactions génétiques-THI sont faibles pour les caractères de production et de santé de la mamelle. Néanmoins, les trajectoires des valeurs génétiques des taureaux montrent quelques reclassements en fonction du THI, indiquant que certains animaux seraient plus adaptés que d'autres à produire malgré des températures élevées. Ces travaux ont été réalisés dans le cadre des projets CAICalor, projet national financé par APIS-GENE et Rumigen, projet financé par le programme de recherche et d'innovation Horizon 2020 de l'Union Européenne sous le numéro de subvention 101000226 et adhérant à EuroFAANG (https://eurofaang.eu). Les auteurs remercient Météo-France pour la base de données Safran.

PAG 2025

Plant and Animal Genome Conference XXXI (San Diego, California, USA — January 10–15, 2025)

24. Sequence-level GWAS of sperm DNA methylation in Holstein bulls

Corentin Fouéré[1,2], Valentin Costes[2,3], Chris Hozé[1,2], Florian Besnard[1,2], Gabriel Costa Monteiro Moreira[3], Aurélie Chaulot-Talmon[3], Francesca Ali[3], Amrita Raja Ravi Shankar[3], Sebastien Fritz[1,2], Mekki Boussaha[2], Didier Boichard[2], Hélène Kiefer[3], Marie-Pierre Sanchez[2]
1 Eliance, 75012 Paris, France,
2 Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78352 Jouy-en-Josas, France,
3 Université Paris-Saclay, INRAE, Ecole Nationale Vétérinaire d’Alfort, BREED, 78352 Jouy-en-Josas, France

In mammals, DNA methylation (DNAm) corresponds to the addition of a methyl group to the 5-carbon of cytosine in CpG dinucleotides. Two waves of epigenetic reprogramming (erasure and de novo methylation) occur (1) during gametogenesis and (2) after fertilization. In humans, genomic variants affecting DNA methylation (meQTLs) have mainly been studied in white blood cells, i.e., after the second wave of reprogramming. Here, we focus on a farm animal species, cattle, and specifically on male gametes that have undergone only the first wave of epigenetic reprogramming. We studied the genetic architecture of sperm DNA methylation in 405 Holstein bulls using Reduced Representation Bisulfite Sequencing (RRBS). Among the 3 million CpG sites revealed by RRBS, we selected those with i) inter-individual variation (methylation rate standard deviation>0.05), ii) no overlap with known genomic variants, and iii) at least 10 reads in 80% of samples. In the end, 166,985 CpGs were analysed. The methylation rate was corrected for age at semen collection. Heritability estimates ranged from 0 to 1 and averaged 0.26 across all selected CpGs, with 76% of estimates above 0.1. Both intergenic regions and low density CpG regions were enriched for CpGs with h2>0.1. A genome-wide association study (GWAS) was performed at the whole-genome sequence level using GCTA software with 12.9 million SNPs imputed from 50K SNP arrays. Cis-meQTLs were defined within 1Mb, while trans-meQTLs were defined as >20Mb intrachromosomal or interchromosomal. The proportion of CpGs influenced by significant cis- and trans-meQTLs was 33% and 5%, respectively. Only 1% of the analyzed CpGs were reported to share both cis- and trans-meQTL. The average absolute distance between a cis-CpG and its top cis-SNP meQTL was 260.8kb. Top SNPs (both cis- and trans-) were enriched in CpG-dense regions, promoters (considered by TSS200 and TSS1500 regions), and ATAC-seq peaks, and they were depleted in low density CpG regions. The enrichment of CpGs with cis- and trans-meQTLs was moderate overall. We identified 8 trans-meQTL hotspots, defined as genomic regions associated with more than 30 trans-CpGs, each influencing between 31 and 553 trans-CpGs. Several of these trans-meQTL hotspots overlapped with genes involved in epigenetic processes, including histone modifiers and genes interacting with DNA methyltransferase enzymes. This work provides new insights into the genetic influence on DNA methylation in cattle, especially relevant with the emergence of the cattle DNAm array as a new tool for cattle breeding.

CF is recipient of a CIFRE PhD grant from ANRT and APIS-GENE. This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 101000226. This work was part of the POLYPHEME project funded by ANR (ANR-21-CE20-0021) and APIS-GENE.

AIDA 2025

XXI Jornadas Sobre Producción Animal (Zaragoza, Spain — May 13–14, 2025)

25. Ordeñando el epigenoma del vacuno lechero: una red de co-metilación para la detección de biomarcadores asociados a estrés metabólico en terneras.

A. López-Catalina[1,2], A. Reverter-Gomez[3], L.R. Porto-Neto[3], P.A. Alexandre[3], L.T. Nguyen[4], O. González-Recio[1]
1INIA-CSIC, Madrid, Spain
2Universidad Politécnica de Madrid (UPM), Madrid, Spain
3CSIRO Agriculture & Food, Queensland Bioscience Precinct, Brisbane, Australia
4University of Queensland, Brisbane, Australia

Los programas de mejora genética permiten mejorar tanto la salud como la productividad del ganado. La genómica ha permitido desarrollar estos modelos y aumentar su precisión. Actualmente, la epigenética es un área de investigación en expansión, con una gran proyección de futuro (Ospelt. 2022). En el ámbito de la mejora genética animal, el estudio de las marcas epigenéticas puede ayudarnos a entender cómo el fenotipo de un animal puede modificarse sin que haya un cambio en su secuencia nucleotídica. Estudiar cómo la dieta, estresores externos e influencias del ambiente alteran el epigenoma del animal nos da una nueva herramienta para mejorar nuestras predicciones genéticas y aquellas prácticas ganaderas que influyan negativamente en el epigenoma (González-Recio. 2012).

EAAP 2025

76th Annual Meeting of the European Federation of Animal Science (Innsbruck, Austria — August 25–29, 2025)

26. Application of Variation Graphs for Genotyping Structural Variants in 14 French Cattle Breeds

M. Mughitz Naji[1], C. Klopp[2], C. Eché[1], A. Di Franco[1], C. Birbes[1], C. Marcuzzo[1], A. Suin, C. Iampietro[1], C. Kuchly[1], C. Vincent[1]., S. Fritz[1], C. Grohs[1], T. Faraut[234], C. Gaspin[234], D. Milan[23], C. Derrien[1]., D. Boichard[1], M. Sanchez[1], M. Boussaha[1]
1INRAE, GABI, Université Paris-Saclay, Jouy-en-Josas, France
2INRAE, GenPhySE, Castanet-Tolosan, France
3Université de Toulouse, INRAE, ENVT, GenPhySE, Toulouse, France
4INRAE, Sigenae & Bioinfo platforms, Toulouse, France

Structural variants (SVs) are genomic variations larger than 50 bp. Long-read (LR) sequencing is preferred over short-read (SR) sequencing to improve SV detection accuracy. Here, we analyzed SVs, focusing on large deletions (DEL) and insertions (INS), using whole-genome sequencing data from 176 LR and 571 SR samples representing 14 French cattle breeds. One sample was sequenced with three LR technologies (PacBio HiFi, Oxford ONT, and PacBio CLR). First, we assessed the performance of three SV detection tools (CuteSV, Pbsv, and Sniffles2) on HiFi data. The tools identified a consensus of 10,000 DEL and 8,866 INS. A further evaluation of SV detection across the three LR technologies, comparing SVs detected from CLR or ONT against HiFi data, revealed that Pbsv showed the highest consistency, with F1 score of 0.91 for DEL and 0.85 for INS. We then compared tool performance by leveraging 154 samples with both LR and SR data. We compared three SV callers (Delly, Lumpy, and Manta) and four SV genotypers (Graphtyper, Svtyper, Paragraph, and VG toolkit) with SR data. Benchmarking these tools against LR-based SVs detected with Pbsv revealed that VG toolkit performed best, achieving an average F1 score of 0.932 for DEL and 0.952 for INS. To explore SV genotyping at the population level, we divided the 154 samples into 6 validation and 148 reference samples. Variation graphs

were incrementally constructed using SVs detected from LR by Pbsv, incorporating data from 1, 2, 3, or all 14 breeds in the reference set. SVs from the validation samples' SR data were then genotyped based on these graphs and compared to their respective LR truth sets. Including breed-specific samples into the variation graph enabled the genotyping of breed-specific SVs and improved recall rates. Finally, we optimized parameters to construct a final variation graph representing 25,191 DEL and 30,118 INS segregating within the 14 breeds. This graph was applied to genotype SVs in 571 SR individuals, enabling population-level profiling of structural genomic variants. This work was funded under CASCAD project by CARNOT France Future Élevage (F2E). Part of the short-read sequence data used in this study were produced by the H2020 RUMIGEN project.

27. Applying the RUMIGEN Methylation Array to Identify Epigenetic Markers Associated to Fetal Development during Maternal Lactation in Cattle

A. López-Catalina, M. Gutiérrez-Rivas, Ó. González-Recio
INIA-CSIC, Crta. de la Coruña km 7.5, 28040 Madrid, Spain

This work was supported by European Union's Horizon 2020 Research and Innovation Program, 101000226. Genotyping arrays have become an essential tool for genetic evaluations. Positions of interest are selected to estimate individuals' genetic merit and identify the best parents to enhance the rate of genetic gain. Alongside improvements in phenotype recording and computation, the genetic rate of gain has doubled over the past decade. However, genomic information and genetic biomarkers such as SNPs explain only part of the phenotypic variance and the heritability of economically relevant traits in livestock. In this context, epigenetic biomarkers, particularly DNA methylation marks, are being explored for their potential role in livestock breeding programmes. These marks provide insights into the heritability of traits and the environmental influences on gene expression, ultimately improving selection strategies. To assess the impact of methylation marks, we epi-genotyped 470 samples from new-born calves from lactating or non-lactating cows using the H2020 RUMIGEN methylation array designed within the RUMIGEN consortium and manufactured by Illumina Inc., which includes approximately 44,000 CpGs, of which 37,337 passed quality control filters. We combined this epigenetic data with genomic information from a 60k SNP chip to identify CpGs associated with maternal lactation status and estimate the genetic control of DNA methylation at different genomic regions using the Gibbsf90+ software. We identified 155 CpGs significantly associated with maternal lactation status (Bonferroni-adjusted p-value < 0.05). These CpGs were linked to genes such as RCAN3 and NOBOX, which are involved in nervous system development, and pre-weaning survival; ALDH1A3, associated with growth traits; WNT7A, which plays a role in embryo development; and several HOX family genes, which are key regulators of embryonic development and body plan organisation. The heritability of DNA methylation at these CpGs ranged from 0.30 to 0.95. Given the nature of DNA methylation marks, this tool could enhance our knowledge of the epigenetic importance on dairy cattle production and may serve as a tool to personalized breeding strategies.

28. Longitudinal Analysis of DNA Methylation in Dairy Calves

A. López-Catalina, M. Gutierrez-Rivas, Ó. González-Recio
INIA-CSIC, Ctra. de La Coruña, km 7, 5, 28040 Madrid, Spain

DNA methylation plays a crucial role in gene regulation and development, yet its temporal dynamics in livestock remain largely unexplored. This study investigates genome-wide DNA methylation changes in dairy calves over five key developmental stages: at birth, six months, one year , 18 months, and after first parturition. Blood samples were collected at each time point and analysed using the RUMIGEN methylation array. Missing valueswere imputed using the MICE algorithm. Principal Component Analysis was performed to explore variance in methylation patterns, revealing clustering trends over time. A linear mixed-effects model was applied to assess the impact of age on CpG methylation levels, with Bonferroni correction for multiple testing. Of the 8,572 CpG sites analysed, 35.04% exhibited increased methylation, while 64.96% became hypomethylated. Gene mapping identified several genes associated with differentially methylated CpG sites, providing functional insights into key biological processes. In cellular regulation, genes such as KCNN3, STK17A, and PARP1 influence excitability, apoptosis, and DNA repair, impacting udder health, while PPAT is crucial for cell growth during lactation. In mammary gland development, NFIA, BCL2L11, and TUSC3 play key roles, with CCDC102A and FAM3Dshowing potential involvement in this process. In immune response, GATA3 and IRF2BPL contribute to immune function and may influence mastitis susceptibility, while STK17A and PARP1 also play immunological roles. Lastly, in milk production, MYO9B may regulate milk secretion, and the microRNA bta-mir-2284c influences gene expression related to milk synthesis. These findings provide new insights into the epigenetic regulation of growth, immune function, and lactation in dairy cattle, highlighting potential biomarkers for future breeding and management strategies.

29. Further characterization of the structural variant upstream of the KIT gene underlying head depigmentation across a diverse panel of cattle breeds

V. Sorin[1], M. P. Sanchez[1], L. Drouilhet[2], G. Tosser-Klopp[2], M. M. Naji[1], D. Boichard[1], H. Pausch[3], M. Boussaha[1], A. Leonard[3]
1 INRAE, Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France
2 INRAE, Université de Toulouse, INRAE, GenPhySE, 31326 Castanet-Tolosan, France,
3 ETH, Animal Genomics, ETH, 8092 Zurich, Switzerland

A previous study based on a taurine pangenome uncovered a structural variant upstream of the KIT gene associated with depigmentation in white-headed cattle. In this study, we aimed to deepen our understanding of this region by analysing a collection of 80 genome assemblies from 20 different cattle breeds, including several not previously investigated. Using a combination of complementary approaches, we constructed a 2 Mb pangenome graph encompassing the KIT gene and the depigmentation-associated region to investigate the spectrum of structural variations within this region. Our analysis revealed a more fragmented structure than previously reported, mainly due to the inclusion of a larger number of assemblies (80 vs. 24) with varying sequencing technologies' quality (CLR and HiFi). Despite this increased complexity, the identified structural variant's alleles closely mirrored those reported in earlier study and depicting clear segregation of the population into two distinct groups. The first group includes breeds characterized by a white-headed phenotype and the second group contains all non-white headed breeds. We further validated this structural variant by aligning short-reads to the pangenome graph and genotyping several hundred animals for this structural variant. Our results provide new insights into the genetic mechanisms underlying head depigmentation in cattle and demonstrate the power of pangenome-based analyses for detecting trait-associated structural variations. This research was supported by funding from the Agreenium programme, with contributions from the Biosphera Graduate School, the ABIES doctoral school, the INRAE Animal Genetics division, and the INRAE GBoS team from the GABI department. Part of the short-read sequence data used in this study were produced by the H2020 RUMIGEN project.

30. Detecting genetic load: A GWAS model accounting for identity by descent

G. Restoux[1], T. Tribout1, R. Bonifazi[2], T. H. Meuwissen[3], J. J. Windig[2], A. Barbat[1], S. Minery[4], J. Vandenplas[2], P. Croiseau[1]
1 INRAE, Uni. Paris-Saclay, AgroParisTech, GABI, , Jouy-en-Josas, France
2 WUR, Animal Breeding and Genomics, , Wageningen, Netherlands
3 NMBU, Department of Animal and Aquaculture Sciences, , Ås, Norway
4IDELE, Paris, France

Inbreeding depression is due to the expression of genetic load through inbreeding and is one of the main threats in livestock as it can lead to reduced fitness and performances. While many strategies to cope with inbreeding depression are based on the avoidance of inbreeding, only few strategies focus on the genetic load itself since it is difficult to estimate it. The primary basis for genetic load is the segregation, at low frequencies, of deleterious and recessive mutations in the populations, making them difficult to detect in the genome. Thus, the goal of this study was to design a model to detect regions of the genome involved in the genetic load. We set a GWAS model considering additive and dominance effects as well as pedigree-based inbreeding to account for whole-genome inbreeding depression effect. Moreover, runs of homozygosity (RoH) located along the genome were added as a covariable in the GWAS model. While very few studies included RoH into GWAS models, even less have discussed the interpretation of considering this covariable in models and what is the added value of doing so. Using simulation, we aimed to better decipher which statistical contrasts should be computed to detect genetic load and their meaning. We assumed that RoH gave access to HBD (homozygosity by descent), thus helping to get rid of weak linkage disequilibrium between causative and targeted SNP, and to cope with low frequency mutations. We applied this model to production (milk yield, fat and protein content) and health (Somatic Cell Score) traits in heifers from 5 local breeds located in France, the Netherlands and Norway. We identified several genomic regions putatively involved in genetic load in at least one breed. Finally, we discuss these regions relative to those detected using other approaches.

31. Considering variances and correlations of offspring breeding values shows beneficial effects on genetic gain and antagonistic correlations

M.P.L. Calus, Y.C.J. Wientjes, P. Bijma, T. Niehoff
Wageningen University & Research, Animal Breeding and Genomics, P.O. Box 338, 6700 AH Wageningen, Netherlands

Determining gametic Mendelian Sampling variances of potential parents allows to predict the variance of the breeding values of their future offspring. Several studies have shown that considering these predicted variances in selection decisions can increase realized genetic gain. Almost all of those studies considered only one single trait. The objectives of our study were to investigate if in the case of multi-trait selection including variances and correlations of breeding values of future offspring in the selection decision can increase genetic gain, and what the impact of this approach is on the genetic correlation between traits. To achieve this objective, we used the gametic Mendelian Sampling variances and correlations, which explain observed genetic variances within traits and correlations between traits among the offspring of individual animals. Given that individual differences in gametic Mendelian Sampling correlations have been observed, we hypothesized that using those correlations in selection decisions may affect future antagonistic genetic correlations. We considered one generation of selection, by simulating one parental and one offspring generation. We directly simulated bivariate distributions of breeding values for two traits, and then performed selection of the parents based on: breeding values (following the current practice), or selection based on the so-called Index5, which in addition to the breeding values also included the individual gametic Mendelian Sampling variances and/or the individual gametic Mendelian Sampling correlations. Including Mendelian Sampling variances and correlations in selection resulted in up to 2% more genetic gain compared to ordinary selection on breeding values, while it was able to change the unfavourable genetic correlation in the next generation from -0.30 to –0.25. When the correlation between traits was ignored, the correlation was unaffected and less genetic gain was achieved. These results provide a first indication that use of individual Mendelian Sampling correlations in selection decisions can be beneficial for genetic gain and may help to mitigate the effect of antagonistic genetic correlations.

32. Constructing haplotype-separated assemblies for Norwegian Red cattle

T. J. Hettasch[1], A. B. Gjuvsland[2], M. P. Kent[1], H. Grove1, D. I. Våge[1]
1 Norwegian University of Life Sciences (NMBU), Ås, Norway
2 Geno SA, Hamar, Norway

High-quality reference genomes improve our understanding of genome structure and enables the identification of genetic variants associated with traits of interest. We have constructed the first set of haplotype-separated assemblies for the Norwegian Red (NR) cattle breed using long-read sequencing and the trio-binning method. This method uses parental sequences to label offspring reads from the maternal and paternal haplotypes, creating fully haplotype-resolved assemblies, which give a more accurate representation of complex SVs and haplotype-specific alleles than assemblies representing both haplotypes in a collapsed manner. Three trios were selected from the NR population. Short-reads (Illumina) were collected from parents and long-reads (PacBio, ONT) were collected from offspring. Computational simulation of parental sequences from genome-wide panels of SNP genotypes was also tested as an alternative way of generating parental data. Offspring long-reads were assembled with Hifiasm using both Illumina reads and simulated sequences as parental input. All six haplotypes, assembled with Illumina reads as parental input, show high continuity (N50 > 73 Mbp) and completeness (BUSCO > 95%). Mapping of bovine satellites suggests that more of the cattle centromere has been assembled than in previous references. Counting the number of haplotype-specific k-mers in the final assemblies shows that only a small fraction [0.46 – 2.52%] are found in the wrong haplotype, suggesting successful separation of offspring sequences into haplotypes. Using simulated sequences as parental input produces assemblies with similar quality and haplotype-resolution as using Illumina short-reads. This enables global haplotype assembly without additional sequencing of parental genomes in livestock species where genotyping is done routinely. A genome-graph will be constructed from the collection of haplotype assemblies, providing a valuable resource for studying genetic diversity and structural variation across the NR population.

33. Harnessing RUMIGEN EpiChip data to predict dairy cow performance

M.P. Sanchez [1], F. Besnard [2],[3] , V. Costes [2], C. Fouéré [1],[2], Clotilde Patry [4], C. Le Danvic [2], G. Costa Monteiro Moreira [5], Alexandre Asset [5], Mekki Boussaha [1], Hélène Jammes [5], Sébastien Fritz [2],[1], Hélène Kiefer [5], Didier Boichard[1]

1 GABI - Génétique Animale et Biologie Intégrative
2 Eliance
3 INRAE - Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement
4 Valogene
5 BREED - Biologie de la Reproduction, Environnement, Epigénétique & Développement

The DNA methylation BeadChip (EpiChip), developed in the H2020 RUMIGEN project, offers a powerful ap¬proach to understanding the epigenetic factors influencing dairy cow performance. This study used the RUMI¬GEN EpiChip on blood samples from 3,087 Holstein cows to evaluate the extent to which epigenetic biomarkers may help predict cow phenotypes. Key performance indicators such as milk production, fertility, and udder health were analyzed using a mixed model that included the random effects of genetic and/or epigenetic values. Genetic covariance structure was computed from SNP genotypes obtained from the EuroGMD chip, while epigenome covariance structure was derived from methylation similarity between individuals estimated from EpiChip data, after adjusting data from leucocyte composition. Three distinct models were applied to estimate genetic and/ or epigenetic variances, all including environmental effects along with i) a genetic component, ii) an epigenetic component, or iii) both genetic and epigenetic components. The population of cows was divided into a training and validation set to estimate prediction accuracy across the three models. Results presented in detail include the proportion of epigenetic variance relative to the total variance for all traits considered in this study. Specifically, we compare the performance of the three models in terms of accuracy and bias in predicting cow performance. These findings pave the way for EpiChip data to improve predictive capacity beyond conventional genetic and environmental factors. This provides a more comprehensive framework for animal performance management, supporting more informed breeding and management decisions to advance efficiency and sustainability in the dairy industry. This project has received funding from the European Union's Horizon 2020 Programme for Re-search & Innovation under grant agreement n°101000226 (RUMIGEN) and from APIS-GENE (PolyPheme).

PAG 2026

Plant & Animal Genome Conference XXXIV (San Diego, Usa — January 10–14, 2026)

34. A Cattle Pangenome Approach Reveals Novel Non-Reference Unique Insertions And Their Impact On Economic Traits.

V. Sorin[1], F. Besnard[1], C. Grohs[1], C. Escouflaire[1], S. Fritz[1], L. Drouilhet[1], G. Tosser-Klopp[2], D. Boichard[1], C. Klopp[2], M. Sanchez[1], M. Boussaha[1]
1INRAE, GABI, Université Paris-Saclay, Jouy-en-Josas, France
2Université de Toulouse, INRAE, ENVT, GenPhySE, Toulouse, France

The current cattle reference genome assembly, which is derived from a single Hereford cow, does not encompass the entire range of genetic variations present within the species. Several megabases of sequences are still missing and these can contain genomic regions potentially associated with key phenotypes. Structural variations (SVs ≥ 50 nucleotides) are difficult to detect using only standard approaches of either short or long-read sequence mapping to the current bovine genome assembly. To overcome this, recent advances in sequencing technologies to produce high-quality long-read sequences combined with the development of pipelines and tools for sequence analysis pave the way for the use of pangenomics in cattle. Those approaches provide thorough comprehension of the extent of genetic diversity among cattle breeds. In this study, we constructed a cattle pangenome graph using 16 high-quality haplotype-resolved genome assemblies originating from 9 French cattle breeds, encompassing both major dairy and beef populations as well as regional breeds, and including the Yak (Bos grunniens) as a close species. A combination of complementary methods led to the detection and characterization of 101,219 structural variations. Of these, 33,634 structural variants were classified as non-reference unique insertions (NRUIs), adding several megabases of new genomic sequences that are not present in the current Hereford reference genome. These NRUIs were subsequently genotyped (presence/absence) in a large multi-breed genotyped cattle population. With these NRUI data, all samples were clustered in distinct groups corresponding to the breed of origin. This result is of particular interest as the NRUIs were able to provide a clear description of the breed structure. Functional characterisation of the NRUIs identified two NRUIs located in the intronic regions of ARMH3 and EPHA5. They were found to be specific to the Normande breed and were significantly associated with milk production and morphological traits, respectively. Our findings underscore the importance of pangenome-based approaches to uncover the full spectrum of genomic variations and establish a potential link with cattle key phenotypes. Part of the long and short-read sequence data used in this study were produced by the H2020 RUMIGEN project.

Scientific papers

2023

1. Oxford nanopore sequencing as an alternative to reduced representation bisulphite sequencing for the identification of CpGs of interest in livestock populations.

López-Catalina, A., Costes, V., Peiró-Pastor, R., Kiefer, H., González-Recio, O. Oxford nanopore sequencing as an alternative to reduced representation bisulphite sequencing for the identification of CpGs of interest in livestock populations. Livestock Science 279, 105377 (2024). https://doi.org/10.1016/j.livsci.2023.105377

Epigenetic marks could potentially explain a portion of the phenotypic variance related to traits of interest in animal breeding models. DNA methylation is the most studied epigenetic mark, involving the addition of a methyl group to the fifth carbon of a cytosine, which transforms it into 5-methylcytosine. This mark is commonly associated with inhibiting gene expression without altering the nucleotide sequence, particularly when located within promoter regions. While bisulphite sequencing is the gold standard for detecting methylation marks, new techniques have emerged to address bisulphite's limitations. Some of these limitations include the confounding effects of bisulphite treatment on DNA methylation and C to T polymorphisms, as well as the inability to distinguish between DNA methylation and hydroxylation. In this context, we propose employing Nanopore sequencing to identify methylation marks. We conducted sequencing on six bull semen samples using both Nanopore and reduced representation bisulphite sequencing. CpGs were filtered based on two coverage thresholds (4x and 7x). Our results indicated that Nanopore methylation data exhibited a correlation exceeding 0.95 with bisulphite sequencing data. The correlation decreased to 0.67 when analysing only CpGs with methylation frequencies ranging from 0.1 to 0.9, due to the lack of enough coverage for ONT. We also examined differentially methylated cytosines identified by each sequencing method. The overlap between the two sets of DMCs and the associated genes was limited, as RRBS library preparation predominantly amplifies gene promoter regions, while ONT covers more intergenic regions. Interestingly, both methods highlighted differentially methylated genes and positions linked to quantitative trait loci of significant economic interest in animal breeding, suggesting promising applications for ONT in the field of agri-genomics.

2. Estimation of genotype by temperature-humidity index interactions on milk production and udder health traits in Montbeliarde cows

Vinet, A., Mattalia, S., Vallée, R. et al. Estimation of genotype by temperature-humidity index interactions on milk production and udder health traits in Montbeliarde cows. Genet Sel Evol 55, 4 (2023). https://doi.org/10.1186/s12711-023-00779-1

Background. Heat stress negatively influences cattle welfare, health and productivity. To cope with the forecasted increases in temperature and heat waves frequency, identifying high-producing animals that are tolerant to heat is of capital importance to maintain milk production.

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

This study, based on the joint analysis of on-farm performance and weather data, had two objectives: (1) to determine the response in production performances (milk, fat and protein yields, fat and protein contents) and udder health (somatic cell score) to temperature-humidity index (THI) variations in Montbeliarde cows, and (2) to estimate the interactions between genotype and THI, to enable the identification of the most adapted animals for facing the expected increases in temperature. Results. Test-day records from first and second lactations from 2016 to 2020 were associated with the average THI during the three days before the test-day record. In total, 446,717 test-day records from 55,650 cows in first lactation and 457,516 test-day records from 58,229 cows in second lactation were analysed. The optimal THI was below 55 (i.e.~12–13 °C) for all traits. Individual responses to THI were estimated by random regression models, which also included individual responses to days in milk. Regardless of the stage of lactation, genetic correlations along the THI gradient were above 0.80, which suggests that genotype-by-THI interactions were weak for production and udder health traits. Nevertheless, a variability in the individual slope of decay could be highlighted at high THI. The genetic correlation between production level at moderate THI and the slope at high THI was negative, while for somatic cell score, it was positive, indicating that heat stress amplifies the susceptibility to mastitis. Conclusions. The optimal THI for French Montbeliarde cows is below 55 for production and udder health traits. Genetic-by-THI interactions are weak in French Montbeliarde cows for production and udder health traits, but not all animals react in the same way to high temperatures. Even if there is little room for improvement, using a heat tolerance index in cattle selection would be relevant to anticipate the expected increases in temperature. Further investigations are needed to interpret this variability on production traits. However, the current selection for mastitis resistance seems appropriate to adapt cattle to rising temperatures.

2024

3. Stress-induced epigenetic effects driven by maternal lactation in dairy cattle: a comethylation network approach

López-Catalina, A., Reverter, A., Alexandre, P.A., Nguyen, L.T., González-Recio, O. Stress-induced epigenetic effects driven by maternal lactation in dairy cattle: a comethylation network approach. Epigenetics 19, 1 (2024). https://doi.org/10.1080/15592294.2024.2381856

Epigenetic marks do not follow the Mendelian laws of inheritance. The environment can alter the epigenotype of an individual when exposed to different external stressors. In lactating cows, the first stages of gestation overlap with the lactation peak, creating a negative energy balance that is difficult to overcome with diet. This negative energy balance could affect early embryo development that must compete with the mammary tissue for nutrients. We hypothesize that the methylation profiles of calves born to nonlactating heifers are different from those of calves born to lactating cows. We found 50,277 differentially methylated cytosines and 2,281 differentially methylated regions between these two groups of animals. A comethylation network was constructed to study the correlation between the phenotypes of the mothers and the epigenome of the calves, revealing 265 regions associated with the phenotypes. Our study revealed the presence of DMCs and DMRs in calves gestated by heifers and lactating cows, which were linked to the dam's lactation and the calves' ICAP and milk EBV. Gene-specific analysis highlighted associations with vasculature and organ morphogenesis and cell communication and signalling. These finding support the hypothesis that calves gestated by nonlactating mothers have a different methylation profile than those gestated by lactating cows.

4. Effect of temperature-humidity index on the evolution of trade-offs between fertility and production in dairy cattle

Vinet, A., Mattalia, S., Vallée, R. et al. Effect of temperature-humidity index on the evolution of trade-offs between fertility and production in dairy cattle. Genet Sel Evol 56, 23 (2024). https://doi.org/10.1186/s12711-024-00889-4

Background. In the current context of climate change, livestock production faces many challenges to improve the sustainability of systems. Dairy farming, in particular, must find ways to select animals that will be able to achieve sufficient overall production while maintaining their reproductive ability in environments with increasing temperatures. With future forecasted climate conditions in mind, this study used data from Holstein and Montbeliarde dairy cattle to: (1) estimate the genetic-by-temperature-humidity index (THI) interactions for female fertility, and (2) evaluate the production-fertility trade-off with increasing values of THI. Results. Two-trait random regression models were fitted for conception rate (fertility) and test-day protein yield (production). For fertility, genetic correlations between different THI values were generally above 0.75, suggesting weak genotype-by-THI interactions for conception rate in both breeds. However, the genetic correlations between the conception rate breeding values at the current average THI (THI = 50, corresponding to a 24-h average temperature of 8 °C at 50% relative humidity) and their slopes (i.e., potential reranking) for heat stress scenarios (THI > 70), were different for each breed. For Montbeliarde, this correlation tended to be positive (i.e., overall the best reproducers are less affected by heat stress), whereas for Holstein it was approximately zero. Finally, our results indicated a weak antagonism between production and fertility, although for Montbeliarde this antagonism intensified with increasing THI. Conclusions. Within the range of weather conditions studied, increasing temperatures are not expected to exacerbate the fertility-production trade-off. However, our results indicated that the animals with the best breeding values for production today will be the most affected by temperature increases, both in terms of fertility and production. Nonetheless, these animals should remain among the most productive ones during heat waves. For Montbeliarde, the current selection program for fertility seems to be adequate for ensuring the adaptation of fertility traits to temperature increases, without adverse effects on production. Such a conclusion cannot be drawn for Holstein. In the future, the incorporation of a heat tolerance index into dairy cattle breeding programs would be valuable to promote the selection of animals adapted to future climate conditions.

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

5. Long-lasting effects of in utero heat stress on subsequent performances of heifers and primiparous cows

Vinet, A., et al. Long-lasting effects of in utero heat stress on subsequent performances of heifers and primiparous cows. Journal of Dairy Science 107, 11065–11077 (2024). https://doi.org/10.3168/jds.2024-25168

The performance of an adult dairy cow may be influenced by heat stress that occurs during her gestation. The present study investigated potential effects of temperature-humidity index (THI) experienced by a cow during pregnancy on the gestated daughter's performance on her first lactation, for the French Holstein and Montbéliarde dairy cattle populations. We analyzed 14 traits, all measured on genotyped cows: 305-d milk, fat, and protein yields; 305-d SCS; clinical mastitis (both occurrence and number of events); body conformation traits; and heifer and cow conception rate. To study the effect of heat stress, we considered the THI experienced by the gestating cow, averaged for each month of her pregnancy and then categorized in 7 classes (≤40, [40,45], [45,50], [50,55], [55,60], [60,65], and >65). These average THI classes were then fitted as categorical covariates in the regression models used for this study, which included other fixed effects, and the GEBV as a covariate, both specific to each trait, the latter being previously obtained from the official French evaluations. The THI effect was therefore estimated as the deviation between the observed and predicted performances. In general, the estimated heat stress effects were small, presenting limited practical impact on the studied traits, and particularly for fertility and udder health, the estimated heat stress effects were not statistically significant. For the production traits (i.e., milk, fat, and protein yields), the estimated effect associated with high THI experienced at the beginning of the gestation was negative, and slightly positive when associated with high THI experienced by the dam at the end of her pregnancy. Finally, our results suggest that under the current French climate conditions, heat stress experienced by cows during any stage of their pregnancy has limited impact on the future performance of their gestated daughters; however, we cannot exclude that a significant in utero heat stress effect may be present in climate conditions warmer than the French.

2025

6. Comprehensive detection of structural variations in long and short reads dataset of French cattle

Naji, M.M., Klopp, C., Eché, C. et al. Comprehensive detection of structural variations in long and short reads dataset of French cattle. Sci Rep 15, 40758 (2025). https://doi.org/10.1038/s41598-025-24477-8

Structural variants (SVs) correspond to different types of genomic variants larger than 50 bp. Many findings suggest the use of long-read (LR) rather than short-read (SR) sequencing to improve the accuracy of SVs detection. Here, we present the results of an in-depth analysis for detection of SVs, mainly large insertions and deletions, in 14 French bovine breeds, based on whole-genome sequence (WGS) data comprising 176 LR and 571 SR samples, with 154 individuals having both LR and SR data available. We first investigated possible biases on the performances of well-known SVs detection tools, namely CUTESV, PBSV, and SNIFFLES, using LR from different technologies, including PacBio HiFi, Oxford ONT, and PacBio CLR. We subsequently highlighted the abilities of tools for detecting SVs (DELLY, LUMPY, and MANTA) and for genotyping known SVs (GRAPHTYPER, SVTYPER, PARAGRAPH, and VG toolkit) using SR data. We then show how the incremental composition of samples in the reference panel affected the SVs genotyping for six validation individuals sequenced in SR. We then searched for the optimal parameters and created the final SVs reference panel consisting of 25,191 deletions and 30,118 insertions. Finally, we emphasized the landscape of the genotyped SVs segregating across 571 SR individuals of 14 breeds.

7. Breeding for thermotolerance in dairy cattle: Production versus fertility traits

Carabaño, M.J., Díaz, C., Ramón, M. Breeding for thermotolerance in dairy cattle: Production versus fertility traits. Journal of Dairy Science 108, 9915–9929 (2025). https://doi.org/10.3168/jds.2025-26258

The decline in production of milk and its components has been extensively studied as an indicator of heat tolerance for genetic evaluations. However, the antagonistic relationship between high production and functionality raises questions about the suitability of using productive traits as indicators of heat tolerance. This study aimed to estimate changes in the relationship between production and fertility under thermoneutral (TN) and heat stress (HS) conditions, to define breeding strategies that enhance adaptation to high heat loads while maintaining both productivity and functionality. The analyzed dataset included records on 100,467 Holstein cows of first-lactation milk, fat, and protein yields (703,574 records for each yield) and conception rate (CR) at first insemination in first lactation of 247,378 cows in Spain. Temperature-humidity indices averaged over the day of milk recording or day of artificial insemination and the 2 previous days for milk traits, or the subsequent 7 d for fertility, were used to measure the heat load associated with each record. Bicharacter sire models were employed, incorporating one of the yield traits and the fertility trait. Models included random regressions with Legendre polynomials for production traits and a broken-line function for fertility to describe the trait responses to increasing heat loads. This approach allowed for the estimation of trait levels under TN and HS conditions, the slope of response under HS as heat tolerance indicators, and the correlations among these variables. The 3 yield traits exhibited estimated negative genetic correlations between their level under TN conditions and their slopes of response under HS, ranging from −0.38 for fat yield to −0.59 for milk yield. For CR, this correlation was close to zero. Estimated genetic correlations between yield traits under TN conditions and the decline in CR under HS were nearly null, ranging from −0.06 for fat yield to 0.07 for protein yield. This suggests that cows with higher production potential under TN conditions are not necessarily more susceptible to fertility decline under HS. Conversely, the correlations between fertility potential under TN conditions and the slopes of production decline under HS were positive, ranging from 0.34 for protein yield to 0.51 for fat yield. This indicates that cows with lower production losses under HS tend to have better fertility performance under TN conditions. Furthermore, the correlations between heat tolerance based

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

on production and fertility declines under HS were positive, ranging from 0.22 for fat yield to 0.65 for milk yield. This suggests that a significant proportion of animals have the potential to maintain both productive and fertility levels under HS. Finally, the genetic correlation between fertility and production traits improved as heat load increased. For milk yield, this correlation shifted from −0.30 under TN to nearly null under extreme heat conditions. Reaction to heat load in functional traits such as fertility should help in selecting animals that show high levels of production under HS due to a better adaptation to hot conditions driven by functional reasons.

8. CRISPR/Cas9-editing of PRNP in Alpine goats

Allais-Bonnet, A., Richard, C., André, M. et al. CRISPR/Cas9-editing of PRNP in Alpine goats. Vet Res 56, 11 (2025). https://doi.org/10.1186/s13567-024-01444-1

Misfolding of the cellular PrP (PrPc) protein causes prion disease, leading to neurodegenerative disorders in numerous mammalian species, including goats. A lack of PrPc induces complete resistance to prion disease. The aim of this work was to engineer Alpine goats carrying knockout (KO) alleles of PRNP, the PrPc-encoding gene, using CRISPR/Cas9-ribonucleoproteins and single-stranded donor oligonucleotides. The targeted region preceded the PRNPTer mutation previously described in Norwegian goats. Genome editors were injected under the zona pellucida prior to the electroporation of 565 Alpine goat embryos/oocytes. A total of 122 two-cell-stage embryos were transferred to 46 hormonally synchronized recipient goats. Six of the goats remained pregnant and naturally gave birth to 10 offspring. Among the 10 newborns, eight founder animals carrying PRNP genome-edited alleles were obtained. Eight different mutated alleles were observed, including five inducing KO mutations. Three founders carried only genome-edited alleles and were phenotypically indistinguishable from their wild-type counterparts. Among them, one male carrying a one base pair insertion leading to a KO allele is currently used to rapidly extend a PRNP-KO line of Alpine goats for future characterization. In addition to KO alleles, a PRNPdel6 genetic variant has been identified in one-third of founder animals. This new variant will be tested for its potential properties with respect to prion disease. Future studies will also evaluate the effects of genetic background on other characters associated with PRNP KO, as previously described in the Norwegian breed or other species.

9. Genetic analysis of the impact of heat stress on fertility traits in dairy cows in the Netherlands

Ojo, T.O., Vandenplas, J., Mulder, H.A., van Pelt, M.L., Calus, M.P.L. Genetic analysis of the impact of heat stress on fertility traits in dairy cows in the Netherlands. Journal of Dairy Science 108, 1699–1713 (2025). https://doi.org/10.3168/jds.2024-25316

This study aimed to use temperature-humidity index (THI) as an indicator for assessing heat stress conditions for fertility traits in the Holstein dairy cattle breed in the Netherlands. Data from AI and calving events of 416,814 first-parity cows from the Netherlands were used, considering different THI definitions based on different numbers of days before and after artificial insemination events. To achieve our aim, we investigated first, at population level, the relationship between different THI definitions and 4 fertility traits: conception rate, interval

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

calving to first insemination, interval first to last insemination, and calving interval. Second, to investigate individual variation in the relationship between THI and fertility, variance components were estimated for each trait using the so-called broken stick model. This model assumes that breeding values are dependent on THI above but not below a THI threshold identified at the population level and explores the presence of genetic variation associated with fertility decline during heat stress. This study revealed considerable changes in fertility traits during periods of heat stress, with a THI threshold of 60 for conception rate and interval first to last insemination and 50 for interval calving to first insemination and calving interval. Interestingly, as THI levels increased, genetic variance and heritability also increased, indicating that at higher THI levels associated with reduced fertility, the genetic variation of fertility traits is greater. Furthermore, significant genotype-by-environment interactions were observed for all 4 fertility traits, suggesting changes in sire rankings between THI levels below and above the threshold. This study provides insights that may help breeding programs and farmers breed animals resilient to heat stress conditions.

10. A Recursive Model Approach to Include Epigenetic Effects in Genetic Evaluations Using Simulated DNA Methylation Effects

López-Catalina, A.^1,2,3, Ragab, M.^1, Reverter, A.^3, González-Recio, O.^1 (2025). A recursive model approach to include epigenetic effects in genetic evaluations using simulated DNA methylation effects. Journal of Animal Breeding and Genetics, 142(5), 550–559. https://doi.org/10.1111/jbg.12925

The advancement of epigenetics has highlighted DNA methylation as an intermediate-omic influencing gene regulation and phenotypic expression. With emerging technologies enabling the large-scale and affordable capture of methylation data, there is growing interest in integrating this information into genetic evaluation models for animal breeding. This study used methylome information from six dairy cows to simulate the methylation profile of 13,183 genotyped animals. The liability to methylation was treated as an additive trait, while a trait moderated by methylation effects was also simulated. A multiomic model (GOBLUP) was adapted to incorporate methylation data in genomic and genetic evaluations, using the traditional BLUP method as a benchmark. The GOBLUP accurately recovered heritability estimates for the liability to methylation in all low, medium and high heritability scenarios and was consistent at estimating the heritability for the epigenetics-moderated trait of interest at a low–medium heritability of 0.14. The genetic variance recovered by the BLUP model was influenced by the h2 of the liability to methylation, and a part of the methylation variance for the phenotypic trait was captured as additive. The h2 of the phenotypic trait partially relies on the h2 value for the methylation windows in the traditional model. A newly proposed estimated epigenetic value (EEV) combines the traditional additive genetic information from genotyping arrays with epigenetic information. The correlation between the traditional estimated breeding value (EBV) and EEV was high (0.92–0.99 depending on the scenario), but the correlation of the EEV with the true breeding value was higher than the correlation between the traditional EBV and the TBV (0.85 vs. 0.75, 0.71 vs. 0.66 and 0.61 vs. 0.62 depending on the scenario). This study

The RUMIGEN project has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

demonstrates that the GOBLUP multiomic recursive model can effectively separates additive and epigenetic variances, enabling improved breeding decisions by accounting for genetic liability to DNA methylation. This enables more informed breeding decisions, optimising selection for desired traits. Emerging sequencing techniques offer new opportunities for cost-effective simultaneous acquisition of genetic and epigenetic data, further enhancing breeding accuracy.

11. Genetic regulation of sperm DNA methylation in cattle through meQTL mapping

Fouéré, C., Costes, V., Hozé, C. et al. Genetic regulation of sperm DNA methylation in cattle through meQTL mapping. BMC Genomics 26, 771 (2025).
https://doi.org/10.1186/s12864-025-11934-x

Background

DNA methylation (DNAm) plays an important functional role and is influenced by genetic variants known as methylation QTLs (meQTLs). The majority of meQTL studies have been conducted in human blood. Despite its unique landscape, the genetic regulation of sperm DNAm remains largely unexplored. In this study, we leveraged DNAm measured in sperm from 405 Holstein bulls using reduced representation bisulfite sequencing (RRBS) and performed sequence-level genome-wide association studies for 166,985 variable CpGs (s.d. >5%). We reported heritability estimates and have mapped both cis-meQTLs and trans-meQTLs.

Results

Heritability estimates ranged from 0 to 1 and averaged 0.26 across all selected CpGs, with 76% of estimates above 0.1. The meQTL mapping revealed that 32.9% of the CpGs had a cis-meQTL, 3.6% had a trans-meQTL and 1.0% had both cis- and trans-meQTLs. The cis-CpGs were located on average 261 kb (absolute mean) from their cis-meQTL top SNPs (defined by the most significant association). MeQTLs were enriched in featured genomic annotations, including regions surrounding transcription start sites and ATAC-seq peaks. We also identified spurious trans-associations by analyzing data across multiple genome assemblies, including the construction of a partial pangenome. Additionally, eight trans-meQTL hotspots, defined as variants associated with at least 30 trans-CpGs, were identified and overlapped with genes involved in epigenetic regulation. Using peripheral blood mononuclear cell DNAm from 54 out of the 405 bulls, we did not observe a similar effect of the trans-meQTL hotspots to that one observed in sperm.

Conclusions

For the first time, meQTLs have been detected and characterized in bovine sperm, contributing to a better understanding of the transmission of paternally inherited DNAm marks. These findings provide useful information for further research aimed at integrating epigenetic information into the prediction of performance traits.

The RUMIGEN project has received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.

12. Comprehensive detection of structural variations in long and short reads dataset of French cattle

Naji, M.M., Klopp, C., Eché, C. et al. Comprehensive detection of structural variations in long and short reads dataset of French cattle. Sci Rep 15, 40758 (2025). https://doi.org/10.1038/s41598-025-24477-8

Structural variants (SVs) correspond to different types of genomic variants larger than 50 bp. Many findings suggest the use of long-read (LR) rather than short-read (SR) sequencing to improve the accuracy of SVs detection. Here, we present the results of an in-depth analysis for detection of SVs, mainly large insertions and deletions, in 14 French bovine breeds, based on whole-genome sequence (WGS) data comprising 176 LR and 571 SR samples, with 154 individuals having both LR and SR data available. We first investigated possible biases on the performances of well-known SVs detection tools, namely CUTESV, PBSV, and SNIFFLES, using LR from different technologies, including PacBio HiFi, Oxford ONT, and PacBio CLR. We subsequently highlighted the abilities of tools for detecting SVs (DELLY, LUMPY, and MANTA) and for genotyping known SVs (GRAPHTYPER, SVTYPER, PARAGRAPH, and VG toolkit) using SR data. We then show how the incremental composition of samples in the reference panel affected the SVs genotyping for six validation individuals sequenced in SR. We then searched for the optimal parameters and created the final SVs reference panel consisting of 25,191 deletions and 30,118 insertions. Finally, we emphasized the landscape of the genotyped SVs segregating across 571 SR individuals of 14 breeds.

2026

13. Pangenome-based association testing between a structural variant located upstream of the KIT gene and head depigmentation across a diverse panel of cattle breeds

Sorin, V., Sanchez, M.P., Drouilhet, L. et al. Pangenome-based association testing between a structural variant located upstream of the KIT gene and head depigmentation across a diverse panel of cattle breeds. Genet Sel Evol 58, 17 (2026).
https://doi.org/10.1186/s12711-026-01037-w

Coat color variation is a key phenotypic trait in domestic animals. Among the genetic factors involved, the KIT gene has frequently been associated with pigmentation diversity across species. In cattle, spotting or piebald phenotypes have been linked to variation in the genomic region encompassing KIT, but the identification of causal variants was not always possible. This is largely due to the regulatory nature of the underlying variants and the structural complexity of this genomic region, which remains difficult to investigate with linear reference genome-based approaches. In the present study, we used a local pangenome strategy to investigate a genomic region on chromosome 6 encompassing KIT which was recently suggested to be associated with head depigmentation in white-headed cattle breeds. We constructed a 2 Mb pangenome graph encompassing the associated region using 79 assemblies from 20 cattle breeds. Through the evaluation of the coverage at the node level on this pangenome graph, we identified a ~ 7 kb structural variant which was supported by 21 assemblies only from breeds exhibiting a white-headed phenotype. To validate these findings, we aligned 564 short-read sequencing data to a local graph of 30 kb, spanning the structural variant identified, and computed normalized coverage across the region. White-headed cattle breeds consistently exhibited higher coverage values, while color-headed breeds displayed nearly zero coverage. Together, these results confirm the association between a structural variant upstream of KIT with the white-headed phenotype. More broadly, our study demonstrates how targeted local pangenome graphs can efficiently resolve complex structural variants (SVs) with phenotypic impact, offering an interesting and computationally feasible alternative to whole-genome graph approaches.

14. Assembly of a pangenome uncovers novel non-reference unique insertion sequences in cattle highlighting their genetic diversity

Sorin, V., Besnard, F., Capitan, A. et al. Assembly of a pangenome uncovers novel non-reference unique insertion sequences in cattle highlighting their genetic diversity. J Animal Sci Biotechnol 17, 47 (2026). https://doi.org/10.1186/s40104-026-01373-3

Background

The current cattle reference genome, derived from a single Hereford cow, does not capture the full spectrum of genetic diversity present within the species. Moreover, detecting structural variations (SVs ≥ 50 nucleotides long) remains challenging using only standard approaches of either short or long-read sequence approaches against a linear reference genome. Recent advances in long-read sequencing technologies and graph-based assembly now enable the construction of breed-specific pangenomes, revealing previously uncharacterized genomic regions that may contribute to important agricultural traits.

Results

In this study we constructed a cattle pangenome graph using 16 high-quality haplotype-resolved genome assemblies originating from nine breeds representing the diversity of French cattle populations, and including yak (Bos grunniens) as a close outgroup species. Using a trio-based strategy combined with complementary sequencing technologies and bioinformatics methods, we identified and characterized 101,219 structural variations. Of these, 33,634 were classified as non-reference unique insertions (NRUIs), adding several megabases of novel genomic sequences absent from the current Hereford reference genome.

Analysis of the distribution of these NRUIs revealed significant genome-wide enrichment within QTL regions associated with milk production and morphological traits, suggesting their contribution to the genetic basis of economically relevant phenotypes. Furthermore, their functional annotation highlighted two NRUIs located within the intronic regions of ARMH3 and EPHA5, both specific to the Normande breed and significantly associated with milk production and morphological traits, respectively.

Conclusions

Our findings demonstrate the value of pangenome approaches to uncover functionally relevant SVs, particularly NRUIs, that are systematically not in the current reference genome. By linking these variants to economically important traits, our work underscores the need to incorporate breed diversity into future genomic analyses and reference-building efforts in cattle.

15. Impact of genomic selection on genetic diversity in five local European cattle breeds

Bonifazi, R., Meuwissen, T. H. E., Croiseau, P., Restoux, G., Minéry, S., Vandenplas, J., & Windig, J. J. (2026). Impact of genomic selection on genetic diversity in 5 local European cattle breeds. Journal of Dairy Science. https://doi.org/10.3168/jds.2025-27562

Genomic selection (GS) has revolutionised animal breeding and accelerated genetic gains in breeding programs. While GS has become common in cosmopolitan dairy cattle breeds, its implementation in local breeds has begun only more recently or is still in progress. However, the introduction of GS in some cosmopolitan breeds has also been associated with increased inbreeding rates, raising concerns about the potential effects of GS on the genetic diversity in smaller or local breeds. Our aim was to investigate the impact of GS on genetic diversity in five (small) local cattle breeds from three European countries. The five breeds evaluated were: Meuse Rhine Yssel (MRY; from the Netherlands), Norwegian Red (from Norway), Abondance, Tarentaise, and Vosgienne (from France). We investigated changes in population demographic structure, as well as trends and rates of kinship and inbreeding, using both pedigree- and genomic-based measures. The population size varied depending on the breed, with Vosgienne being the smallest and Norwegian Red being the largest. Single nucleotide polymorphisms (SNPs) genotypes were available for 4,645 MRY, 193,489 Norwegian Red, 16,387 Abondance, 8,578 Tarentaise, and 4,472 Vosgienne animals. Animals were genotyped with more than 40,000 SNPs. Overall, following the implementation of GS in these breeds, we observed a reduction of up to 4 years in generation intervals for sires, fewer calves that later became sires, and, for the French breeds, a broader sire usage. Such changes were likely due to GS enabling the preselection and screening of more young bulls. Additionally, the contributions of the top ten sires were more evenly distributed after the introduction of GS. Although changes in inbreeding and kinship rates occurred after the introduction of GS, there was no consistent pattern across breeds: pedigree (and genomic ROH-based) inbreeding rates per generation increased in MRY from -0.67 before GS to 0.51 after GS (from -1.12 to 0.93) and Tarentaise from 0.35 to 0.93 (from 0.68 to 0.86), but decreased in Norwegian Red from 0.26 to 0.05 (from 0.10 to 0.06), Abondance from 1.19 to 0.99 (from 2.39 to 0.58), and Vosgienne from 0.53 to 0.23 (from 0.88 to 0.19). Moreover, analysis of genomic ROH-based inbreeding by length class showed that, after the implementation of GS, the largest changes in inbreeding level and inbreeding rates per generation occurred for shorter ROH segments. Our study suggests that changes and increases in inbreeding rates may occur after the introduction of GS, although they may not be directly due to the introduction of GS per se, but rather due to population management strategies, such as optimal contribution selection. Our findings emphasise the importance of monitoring changes in both genetic diversity and population demographic structure after implementing GS in local breeds, as well as adjusting breeding strategies when needed to ensure long-term sustainability.

16. Haplotype-resolved genome assemblies for Norwegian Red cattle (preprint at the moment)

T. J. Hettasch[1], M. P. Kent[1], A. B. Gjuvsland[2], M. Árnyasi[1], D. I. Våge[1]
1Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences (NMBU), Oluf Thesens vei 6, 1433 Ås, Norway
2Geno SA, Storhamargata 44, 2317 Hamar, Norway
https://doi.org/10.64898/2025.12.18.695069

Norwegian Red (NR) cattle are the main dairy breed in Norway, bred according to a broad breeding goal including health and fertility since the 1970s. Genomic studies on NR cattle have relied on the public Hereford reference, thus increasing the risk of missing or misrepresenting NR breed-specific variation. Moreover, the Hereford reference is a pseudohaploid assembly, representing homologous chromosomes in a collapsed manner, which results in loss of haplotype-specific alleles and misrepresentation of complex variants. To develop more refined NR specific resources, we utilised long-read sequencing (PacBio HiFi + ONT) and trio-binning to construct six new haplotype-resolved assemblies representing NR genomes. These six NR2025 assemblies show high completeness (BUSCO: 95.82-98.11%) and contiguity (N50: 73.8-88.5 Mb) and are accurately phased (hamming error rate: 0.46-2.52%). Most autosomes have been assembled into the acrocentric centromere, and mapping of bovine satellite sequences reveal distinct organisational patterns of different satellite units across this highly repetitive region. The NR2025 assemblies provide a valuable resource for identification of novel variants and haplotypes in the NR population, which will enable more accurate association studies of genotype-to-phenotype relations and genomic predictions in NR cattle, ultimately enhancing the efficacy of selection for desirable traits.

The RUMIGEN project has received funding from the European Union’s Horizon 2020 research and innovation program under Grant Agreement No 101000226. This publication reflects the views only of the author, and the European Union cannot be held responsible for any use which may be made of the information contained therein.