

SMARTER

SMALL RuminanTs breeding for Efficiency and Resilience

Research and Innovation action: H2020 – 772787

Call: H2020-SFS-2017-2

Type of action: Research and Innovation Action (RIA)

Work programme topic: SFS-15-2016-2017

Duration of the project: 01 November 2018 – 31 October 2022

A manuscript on the development and testing of genomic evaluation tools to improve resilience and efficiency accounting for the impact of environmental challenge and using novel phenotypes

Masoud Ghaderi Zefreh, Andrea B. Doeschl-Wilson,

Valentina Riggio, Oswald Matika and Ricardo Pong-Wong*

The Roslin Institute, The University of Edinburgh.

Easter Bush Campus, Midlothian, EH25 9RG. Scotland UK

* Deliverable leader – Contact: ricardo.pong-wong@roslin.ed.ac.uk

DELIVERABLE D5.5

Workpackage N°5

Due date: M56

Actual date: 10/07/2023

Dissemination level: Confidential or Public

About the SMARTER research project

SMARTER will develop and deploy innovative strategies to improve Resilience and Efficiency (R&E) related traits in sheep and goats. SMARTER will find these strategies by: i) generating and validating novel R&E related traits at a phenotypic and genetic level ii) improving and developing new genome-based solutions and tools relevant for the data structure and size of small ruminant populations, iii) establishing new breeding and selection strategies for various breeds and environments that consider R&E traits.

SMARTER with help from stakeholders chose several key R&E traits including feed efficiency, health (resistance to disease, survival) and welfare. Experimental populations will be used to identify and dissect new predictors of these R&E traits and the trade-off between animal ability to overcome external challenges. SMARTER will estimate the underlying genetic and genomic variability governing these R&E related traits. This variability will be related to performance in different environments including genotype-by-environment interactions (conventional, agro-ecological and organic systems) in commercial populations. The outcome will be accurate genomic predictions for R&E traits in different environments across different breeds and populations. SMARTER will also create a new cooperative European and international initiative that will use genomic selection across countries. This initiative will make selection for R&E traits faster and more efficient. SMARTER will also characterize the phenotype and genome of traditional and underutilized breeds. Finally, SMARTER will propose new breeding strategies that utilise R&E traits and trade-offs and balance economic, social and environmental challenges.

The overall impact of the multi-actor SMARTER project will be ready-to-use effective and efficient tools to make small ruminant production resilient through improved profitability and efficiency.

Table of contents

About the SMARTER research project	1
1 Summary	3
2 Introduction.....	3
3 Methods	3
4 Results	3
5 Deviations or delays.....	4
6 References.....	4
7 Appendix	4

1 Summary

This deliverable is a manuscript presenting the results from a theoretical study on the value of genomic prediction using a reaction norm model to simultaneously improve production and resilience. The factors affecting the accuracy of the evaluation were identified, to optimise the design of breeding scheme aiming to improve both traits. The potential selection response over several generation. The results from the study has been published in a peer reviewed journal (Ghaderi Zefreh et al (2023), Exploring the value of genomic predictions to simultaneously improve production potential and resilience of farmed animals. *Front. Genet.* 14:1127530. doi: 10.3389/fgene.2023.1127530).

2 Introduction

Farmed animals are exposed a wide range of environmental challenges during the lifespan. Such challenges have a detrimental effect on both the animal's welfare and their productivity. Thus, sustainable livestock production requires for animals to have a high production potential but are also highly resilient to maintain such production when exposed to environmental challenges. In this study we evaluated the benefit of genomic prediction to predict breeding values for both production potential and resilience. Factors which may impact the quality of the evaluation were identified. The potential selection response when the aim is to improve both traits was quantified. Full details of the study and results are given in the published paper (doi: 10.3389/fgene.2023.1127530) given appendix.

3 Methods

Simulation was carried out to quantify the benefit of genomic prediction on the accuracy of breeding values for production potential and resilience using a reaction norm model.

To aid the optimisation of breeding scheme design, the effect of several factors including the population parameters of the traits, the distribution of phenotypes across environment and the uncertainty of the level of challenge, were compared.

The feasibility of selection for improving both production potential and resilience was assessed by simulating a population undergoing selection over several generations based on genomic estimated breeding values obtained implementing a reaction norm model.

4 Results

The results from this study showed that genomic prediction can double the accuracy of estimated breeding values. The accuracy of production potential was affected by the distribution of phenotypes across the environment, where the maximum accuracy is obtained when families are distributed across the whole range of environment. The level of uncertainty on the level of environmental challenge can a significant effect on the quality of the evaluation, where accuracy can reduce as much as a third when the level of uncertainty is 50%. The selection response for production potential and resilience was also affected by the degree of challenge of environments where candidates are recorded: selection

response for resilience increased when the candidates were recorded in environments with high level of challenge.

See manuscript in appendix for further details.

5 Deviations or delays

Submitted on the 9th of June, 2023.

6 References

Ghaderi Zefreh M, Wilson AB, Riggio V, Matika O and Pong-Wong R (2023). Exploring the value of genomic predictions to simultaneously improve production potential and resilience of farmed animals. *Front. Genet.* 14:1127530. doi: 10.3389/fgene.2023.1127530).

7 Appendix

The manuscript.



OPEN ACCESS

EDITED BY

Johann Sölkner,
University of Natural Resources and Life
Sciences Vienna, Austria

REVIEWED BY

Florence Phocas,
INRAE Centre Jouy-en-Josas, France
Suresh Neethirajan,
Farmworx Research Institute,
Netherlands

*CORRESPONDENCE

Masoud Ghaderi Zefreh,
✉ mghaderi@aed.ac.uk

RECEIVED 19 December 2022

ACCEPTED 02 May 2023

PUBLISHED 12 May 2023

CITATION

Ghaderi Zefreh M, Doeschl-Wilson AB,
Riggio V, Matika O and Pong-Wong R
(2023), Exploring the value of genomic
predictions to simultaneously improve
production potential and resilience of
farmed animals.

Front. Genet. 14:1127530.

doi: 10.3389/fgene.2023.1127530

COPYRIGHT

© 2023 Ghaderi Zefreh, Doeschl-Wilson,
Riggio, Matika and Pong-Wong. This is an
open-access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Exploring the value of genomic predictions to simultaneously improve production potential and resilience of farmed animals

Masoud Ghaderi Zefreh^{1*}, Andrea B. Doeschl-Wilson¹,
Valentina Riggio¹, Oswald Matika^{1,2} and Ricardo Pong-Wong¹

¹The Roslin Institute and R(D)SVS, University of Edinburgh, Edinburgh, United Kingdom, ²Centre for Tropical Livestock Genetics and Health (CTLGH), The Roslin Institute and R(D)SVS, University of Edinburgh, Edinburgh, United Kingdom

Sustainable livestock production requires that animals have a high production potential but are also highly resilient to environmental challenges. The first step to simultaneously improve these traits through genetic selection is to accurately predict their genetic merit. In this paper, we used simulations of sheep populations to assess the effect of genomic data, different genetic evaluation models and phenotyping strategies on prediction accuracies and bias for production potential and resilience. In addition, we also assessed the effect of different selection strategies on the improvement of these traits. Results show that estimation of both traits greatly benefits from taking repeated measurements and from using genomic information. However, the prediction accuracy for production potential is compromised, and resilience estimates tends to be upwards biased, when families are clustered in groups even when genomic information is used. The prediction accuracy was also found to be lower for both traits, resilience and production potential, when the environment challenge levels are unknown. Nevertheless, we observe that genetic gain in both traits can be achieved even in the case of unknown environmental challenge, when families are distributed across a large range of environments. Simultaneous genetic improvement in both traits however greatly benefits from the use of genomic evaluation, reaction norm models and phenotyping in a wide range of environments. Using models without the reaction norm in scenarios where there is a trade-off between resilience and production potential, and phenotypes are collected from a narrow range of environments may result in a loss for one trait. The study demonstrates that genomic selection coupled with reaction-norm models offers great opportunities to simultaneously improve productivity and resilience of farmed animals even in the case of a trade-off.

KEYWORDS

resilience, robustness, reaction norm, genomic prediction, genomic selection, GxE, trade-off

1 Introduction

Farmed animals can be exposed to a wide range of environmental challenges during their development and lifespan. Climate change is known to exacerbate these challenges, e.g., through increased temperature fluctuations or more frequent occurrence of extreme weather conditions (Thornton et al., 2007; Baumgard et al., 2012) and associated food shortages or increased burden of infectious pathogens (Tomley and Shirley, 2009). It is thus desirable, for

both animals' welfare and productivity, that animals not only have a high production potential in ideal environmental conditions, but are also able to maintain it when exposed to environmental or infectious challenges. In the animal breeding community, this capacity has been defined as resilience and has been considered as an important breeding goal since decades (Bisset and Morris, 1996; Hermes and Dominik, 2014; Knap and Doeschl-Wilson, 2020). In the context of intensive livestock production, resilience and robustness are often used interchangeably (Colditz and Hine, 2016; Friggens et al., 2016). Here, we follow the definition of Knap (2005), who defined robustness as the combination of high production potential with high resilience to external stressors, allowing for unproblematic expression of that production potential in a wide variety of environmental conditions. The performance of an animal in an environment is thus determined by its production potential, resilience, and the level of the overall challenge the animal faces in that environment.

Neither the resilience nor the production potential of an animal in ideal environmental conditions are directly measurable. Instead, estimates for both traits can be obtained by using so called reaction norm (RN) models in which the performance of an animal is regressed against the environmental challenge level (de Jong, 1990). In the case of a linear RN model, the production potential then refers to the model intercept and resilience refers to the inverse of the slope (Mulder, 2016). Genetic parameter estimates and estimated breeding values for both traits can then be obtained through random regression approaches (Strandberg, 2006). Such random regression models have been used to estimate animals' resilience to temperature or other quantifiable climatic measures (Knap and Su, 2008; Nguyen et al., 2016; Bohlouli et al., 2019; Sánchez-Molano et al., 2019).

One of the hurdles for estimating resilience is that in many cases the environmental challenge level is unknown or difficult to quantify. This is particularly the case when animals are exposed to infectious pathogens (Knap and Doeschl-Wilson, 2020) or a whole cocktail of different environmental stressors (e.g., of multiple pathogens, sub-optimal nutritional resources, etc.). In these cases, the common approach is to use the contemporary group mean performance as a proxy for the environmental challenge level in the RN models (Knap, 2005; Strandberg, 2006; Rashidi et al., 2014; Hermes et al., 2015). To avoid bias in the genetic parameter estimates for the regression parameters, the contemporary group means (e.g., herd effect or herd-season-year effect) representing the environmental value in the regression models are usually estimated together with the regression parameter estimates in an iterative procedure (Calus et al., 2004; Knap and Su, 2008). Given that the average group performance is most likely a combination of many factors, this proxy may indeed provide a good overall description of the type of environment the animals are exposed to (Strandberg, 2006). However, it is important to note that this measure is only an approximation of the true environmental challenge that each individual is exposed to, and that the ideal environment and thus also the deviation from it, i.e., the challenge level, may vary between individuals. However, it is currently not known how inaccuracies in the estimates of the actual environmental challenge levels of individuals affect the prediction accuracies for production potential and resilience.

Another obstacle in estimating resilience through a RN approach is the lack of performance data of an individual across multiple

environments. Many phenotypes have limited measurements per animal, e.g., carcass weight and hence rely on phenotypic records of related individuals (Strandberg, 2006). The sparsity of phenotypic records for an individual can have a negative impact on the accuracy of estimated breeding values (EBVs) (Cameron, 1997). However, genomic prediction has proven to be beneficial in improving prediction accuracies under these conditions (Meuwissen et al., 2001; Mulder, 2016), with a few studies having assessed the benefit of genomic prediction on the accuracies of resilience and production potential. In particular, Calus et al. (2004) showed that prediction accuracies of RN model parameters strongly depend on how related individuals are distributed across different environments. Furthermore, in scenarios where related individuals are reared in the same environment, accuracy of EBVs for RN slope and intercept are adversely affected (Calus et al., 2004). However, it is not known whether genomic prediction can partly or completely overcome this issue, because the genomic relationship matrix accounts for similarities between related as well as unrelated individuals.

Numerous studies have applied RN models to predict animals' resilience or performance under different environmental challenge conditions (Knap and Su, 2008; Herrero-Medrano et al., 2015; Li and Hermes, 2016; Oliveira et al., 2019; Song et al., 2020; Garcia-Baccino et al., 2021). One of the clear benefits of linear RN models is that they provide genetic parameter estimates for both breeding goal traits, production potential and resilience. Therefore, depending on the genetic parameters, it may be possible to concomitantly select for both traits given appropriate indices. However, RN models are also known to require a large amount of data for convergence (Knap and Su, 2008). Therefore, many studies continue to use conventional genetic evaluation models for production performance, that don't explicitly quantify the environmental challenge level, but account for potential differences in these by fitting fixed group effects. Although few studies have assessed selection response in performance in different environments for RN models compared to conventional models (Kolmodin and Bijma, 2004; Mulder, 2016; Mulder and Rashidi, 2017), the genetic improvement in production potential and resilience as breeding goal traits, that can be achieved by either approach, has not been explicitly assessed. These are likely to also depend on the environmental conditions under which phenotypic performance records are collected (Calus et al., 2004; Mulder, 2016; Le et al., 2022).

Therefore, the first objective of this study was to assess the benefits of using genomics over pedigree-based predictions in obtaining accurate and unbiased breeding value estimates for production potential and resilience from linear RN models. In particular, we assessed how prediction accuracies depend on the number of phenotypic records per individual, the distribution of related animals across environments, and on the ability to accurately quantify the environmental challenge level. The second objective was to assess the feasibility of RN models to improve production potential and resilience and compare these to the response to selection that can be achieved by using conventional models under a range of conditions.

2 Methods

The potential benefit of genomic selection to select for production potential and resilience under a linear RN model was

assessed using simulations of an outbred population of farmed animals. Without loss of generality, in this study the genomes of sheep populations were simulated, but it is expected that results would be similar for other population of terrestrial farmed animals. To assess the impact of genomics and other factors on prediction accuracies and bias, a sheep population undergoing random selection was simulated. To assess response to selection, a population undergoing selection for 10 generations under different selection strategies was simulated. Each simulation scenario was replicated 100 times.

2.1 Genetic models for production potential and resilience

The phenotype was defined assuming a RN model, where the individual's performance is affected by two components, simulated as random effects: one related to the performance under ideal condition and another being linearly related to an environmental challenge level the individual is exposed to.

Here, the respective components are denoted as production potential (subscripts 0) and resilience (subscripts R) and both are assumed to be under genetic (A) and environmental (E) control. Hence, the performance, y , of individual j in an environment with challenge level X is:

$$y_{j,k}(X) = (\mu_0 + A_{0j} + E_{0jk}) + X(\mu_R + A_{Rj} + E_{Rjk}), \quad (1)$$

where the subscript k corresponds to the time of measurement in case of multiple records for an individual. In the above equation, μ_0 refers to the population mean performance in the absence of challenge, μ_R is the population average rate of reduction in performance (as the environmental challenge generally has adverse effect on performance, μ_R is negative). The terms A_{0j} and A_{Rj} are the additive genetic effects for production potential and resilience and E_{0jk} and E_{Rjk} are their environmental deviations from the following normal distributions

$$\begin{pmatrix} A_{0j} \\ A_{Rj} \end{pmatrix} \in \mathcal{N}\left(0, \begin{pmatrix} \sigma_{A_0}^2 & \sigma_{A_0 A_R} \\ \sigma_{A_0 A_R} & \sigma_{A_R}^2 \end{pmatrix}\right), \quad \begin{pmatrix} E_{0jk} \\ E_{Rjk} \end{pmatrix} \in \mathcal{N}\left(0, \begin{pmatrix} \sigma_{E_0}^2 & 0 \\ 0 & \sigma_{E_R}^2 \end{pmatrix}\right), \quad (2)$$

where $\sigma_{A_0}^2$, $\sigma_{A_R}^2$ and $\sigma_{A_0 A_R}^2$ are the genetic variance of production potential, resilience and the genetic covariance and $\sigma_{E_0}^2$ and $\sigma_{E_R}^2$ are the environmental variance of production potential and resilience, respectively.

With the parameterisation in Eq. 1, the heritabilities of the production potential and resilience are

$$h_0^2 = \frac{\sigma_{A_0}^2}{\sigma_{A_0}^2 + \sigma_{E_0}^2}, \quad h_R^2 = \frac{\sigma_{A_R}^2}{\sigma_{A_R}^2 + \sigma_{E_R}^2}. \quad (3)$$

The heritability of the performance at a given environmental challenge level X follows (Kolmodin and Bijma, 2004)

$$h^2(X; \sigma^2) = \frac{\sigma_{A_0}^2 + X^2 \sigma_{A_R}^2 + 2X \sigma_{A_0 A_R}}{\sigma_{A_0}^2 + X^2 \sigma_{A_R}^2 + 2X \sigma_{A_0 A_R} + \sigma_{E_0}^2 + X^2 \sigma_{E_R}^2}. \quad (4)$$

We used low heritability values (0.1) for h_0^2 and h_R^2 (see Table 1), however, the profile of changes with respect to other parameters did

not change for other heritabilities (result shown in Supplementary Table SA1).

2.2 Simulation of the population

2.2.1 Genome in linkage disequilibrium

A similar approach as in Lee and van der Werf (2006); Sánchez-Mayor et al. (2022) was adopted: first, a founder population with its genome in linkage disequilibrium (LD) was simulated using a mutation-drift algorithm (Meuwissen et al., 2001). In this regard, a population with a genome divided in several chromosomes is allowed to evolve. Mutations appear, and drift causes them to be lost or to increase in frequency, and after many generations the genome reaches an equilibrium with segregating SNP at a specific LD pattern determined by the parameters used in the simulation, such as population size and mutation rate. To simulate the pattern of a typical farmed sheep population, the initial population was assumed to have 100 individuals (half males and half females) which were allowed to reproduce for 10,000 generations. Their genome was composed of 26 autosomal chromosomes, each of one Morgan length with 200,000 biallelic loci located equidistantly with a mutation rate of 10^{-5} . The mutation rate was tuned such that the LD profile of the final generation matched that of real data for a sheep breed (Kijas et al., 2014). Then, the final generation was expanded within 5 generations to a larger population which served as the gene-pool for sampling the base population for each replicate. The final expanded founder population has 10,000 individuals, 2 haplotypes per individual and per chromosome, 26 chromosomes, and an average number of 4,800 segregating SNPs per chromosome. Hence, for a given replicate, the genome of the base generation was simulated by randomly sampling haplotypes from the expanded founder population; and thereafter the genome of animals from further generations were sampled by dropping haplotypes from offspring to parents assuming Mendelian inheritance law. Two scenarios for pedigree structure and base population was considered as discussed in Section 2.4. Base populations with size 1,200 (1,100 females and 100 males) and 2,970 (2,700 females and 270 males) were simulated from the founder population. In simulation of N individuals of base population from the founder population (N either 1,200 or 3,970), $2N$ haplotypes were randomly drawn from the pool of 20,000 haplotypes. In addition, sampling of haplotypes was done independently for each chromosome to ensure independency of replicates. As shown in the Supplementary Materials, the probability that two drawn samples are dependent is very small. Moreover, as discussed in the next section, location of SNPs and QTLs as well as their effect was independently sampled for each replicate.

2.2.2 Genetic architecture

Once the base generation was sampled in each replicate, 1,500 segregating loci with the highest minimum allele frequency were selected for each chromosome: 500 being randomly assigned to be quantitative trait locus (QTL) affecting the traits and 1,000 being part of the SNP chip array used to calculate the genomic relationship matrix (GRM) needed in the genomic evaluation. The total number

TABLE 1 Parameter values used in simulating phenotypes.

Parameter	Value	Alternative values
μ_0	10	—
μ_R	−3	—
$\sigma_{A_0}^2$	0.1	0.3, 0.6
$\sigma_{A_R}^2$	0.1	0.3, 0.6
$\sigma_{E_0}^2$	0.9	0.7, 0.4
$\sigma_{E_R}^2$	0.9	0.7, 0.4
$\rho_{A_0 A_R}$	−0.5	0, 0.5
F_N	10	—
ϵ	0.1	0.2, 0.3, 0.5
Range of X	[0,2]	[0, 1], [1, 2]

μ_0 : mean production potential, μ_R : mean resilience, $\sigma_{A_0}^2$: additive genetic variance for production potential, $\sigma_{A_R}^2$: additive genetic variance for resilience, $\sigma_{E_0}^2$: environmental variance for production potential, $\sigma_{E_R}^2$: environmental variance for resilience, $\rho_{A_0 A_R}$: genetic correlation, F_N : number of flocks, ϵ : relative range of environmental challenge levels within each flock compared to the entire range of environmental conditions.

of QTLs and SNPs across the whole genome were 13,000 and 26,000, respectively.

The true breeding values (TBV) were calculated as the sum of all QTL effects given the individual's genotypes. The QTLs were assumed to have pleiotropic additive effects on both the production potential and the resilience traits (A_0 and A_R) to allow for non-zero genetic correlation between both traits. The QTL effects for A_0 and A_R were sampled from a bivariate normal distribution with a unit standard deviation and a correlation equal to the targeted genetic correlation between A_0 and A_R (Wientjes et al., 2017), and then they were re-scaled so that the TBV variances for A_0 and A_R in the base generation were equal to the targeted genetic variances. Finally, the environmental deviations E_0 and E_R were sampled to simulate the individuals' phenotypic performance (using Eq. 1) given the level of exposed environmental challenge at the time of phenotype measurement.

2.2.3 Level of environmental challenge

To simulate the level of environmental challenge (X), individuals were assumed to be phenotyped in flocks. These flocks represent micro-environments, defined by a range of environmental challenge levels around a flock specific average challenge level. Thus, for an overall environmental challenge level $[0, X_{\max}]$, individuals in a given flock f were exposed to a subset ranging between $(\hat{X}_f - X_{\max} \frac{\epsilon}{2})$ and $(\hat{X}_f + X_{\max} \frac{\epsilon}{2})$, where $\hat{X}_f$ is the average challenge effect of flock f and ϵ is a parameter determining the heterogeneity of the flock conditions. If multiple records were required for an animal, multiple environmental challenge levels were sampled from the same flock.

The simulation of the level of environmental challenge related to a given recorded phenotype was done as follows: firstly, the average challenge level ($\hat{X}_f$) for each flock (f) was sampled from a uniform distribution within $[0 + X_{\max} \frac{\epsilon}{2}, X_{\max} - X_{\max} \frac{\epsilon}{2}]$; secondly, each individual was allocated to a given flock and all their phenotypic records were assumed to be recorded in it; third the challenge level X

for a given phenotype was sampled from a uniform distribution within the range $[\hat{X}_f - X_{\max} \frac{\epsilon}{2}, \hat{X}_f + X_{\max} \frac{\epsilon}{2}]$, and this value was used to calculate the phenotypic performance using Eq. 1. Hence, the level of the exposed environmental challenge is different for all individuals (and among records if an individual has more than one recorded phenotype) but more alike between individuals allocated to the same flock. In this study $X_{\max}$ was set to 2, which corresponded to an average difference of 6 phenotypic standard deviations in performance compared to the average performance at optimum condition ($X = 0$) (Table 1). Having a relatively large environmental condition allows to define flocks within where we can investigate the effect of, e.g., heterogeneity of flocks and unknown environmental challenge levels.

In this study, ten flocks were simulated, all having equal within-flock heterogeneity, ϵ . The magnitude of ϵ was set to 0.1, but other values were simulated (Table 1) to assess its impact on the genomic prediction with the RN approach when the environmental challenge level is unknown (see below).

Additionally, several methods for allocating individuals across flocks were used to assess the impact of the distribution of phenotypes across environments: (i) random (RND), where individuals were randomly distributed across all flocks, (ii) clustered (CLS), where individuals from the same half-sib family were allocated into one random flock, (iii) assortative (AST), where individuals from the same half-sib family were allocated to the same flock such that offspring from a sire with the highest breeding values for production potential were allocated to the flock with the lowest level of environmental challenge (i.e., the genetically best animals were allocated to the best environment), and (iv) disassortative (DIS), i.e., similar to AST scenario but with reverse order of sires.

2.3 Genetic evaluation

Genetic evaluations were carried out using the best linear unbiased predictor (BLUP) or genomic BLUP (GBLUP)

(Meuwissen et al., 2001) defined by the relationship matrix used in the analysis (i.e., the NRM based on pedigree information for BLUP, and the GRM based on genotype information with GBLUP). The breeding values for production potential and resilience were estimated using random regression reaction norm models (Eq. 1) under two scenarios with respect to environmental challenge level: known or unknown.

When the environmental challenge levels were unknown, the genetic evaluation was carried out in a two-step RN approach (Calus et al., 2002; Kolmodin et al., 2002): initially, a genetic evaluation ignoring resilience was performed to estimate the flock effects, and subsequently a random regression was performed using the obtained flock effect estimates as proxy for the challenge level (Kolmodin et al., 2002; Mulder and Rashidi, 2017; Carvalheiro et al., 2019). To calculate the proxy, the estimated flock effect was rescaled such that the flock with highest average performance was assigned as having no challenge (i.e., $\bar{X} = 0$) and the lowest flock average is assigned the maximum challenge. This proxy value was then assumed to be the challenge level for all phenotypes recorded in the flock in question.

Hence, the use of the proxy with the RN approach introduces two sources of error into the analysis: (i) error in the estimates of the flock effects in the first step of the evaluation, and (ii) error by assigning the same environmental challenge level to all animals in a flock, thus ignoring the within flock heterogeneity in challenge. Therefore, as flocks become more heterogeneous, the estimated proxy becomes more uncertain and less representative of the true challenge level, potentially affecting the quality of the estimates.

In addition, for the comparison of responses to selection based on different breeding criteria traits, EBVs for production performance were calculated using a conventional model without the term associated to the environmental challenge (i.e., the resilience trait) but accounting for the environmental flock effects,

$$y_{jl} = \hat{\mu} + \hat{F}_l + \hat{A}_j + \hat{E}_j, \quad (5)$$

where $\hat{F}_l$ is the fixed effect of the flock l that animal j belongs to, $\hat{\mu}$ is the estimated population average of the production, $\hat{A}_j$ is the estimated breeding value of the animal j and $\hat{E}_j$ is the random environmental effect.

An in-house software was used to estimate the variance components and (G)EBVs using restricted residual maximum likelihood (Patterson and Thompson, 1971; Lee and van der Werf, 2006) and BLUP (Henderson, 1975; Gumedze and Dunne, 2011), respectively. The software is available on <https://github.com/mghaderizefreh/GenEval>. The genomic relationship matrix was calculated following (VanRaden, 2008). The genetic correlation between production potential and resilience was estimated even when the true genetic correlation was zero.

The quality of the genetic evaluation was assessed by the accuracy and bias of (genomic) estimated breeding values, (G)EBV. The prediction accuracy is defined as the correlation between (G)EBV and the TBV. The bias ($\tilde{b}$) is defined as the standardised regression coefficient of TBV on EBV as explained in (Lipschutz-Powell et al., 2012),

$$\tilde{b} = \begin{cases} 1/b - 1 & b \geq 1 \\ b - 1 & b < 1 \end{cases}, \quad (6)$$

where b is the regression coefficient between (G)EBVs and TBVs. Hence, unbiased estimation has a zero value for the standardised

regression coefficient, whereas overestimated or underestimated (G)EBVs have positive or negative standardised coefficients, respectively.

2.4 Scenarios compared

As mentioned above, two different population structures were considered to achieve the objectives of this study: To assess the impact of genomics and other factors on prediction accuracies and bias, a sheep population undergoing random selection was simulated with 3 generations and ~1,000 animals per generation. To assess response to selection, a population undergoing selection for 10 generations under different selection strategies was simulated. The number of animals per generation in the second population was ~3,000.

Table 1 shows the values of parameters used in the simulations that were common for both populations. For the genetic correlations between production potential and resilience, three values were assumed corresponding to a favourable, zero and antagonistic relationship between the traits. Results are mostly shown for the case of negative genetic correlation as the worst-case scenario, because it has the lowest heritability for production performance across environments (c.f. Eq. 4). Results corresponding to medium and high heritability are provided in of the [Supplementary Table S1A](#).

2.4.1 Population under random selection

The scenarios considered here aim to assess the benefit of genomic prediction in improving the quality of the prediction of production potential and resilience as well as to assess factors affecting the predictions. The population consisted of three non-overlapping generations with a half-sib structure. The base or first generation started with 100 males and 1,100 females, with each male being mated to 11 females to produce one offspring per female, resulting in one male and 10 female offspring (i.e., 1,100 individuals in the second generation, 100 males and 1,000 females). In the second generation, each sire was mated with 10 dams to produce 1,000 individuals in the third generation.

This population was used to assess the impact of different factors on the benefit of genomic prediction to evaluate production potential and resilience. They include:

- The effect of population parameters on the benefit of GBLUP over BLUP in terms of improved accuracy.
- The effect of the distribution of individuals across environments when comparing the GEBV accuracies obtained from individuals across flocks allocated assuming the RND, CLS and AST scenarios.
- The effect of uncertainty of the environment challenge level in a RN model situation by studying the impact of the within-flock level heterogeneity on the range of environmental challenge.

Further, we assume that the third generation does not have phenotypes and we assess the (G)EBVs for this generation only. We use these assumption in order to have generally low accuracy so that the above effects are more pronounced.

2.4.2 Population undergoing multiple generations of selection

Performance records in this section were simulated assuming a genetic correlation of -0.5 between resilience and production potential,

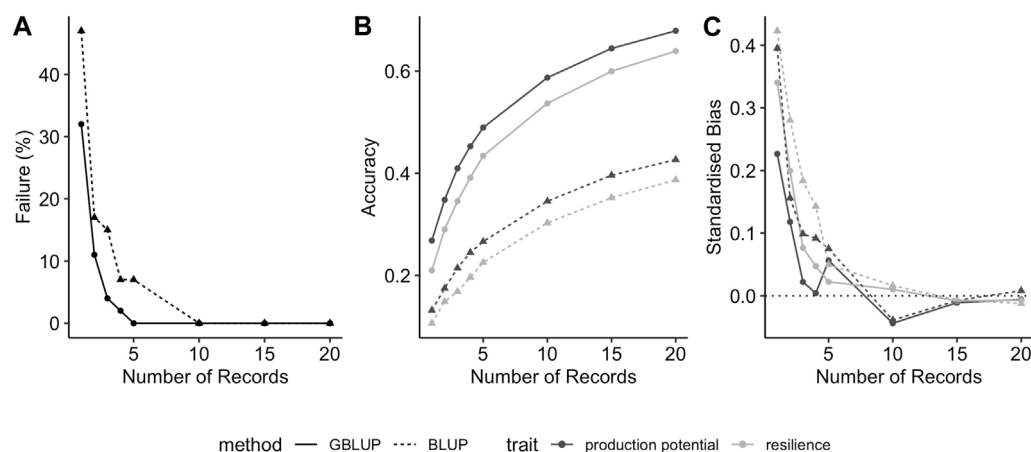


FIGURE 1

Effect of number of records per individual on (A) percentage of simulations where REML failed to converge to feasible variance components (B) prediction accuracy of production potential and resilience, and (C) standardised bias in EBV for production potential and resilience, when using GBLUP (solid lines and circles) or BLUP (dashed lines and triangles).

with flocks covering 10% of the whole environment and one record for each individual in a random distribution scenario. This population structure was used in scenarios that quantified the selection response for production potential and resilience when using a RN model. The population attempted to mimic a sheep population with large half-sib and small full-sib families. The population structure for each generation consisted of 8,100 individuals (2,700 males and 5,400 females), all genotyped with the SNP chip array and having single phenotypic record. At each generation, half of the females and a tenth of males were selected to mate at random (10 females per 1 male) and a litter size of 3 including 1 male and 2 females was assumed. The assumption of discrete generations and litter size of 3, albeit not very common, (Kenyon et al., 2019), was made to maintain the population size across different generation while allowing for a selection intensity for the females. Genetic evaluations were performed using either the RN approach (with unknown environmental challenge level) or the conventional method without RN (i.e., using Eq. 5).

When using RN approach, a selection index was defined as $I = (1 - \alpha)\hat{A}_0 + \alpha\hat{A}_R$ where α ($0 \leq \alpha \leq 1$) is the weight given to resilience. Hence, $\alpha = 0$ corresponds to selection for production potential only, whereas $\alpha = 1$ corresponds to selection for resilience only. The selection scheme was done over 10 discrete generations and individuals were allocated using the RND allocation scenario. We also contrasted scenarios where performance records were generated in a narrow compared to a wide range of environments (Table 1).

3 Results

3.1 Random selection scenario

3.1.1 Comparison between pedigree based and genomic predictions

Figure 1 shows the BLUP and GBLUP results for the genetic evaluation of production potential and resilience depending on the number of phenotypic performance records available per individual. As expected, the extra information from increasing the number of

records per individual resulted in better estimation leading to less convergence failures in the REML analysis (Figure 1A), and higher GEBVs accuracies (Figure 1B) with lower standardised biases (Figure 1C) in the GBLUP/BLUP analyses. The average prediction accuracies of EBVs from BLUP for production potential and resilience incremented as much as 70% when the number of records per individual was increased from one to three (i.e., the accuracies for production potential and resilience with one record per individual were 0.13 and 0.10, respectively, compared to 0.22 and 0.17 with three records per individual). Consistently across the scenarios, the GBLUP evaluation outperformed BLUP. For instance, when assuming three records for each animal, the REML analysis using pedigree information failed to converge in 15% of the replicates whereas using genomic information only 4% of the replicates failed to converge (Figure 1A). For the same scenario the GEBV accuracy was 91% and 105% higher than that of EBVs for production potential and resilience, respectively (Figure 1B). In particular, for both traits, prediction accuracy assuming one record per animals with GBLUP was as high as that of BLUP with five records. Regardless of the method used (pedigree or genomic), estimated breeding values for production potential were in general more accurate than those for resilience, e.g., for three records per individuals, prediction accuracies for production potential were 19% higher than those for resilience. Some degree of bias in the estimates were observed in scenarios considering one or two records available per individual, but this tended to be slightly less on estimates from GBLUP than with those estimates obtained using BLUP (Figure 1C). Since the frequency of variance component estimation giving non-zero estimates was lower for $n = 3$, all our comparison based on this pedigree structure were done assuming 3 records per individual.

Moreover, the observed trends were consistent across different genetic correlations between production potential and resilience and between BLUP and GBLUP (results not shown). In particular, as would be expected, prediction accuracies for both traits increase with increasing genetic correlations (as shown for GEBVs in Figure 2A): the average prediction accuracies of production potential and resilience

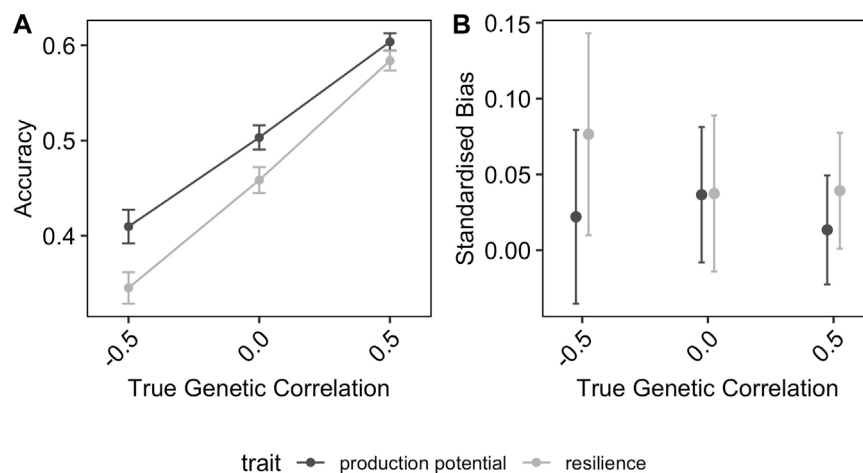


FIGURE 2

Effect of true genetic correlation between production potential and resilience on (A) the prediction accuracy of GEBVs and (B) the standardised bias in GEBV (positive values correspond to over-estimated GEBVs). The error-bars are the 95% confidence intervals.

increases from 0.41 to 0.60, and from 0.34 to 0.58, respectively. In addition, the resilience tends to be upward biased, and this intensifies when the genetic correlation is negative (Figure 2B).

For the remainder of this section (Random selection scenario), we will only show results for GEBVs assuming three records per individual and genetic correlation of -0.5 between production potential and resilience. This scenario was chosen to show how potential trade-offs between the traits may affect prediction accuracies, and because it has a low rate of failure to converge in the REML estimation ensuring that the results are true reflection of the situations compared and not to a lack of information due to inadequate population size. Furthermore, the observed effects of distribution of phenotypes across environments and uncertainty in the level of environmental challenge described below were similar for EBVs and GEBVs.

3.1.2 Effect of distribution of phenotypes across environments

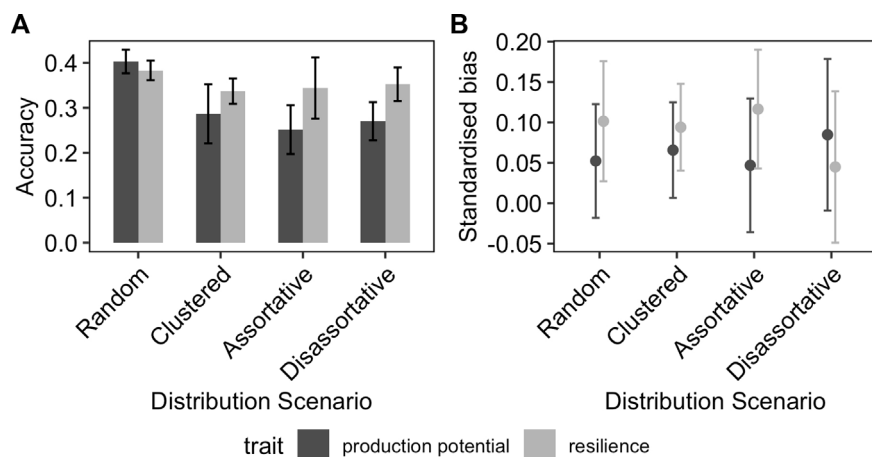
Figure 3 shows the effect of the distribution of phenotyped animals across environments with different average challenge levels on the accuracy and the bias of the GEBV for resilience and production potential. In general, the distribution of phenotypes across environments affected the quality of the genetic prediction of both traits but in different ways. The accuracy of the production potential was affected by the distribution of phenotypes across environments, with the highest accuracy (0.41) observed in the RND distribution scenarios and the lowest in AST and DIS scenarios (0.32). This was not the case for the resilience GEBVs, to which their accuracies were almost identical for the different allocation scenarios. In particular, even though the accuracy of production potential is generally higher than that of resilience, in the scenarios AST and DIS, resilience is estimated on average more accurately than the production potential. On the other hand, the distribution of phenotypes across different environments affected the resilience GEBV mainly, on the degree of bias, with the AST scenarios introducing the highest standardised bias of 15% (Figure 3B).

3.1.3 Effect of uncertainty in the level of environmental challenge

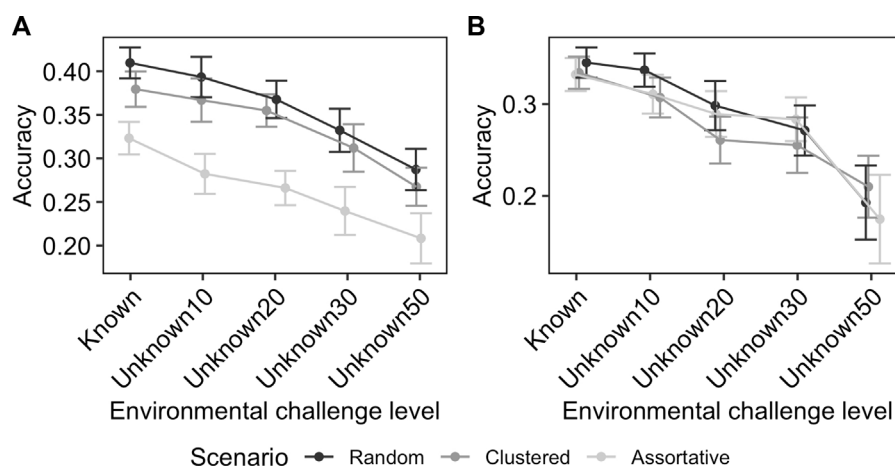
Figure 4 shows the effect of uncertainty in the knowledge of challenge level on the GEBV accuracies. In the ideal scenario, where the challenge levels are known without error, the GEBV accuracy was 0.41 for production potential and 0.35 for resilience. When the heterogeneity of flocks was 10%, as modelled here by flocks exposed to larger ranges of environmental challenge levels with average performance used as a proxy, the reduction on the GEBV accuracy was small (e.g., for RND scenario the average prediction accuracy for resilience and production potential changes from 0.44 to 0.49 to 0.46 and 0.50, respectively). However, as the level of uncertainty increased the GEBV accuracy further decreased, to the level that for the scenario Unkown50 (where each flock covered up-to 50% of full environmental challenge range) the GEBV accuracy was 0.28 and 0.21 for production potential and resilience, respectively. The rate of reduction in GEBV accuracy with increasing uncertainty was similar for different distribution scenarios (Figure 4).

3.2 Directional selection scenario

Figure 5 shows the gain in production potential and resilience after ten generations of selection when phenotypes are collected from a wide range of environmental challenge levels, spanning over six phenotypic standard deviations in average performance. The genetic gain is shown for different weights for resilience in the selection index. The response to selection for the two traits when the conventional model for performance (i.e., ignoring the environmental challenge and only fitting the flock as fixed effect) (Eq. 5) is used, is indicated with a point and is moved horizontally to lie on the selection index profile. Figure 5 shows a smooth transition of the genetic gain for resilience from negative to positive values as the index for the resilience is varied from 0 to 1, and the reverse is true for production potential. The maximum achievable relative

**FIGURE 3**

Effect of distribution of phenotyped animals on the (A) prediction accuracy and (B) standardised bias of GEBVs. The error-bars are the 95% confidence intervals.

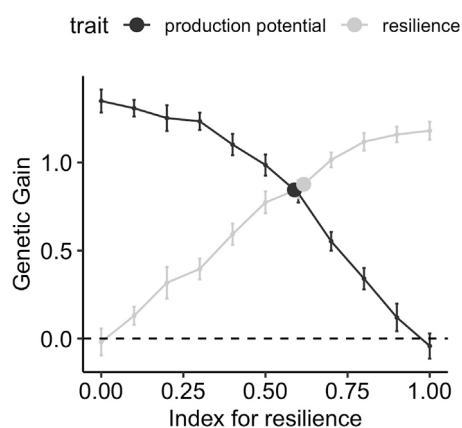
**FIGURE 4**

Effect of unknown environmental challenge in RN model on the accuracy of GEBV for production potential (A) and resilience (B). The numbers after the word "Unknown" on the x-axis indicate the heterogeneity of flocks in terms of percentile coverage of the entire range of environmental conditions. The effect is shown for different distribution scenarios with different colours. The error-bars indicate the 95% confidence intervals.

gains for production potential and resilience using corresponding indices are 1.4 and 1.3, respectively. This means that using an index with all weights on production potential, the average production potential is increased from 10 to 11.4 after 10 generations. Similarly, $\alpha = 1$ in the selection index increases the resilience by 43.3% from -3 to -1.7 after 10 generations (cf. Table 1). The genetic gain from the conventional model indicates that in such a scenario both resilience and production potential can be improved simultaneously with a relative gain of 0.78 for both traits, equivalent to an index with weights 0.58 and 0.42 for resilience and production potential, respectively.

The response to selection after 10 generations of selection depends strongly on the range of environments in which

phenotypes are collected. Figure 6 shows the results when phenotypes are collected in a narrow range of good or bad environmental conditions, corresponding to ranges of [0,1] and [1,2], respectively. In good environmental conditions (i.e., low challenge levels, Figure 6A), the maximum achievable gain for resilience (i.e., an index with all weights on resilience) is 0.7 whereas the maximum relative genetic gain in production potential (i.e., an index with all weight on production potential) is up to 1.6. Using a conventional model when phenotypes are mainly collected from a narrow range of environmental conditions with low challenge, only production potential is improved, whereas the relative genetic gain in resilience is zero. Conversely, collecting data mainly from bad environments (Figure 6B) results in higher

**FIGURE 5**

Genetic gain after 10 generations of selection using different strategies, when performance records are obtained from a wide range of environmental challenge levels ([0,2]). The lines correspond to using a selection index comprising production potential and resilience with different relative weights on resilience on the x-axis, and points correspond to selection on production performance GEBVs obtained using the conventional model (Eq. 5)

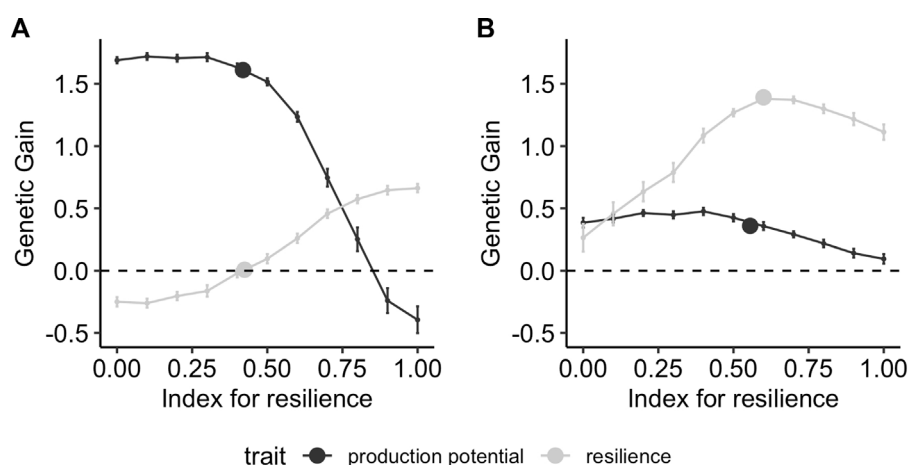
gains in resilience. The maximum available gain for production potential with a selection index reduces to 0.5 in this case (i.e., one-third of what is observed when phenotypes are collected in good environments), whereas that of resilience increases to 1.4 (i.e., twice as much as when phenotypes are from good environments). Similarly, the genetic gain that can be achieved when analysing phenotypes collected in bad environments with the conventional model results in improvement of resilience mainly, with a relative gain of 1.5 compared to the 0.35 gain for production potential. It is also noteworthy that the conventional model yields close to maximum genetic gain in resilience if phenotypic records are

collected under high challenge conditions, whereas it leads to zero improvement of resilience when only phenotypic records in low challenge levels are available. In the latter case, genetic gain in resilience can only be achieved using reaction-norm models.

4 Discussion

In this study, we demonstrated the benefit of genomic prediction on the accuracy and bias of the estimated breeding value of resilience and production potential traits for data simulated with RN models. We also showed that the number of phenotypic records, the distribution of phenotypes across different environments, and the degree of uncertainty on the challenge levels can affect GEBV accuracy and bias. However, even for sub-optimal distribution of phenotypes or highly heterogeneous and unknown challenge levels, accuracies are sufficiently high, and bias is sufficiently low to select simultaneously for both traits.

It has been widely reported in literature that the use of genomic information improves the accuracy of breeding values in genetic evaluation for different species (VanRaden et al., 2009; Riggio et al., 2014; Wolc et al., 2016; Tsai et al., 2016; Sánchez-Mayor et al., 2022). Evidence on the beneficial effect of genomic information when using a RN model is also available from real data. For example, Silva et al. (2014) reported that the GEBVs for total number of born in pigs at a given environment were about 70% more accurate than EBV accuracy obtained using pedigree information. Moreover, recently it has been shown that ssGBLUP implementing a RN model on production traits in pigs yields prediction which were 10–25% more accurate than a RN using pedigree alone (Fragomeni et al., 2016; Song et al., 2020). Whilst these studies estimated GEBVs for production potential and resilience, their criterion of comparison was based on the accuracy to predict performance at a given environment with a specific level of challenge. The emphasis in this study, however, was to evaluate the benefits of using genomic

**FIGURE 6**

Genetic gain after 10 generations of selection using different strategies on performance records obtained in narrow environmental conditions with mainly (A) low ($X \in [0,1]$) and (B) high ($X \in [1,2]$) challenge levels. The lines correspond to using a selection index comprising production potential and resilience, with different relative weights on resilience on the x-axis, and the points correspond to selection on production performance GEBVs obtained using the conventional model (Eq. 5).

information on prediction accuracies for the two traits, production potential and resilience, themselves.

We observed that the benefit of GBLUP in increasing GEBV accuracies was greater for resilience than for production potential, which is partly explained by the fact that the resilience trait has lower accuracy from the start. It is important to note that the beneficial effect of GBLUP over BLUP was the same regardless of the number of phenotypic records available per individual. Hence, one may extrapolate that similar benefits of using genomic information should be expected in other populations that may differ in size or the number of phenotypic records available.

Whilst the GEBV accuracy for production potential was generally greater than for resilience, production potential GEBVs were more sensitive to the distribution of animals across environments, compared to the resilience trait, whose accuracy remained barely unchanged across the different scenarios on how the phenotypes are distributed across the environments. It is important to note that for the extreme case of AST scenario, the production potential GEBVs were severely affected to the point that their accuracy was on average lower than that for resilience. It is known that accuracy of EBVs for production potential (estimated with pedigree) are lowest when animals are reared in clusters and the correlation between the exposed challenge level and the production level is negative (i.e., the AST scenario) (Calus et al., 2004). The most plausible explanation for the lower accuracy in the AST scenario may be due to genetic and environment factors being confounded and that genetic connectedness is reduced between herds (Foulley and Quaas, 1995). We observed that this trend remains true also for GEBVs. For a BLUP evaluation, information of EBV for an individual comes from its own performance plus that of closely related individuals, so clustering them in environments with similar challenge level would be expected to affect EBV accuracy (especially for AST scenario where the best families are reared in the best environment). However, for GBLUP evaluation, the GEBVs for a given individual are based on their own information and from all recorded individuals, related or unrelated; hence one may expect that GEBV accuracies being less sensitive to the distribution of phenotypes. However, our results show that this is not the case, and the use of genomic information would not overcome the negative impact of member of the same family being evaluated in similar environments as in CLS or AST. Note that although the GEBV accuracy for resilience is relatively unaffected by the system of recording performance across the environment, some degree of bias in the resilience estimate tend to appear as the scenario become more like the CLS or AST. Therefore, for estimating both the production potential and resilience, it is important that the phenotypes from related individuals are collected from different environmental conditions as was represented by the RND scenario in this study. Unfortunately, the pattern of distribution of phenotype across environments is heavily influenced by the production system. Populations where natural mating is common would lead to scenarios similar to AST or CLS, but the widespread use of artificial insemination may allow for a more mixing of families across environments. Unfortunately, artificial insemination has been proved to be challenging for some species like sheep (Carta et al., 2009). Additionally, scenarios with large use of artificial insemination may also lead to clustered distribution of phenotypes, where good farms with high management input may result in environments with low challenge but also having the resources to buy the semen from the best sires (hence resulting in AST situation).

The results from our study further highlight the need to promote as much as possible an un-clustered distribution scenario if resilience and production potential are to be accurately evaluated without bias. Based on our results, as shown in Figure 3, for breeding values estimated with GBLUP, the use of genomic evaluation would not counterbalance this detrimental clustering effect and the accuracies are lower when animals are clustered.

The RN models are the most common approach in disentangling the environmental effect from the genetic effect on the phenotype (Calus et al., 2004; Mulder, 2016). In many real-world scenarios, the environmental challenge is unknown or simply not measured. In order to fit random effects for the level of response to environment, i.e., the resilience, a proxy is estimated for the challenge level using the same data to be analysed. The drawback is that this approach requires to separate performance into discrete classes (e.g., contemporary groups), where all individuals in each class are assumed to be affected by the same challenge level, so the genetic evaluation is done assigning the same proxy value to all individuals in the class. We show that accuracy of GEBVs obtained from this approach are not compromised if the discrete classes are relatively homogeneous, i.e., cover a small range of environments. This is likely not to be the case, particularly for infectious diseases, where exposure to infectious pathogens is highly stochastic in nature (Bishop and Woolliams, 2010). The error in the assigned challenge level would be expected to affect the accuracy of the GEBVs. Furthermore, the magnitude of this detrimental effect is expected to be related to the within-group heterogeneity in the challenge level. Our results showed that the loss in accuracy due to uncertainty on the level of the challenge should be minimal, if classes to calculate the proxy are created such as that their heterogeneity within (defined as the range of possible environmental challenge of within the class, relative to the whole) is under 10%. However, although the large impact that uncertainty on the challenge level can have on the quality of the prediction, our results showed that relatively good genetic estimates can still be achieved even when these environments are very heterogeneous (e.g., 50%). Whilst the implementation of an RN evaluation should have extra care in the allocation of performance records in discrete classes, (to reduce environmental heterogeneity within class), we have shown that even when these classes are very heterogeneous, the GEBV can be of sufficient accuracy so their use can result in successful genetic progress in practical breeding programmes.

Over the recent decades, the animal breeding community has given much attention to breeding for resilience. However, it is desirable that by selecting for increased resilience, one does not accidentally select for poor performers when conditions are good. It has been postulated (Falconer and Mackay, 1996), and shown here, that selecting for animals that perform well in bad environments increases resilience when a conventional model is used without RN. However, this may result in little gain for the production potential if not a decrease when the genetic correlation between production potential and resilience is negative. Improving both traits simultaneously is easier when there is no trade-off between traits (Mulder, 2016). However, many studies report an antagonistic genetic correlation between resilience and production potential (Carabaño et al., 2017; Sánchez-Molano et al., 2020; Freitas et al., 2021). Additionally, it has been shown, using a more complex model, that the estimates of heritability are smaller when animals are

exposed to micro-environmental disturbances, $X_d > 0$, compared to an environment with no disturbance, $X_{wd} = 0$, (Le et al., 2022). This translates to heritability of performance at X_{db} i.e., $h^2(X_d)$, being smaller than heritability at X_{wdb} which is $h^2(X_{wd}) = h^2(0)$ in Equation 4, which occurs always for negative genetic correlation and small X_{wd} .

In this study, linear reaction models were used. Changes in performance of animals may however not always be linear with respect to challenge level. Therefore, non-linear reaction norms (e.g., quadratic) can predict phenotypes at specific environments more accurately Pollott and Greeff (2004); Sánchez-Molano et al. (2020). To improve resilience with such models a more complex selection index may be needed because all the coefficients from the model contribute to the sensitivity of an animal to a challenge and hence need to be incorporated in the selection index. Furthermore, in our study, we used a linear heterogeneous residual term and continuous challenge level. Heterogeneous residual variances can be modelled with exponential function (Hill and Mulder, 2010), as is the case in uniformity studies. While the definition of the heritability for resilience is more straightforward with our model, it is important that the conclusions drawn in this study are tested in other scenarios.

Lastly, in this paper, we defined resilience as the ability of an animal to maintain high production performance when exposed to challenge, in line with Bisset and Morris (1996). Resilience was simulated as a reaction norm, where one performance record for an animal corresponds to a given environmental challenge, X . This approach is valid when there is a relatively constant challenge level over a significant period of animals' life, e.g., prolonged nutritional shortage or heat stress, and when the effects on the animal phenotype of interest can be captured by a constant value, e.g., growth rate over a specific time period, or carcass weight. Several more recent studies, including the EU Horizon 2020 Smarter consortium (SMARTER, 2018) define resilience as the ability of an animal to either maintain or revert quickly to high production or health status when exposed to challenge or micro-environmental disturbances (Colditz and Hine, 2016; Berghof et al., 2019; Le et al., 2022). To estimate resilience according to this refined definition of resilience, longitudinal performance measures of individual animals would be required to capture, e.g., ability of the animal to return to its pre-challenge state after exposed to a challenge (Colditz and Hine, 2016; Berghof et al., 2019; Knap and Doeschl-Wilson, 2020). Thanks to the automated phenotype measurement devices that the animal industry is adopting, daily measurements of performances are becoming more available (Poppe et al., 2020; Neethirajan and Kemp, 2021). Recent studies have produced and evaluated novel resilience indicators from such daily measurements (Putz et al., 2019; Nguyen-Ba et al., 2020; Poppe et al., 2021). For example, Le et al. (2022) defined and used resilience and resistant terms in simulated longitudinal body weight of growing pigs and showed environmental disturbances affect estimates of breeding values. Future studies are warranted to assess the benefits of genomics and the influence of diverse factors as those assessed here, on these novel resilience indicators.

5 Conclusion

In this paper, we studied the benefits of using genomic information, random regression RN models and phenotyping strategies on genetic evaluations of resilience and production potential. We showed that simultaneous improvement of both traits is possible even under unknown environmental challenge

conditions if genomic information is used in a RN model and phenotyped animals are not in family clusters. In contrast, genetic improvement in resilience without the use of RN models is possible if genotyped and phenotyped animals are reared in wide range of environmental conditions. Therefore, for the best genetic evaluation, use of artificial insemination or other methods that increase connectedness among genotypes in various environmental conditions, is recommended. Additionally, higher between-variance for different environmental conditions facilitates simultaneous genetic improvement for resilience and production potential.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

RP-W, AD-W, OM, and VR contributed to conception and design of the study. MGZ performed the simulations and analyses. MGZ wrote the first draft of the manuscript. AD-W and RP-W wrote sections of the manuscript. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication. All authors contributed to the article and approved the submitted version.

Funding

This project is funded by the European Union's Horizon 2020 research and innovation program under the Grant Agreement No. 772787 (SMARTER). AD-W's and RP-W's contributions were partly funded by the BBSRC Institute Strategic Programme Grant ISP3: Improving animal production and welfare, Theme 1: Genetic improvement of farm animals BBS/E/D/30002275 and Theme 2: Complex phenotypes and GxE interactions BBS/E/D/30002276.

Acknowledgments

The first author would like to thank Fatemeh Ghaderi Zefreh for their help in deriving the mathematical expressions in the Supplementary Material S1.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their

affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Baumgard, L. H., Rhoads, R. P., Rhoads, M. L., Gabler, N. K., Ross, J. W., Keating, A. F., et al. (2012). "Impact of climate change on livestock production," in *Environmental stress and amelioration in livestock production* (Berlin, Heidelberg: Springer Berlin Heidelberg), 413–468. doi:10.1007/978-3-642-29205-7_15
- Berghof, T. V. L., Poppe, M., and Mulder, H. A. (2019). Opportunities to improve resilience in animal breeding programs. *Front. Genet.* 9, 692. doi:10.3389/fgene.2018.00692
- Bishop, S. C., and Woolliams, J. A. (2010). On the genetic interpretation of disease data. *PLoS ONE* 5, e8940. doi:10.1371/journal.pone.0008940
- Bisset, S., and Morris, C. (1996). Feasibility and implications of breeding sheep for resilience to nematode challenge. *Int. J. Parasitol.* 26, 857–868. doi:10.1016/S0020-7519(96)80056-7
- Bohlouli, M., Alijani, S., Naderi, S., Yin, T., and König, S. (2019). Prediction accuracies and genetic parameters for test-day traits from genomic and pedigree-based random regression models with or without heat stress interactions. *J. Dairy Sci.* 102, 488–502. doi:10.3168/jds.2018-15329
- Calus, M. P. L., Bijma, P., and Veerkamp, R. F. (2004). Effects of data structure on the estimation of covariance functions to describe genotype by environment interactions in a reaction norm model. *Genet. Sel. Evol.* 36, 489–507. doi:10.1186/1297-9686-36-5-489
- Calus, M. P. L., Groen, A. F., and de Jong, G. (2002). Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. *J. Dairy Sci.* 85, 3115–3123. doi:10.3168/jds.S0022-0302(02)74399-3
- Cameron, N. D. (1997). *Selection indices and prediction of genetic merit in animal breeding*. Oxford: CAB International.
- Carabaño, M. J., Ramón, M., Díaz, C., Molina, A., Pérez-Guzmán, M. D., and Serradilla, J. M. (2017). Breeding and genetics symposium: Breeding for resilience to heat stress effects in dairy ruminants. A comprehensive review. *J. Animal Sci.* 95, 1813–1826. doi:10.2527/jas.2016.1114
- Carta, A., Casu, S., and Salaris, S. (2009). Invited review: Current state of genetic improvement in dairy sheep. *J. Dairy Sci.* 92, 5814–5833. doi:10.3168/jds.2009-2479
- Carvalho, R., Costilla, R., Neves, H. H. R., Albuquerque, L. G., Moore, S., and Hayes, B. J. (2019). Unraveling genetic sensitivity of beef cattle to environmental variation under tropical conditions. *Genet. Sel. Evol.* 51, 29. doi:10.1186/s12711-019-0470-x
- Colditz, I. G., and Hine, B. C. (2016). Resilience in farm animals: Biology, management, breeding and implications for animal welfare. *Animal Prod. Sci.* 56, 1961–1983. doi:10.1071/AN15297
- de Jong, G. (1990). Quantitative Genetics of reaction norms. *J. Evol. Biol.* 3, 447–468. doi:10.1046/j.1420-9101.1990.3050447.x
- Falconer, D. S., and Mackay, T. F. C. (1996). *Introduction to quantitative genetics*. 4th edn. Harlow: Pearson.
- Foulley, J., and Quaas, R. (1995). Heterogeneous variances in Gaussian linear mixed models. *Genet. Sel. Evol.* 27, 211. doi:10.1186/1297-9686-27-3-211
- Fragomeni, B. O., Lourenco, D. A. L., Tsuruta, S., Bradford, H. L., Gray, K. A., Huang, Y., et al. (2016). Using single-step genomic best linear unbiased predictor to enhance the mitigation of seasonal losses due to heat stress in pigs. *J. Animal Sci.* 94, 5004–5013. doi:10.2527/jas.2016-0820
- Freitas, P. H. F., Johnson, J. S., Chen, S., Oliveira, H. R., Tiezzi, F., Lázaro, S. F., et al. (2021). Definition of environmental variables and critical periods to evaluate heat tolerance in large white pigs based on single-step genomic reaction norms. *Front. Genet.* 12, 717409. doi:10.3389/fgene.2021.717409
- Friggens, N., Duvaux-Ponter, C., Etienne, M., Mary-Huard, T., and Schmidely, P. (2016). Characterizing individual differences in animal responses to a nutritional challenge: Toward improved robustness measures. *J. Dairy Sci.* 99, 2704–2718. doi:10.3168/jds.2015-10162
- García-Baccino, C. A., Marie-Etancelin, C., Tortoreau, F., Marcon, D., Weisbecker, J.-L., and Legarra, A. (2021). Detection of unrecorded environmental challenges in high-frequency recorded traits, and genetic determinism of resilience to challenge, with an application on feed intake in lambs. *Genet. Sel. Evol.* 53, 4. doi:10.1186/s12711-020-00595-x
- Gumedze, F. N., and Dunne, T. T. (2011). Parameter estimation and inference in the linear mixed model. *Linear Algebra Its Appl.* 435, 1920–1944. doi:10.1016/j.laa.2011.04.015
- Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* 31, 423–447. doi:10.2307/2529430
- Hermesch, S., and Dominik, S. (2014). *Breeding focus 2014 - improving resilience (Animal genetics and breeding unit, UNE with support from Pork CRC)*.
- Hermesch, S., Li, L., Doeschl-Wilson, A. B., and Gilbert, H. (2015). Selection for productivity and robustness traits in pigs. *Animal Prod. Sci.* 55, 1437. doi:10.1071/AN15275
- Herrero-Medrano, J. M., Mathur, P. K., Napel, J. t., Rashidi, H., Alexandri, P., Knol, E. F., et al. (2015). Estimation of genetic parameters and breeding values across challenged environments to select for robust pigs. *J. Animal Sci.* 93, 1494–1502. doi:10.2527/jas.2014-8583
- Hill, W. G., and Mulder, H. A. (2010). Genetic analysis of environmental variation. *Genet. Res.* 92, 381–395. doi:10.1017/S0016672310000546
- Kenyon, P. R., Roca Fraga, F. J., Blumer, S., and Thompson, A. N. (2019). Triplet lambs and their dams – A review of current knowledge and management systems. *N. Z. J. Agric. Res.* 62, 399–437. doi:10.1080/00288233.2019.1616568
- Kijas, J. W., Porto-Neto, L., Dominik, S., Reverter, A., Bunch, R., McCulloch, R., et al. (2014). Linkage disequilibrium over short physical distances measured in sheep using a high-density SNP chip. *Anim. Genet.* 45, 754–757. doi:10.1111/age.12197
- Knap, P. W. (2005). Breeding robust pigs. *Aust. J. Exp. Agric.* 45, 763–773. doi:10.1071/EA05041
- Knap, P. W., and Doeschl-Wilson, A. (2020). Why breed disease-resilient livestock, and how? *Genet. Sel. Evol.* 52, 60. doi:10.1186/s12711-020-00580-4
- Knap, P. W., and Su, G. (2008). Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. *Animal* 2, 1742–1747. doi:10.1017/S1751731108003145
- Kolmodin, R., and Bijma, P. (2004). Response to mass selection when the genotype by environment interaction is modelled as a linear reaction norm. *Genet. Sel. Evol.* 36, 435–454. doi:10.1186/1297-9686-36-4-435
- Kolmodin, R., Strandberg, E., Madsen, P., Jensen, J., and Jorjani, H. (2002). Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. *Acta Agric. Scand. Sect. A - Animal Sci.* 52, 11–24. doi:10.1080/09064700252806380
- Le, V., Rohmer, T., and David, I. (2022). Impact of environmental disturbances on estimated genetic parameters and breeding values for growth traits in pigs. *Animal* 16, 100496. doi:10.1016/j.animal.2022.100496
- Lee, S. H., and van der Werf, J. H. (2006). An efficient variance component approach implementing an average information REML suitable for combined LD and linkage mapping with a general complex pedigree. *Genet. Sel. Evol.* 38, 25–43. doi:10.1186/1297-9686-38-1-25
- Li, L., and Hermesch, S. (2016). Environmental variation and breed sensitivity for growth rate and backfat depth in pigs. *Animal Prod. Sci.* 56, 61. doi:10.1071/AN14066
- Lipschutz-Powell, D., Woolliams, J. A., Bijma, P., Pong-Wong, R., Bermingham, M. L., and Doeschl-Wilson, A. B. (2012). Bias, accuracy, and impact of indirect genetic effects in infectious diseases. *Front. Genet.* 3, 215. doi:10.3389/fgene.2012.00215
- Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. doi:10.1093/genetics/157.4.1819
- Mulder, H. A. (2016). Genomic selection improves response to selection in resilience by exploiting genotype by environment interactions. *Front. Genet.* 7, 178. doi:10.3389/fgene.2016.00178
- Mulder, H. A., and Rashidi, H. (2017). Selection on resilience improves disease resistance and tolerance to infections. *J. Animal Sci.* 95, 3346–3358. doi:10.2527/jas.2017.1479
- Neethirajan, S., and Kemp, B. (2021). Digital phenotyping in livestock farming. *Animals* 11, 2009. doi:10.3390/ani11072009
- Nguyen, T. T., Bowman, P. J., Haile-Mariam, M., Pryce, J. E., and Hayes, B. J. (2016). Genomic selection for tolerance to heat stress in Australian dairy cattle. *J. Dairy Sci.* 99, 2849–2862. doi:10.3168/jds.2015-9685
- Nguyen-Ba, H., van Milgen, J., and Taghipoor, M. (2020). A procedure to quantify the feed intake response of growing pigs to perturbations. *Animal* 14, 253–260. doi:10.1017/S1751731119001976

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2023.1127530/full#supplementary-material>

- Oliveira, H., Brito, L., Silva, F., Lourenco, D., Jamrozik, J., and Schenkel, F. (2019). Genomic prediction of lactation curves for milk, fat, protein, and somatic cell score in Holstein cattle. *J. Dairy Sci.* 102, 452–463. doi:10.3168/jds.2018-15159
- Patterson, H. D., and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545–554. doi:10.1093/biomet/58.3.545
- Pollett, G. E., and Greeff, J. C. (2004). Genotype x environment interactions and genetic parameters for fecal egg count and production traits of Merino sheep. *J. Animal Sci.* 82, 2840–2851. doi:10.2527/2004.82102840x
- Poppe, M., Mulder, H., and Veerkamp, R. (2021). Validation of resilience indicators by estimating genetic correlations among daughter groups and with yield responses to a heat wave and disturbances at herd level. *J. Dairy Sci.* 104, 8094–8106. doi:10.3168/jds.2020-19817
- Poppe, M., Veerkamp, R., van Pelt, M., and Mulder, H. (2020). Exploration of variance, autocorrelation, and skewness of deviations from lactation curves as resilience indicators for breeding. *J. Dairy Sci.* 103, 1667–1684. doi:10.3168/jds.2019-17290
- Putz, A. M., Harding, J. C. S., Dyck, M. K., Fortin, F., Plastow, G. S., Dekkers, J. C. M., et al. (2019). Novel resilience phenotypes using feed intake data from a natural disease challenge model in wean-to-finish pigs. *Front. Genet.* 9 (660). doi:10.3389/fgene.2018.00660
- Rashidi, H., Mulder, H. A., Mathur, P., van Arendonk, J. A. M., and Knol, E. F. (2014). Variation among sows in response to porcine reproductive and respiratory syndrome. *J. Animal Sci.* 92, 95–105. doi:10.2527/jas.2013-6889
- Riggio, V., Abdel-Aziz, M., Matika, O., Moreno, C., Carta, A., and Bishop, S. (2014). Accuracy of genomic prediction within and across populations for nematode resistance and body weight traits in sheep. *Animal* 8, 520–528. doi:10.1017/S1751731114000081
- Sánchez-Molano, E., Kapsona, V. V., Ilska, J. J., Desire, S., Conington, J., Mucha, S., et al. (2019). Genetic analysis of novel phenotypes for farm animal resilience to weather variability. *BMC Genet.* 20, 84. doi:10.1186/s12863-019-0787-z
- Sánchez-Molano, E., Kapsona, V. V., Oikonomou, S., McLaren, A., Lambe, N., Conington, J., et al. (2020). Breeding strategies for animal resilience to weather variation in meat sheep. *BMC Genet.* 21, 116. doi:10.1186/s12863-020-00924-5
- Sánchez-Mayor, M., Riggio, V., Navarro, P., Gutiérrez-Gil, B., Haley, C. S., De la Fuente, L. F., et al. (2022). Effect of genotyping strategies on the sustained benefit of single-step genomic BLUP over multiple generations. *Genet. Sel. Evol.* 54, 23. doi:10.1186/s12711-022-00712-y
- Silva, F. F., Mulder, H. A., Knol, E. F., Lopes, M. S., Guimarães, S. E. F., Lopes, P. S., et al. (2014). Sire evaluation for total number born in pigs using a genomic reaction norms approach. *J. Animal Sci.* 92, 3825–3834. doi:10.2527/jas.2013-6486
- SMARTER (2018). *Small ruminants breeding for efficiency and resilience*.
- Song, H., Zhang, Q., Misztal, I., and Ding, X. (2020). Genomic prediction of growth traits for pigs in the presence of genotype by environment interactions using single-step genomic reaction norm model. *J. Animal Breed. Genet.* 137, 523–534. doi:10.1111/jbg.12499
- Strandberg, E. (2006). “Analysis of genotype by environment interaction using random regression models,” in Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13–18, 2006.25.05
- Thornton, P. K., Herrero, M. T., Freeman, H. A., Okeyo Mwai, A., Rege, J., Jones, P. G., et al. (2007). Vulnerability, climate change and livestock-opportunities and challenges for the poor. *J. Semi-Arid Trop. Agric. Res.*
- Tomley, F. M., and Shirley, M. W. (2009). Livestock infectious diseases and zoonoses. *Philosophical Trans. R. Soc. B Biol. Sci.* 364, 2637–2642. doi:10.1098/rstb.2009.0133
- Tsai, H.-Y., Hamilton, A., Tinch, A. E., Guy, D. R., Bron, J. E., Taggart, J. B., et al. (2016). Genomic prediction of host resistance to sea lice in farmed Atlantic salmon populations. *Genet. Sel. Evol.* 48, 47. doi:10.1186/s12711-016-0226-9
- VanRaden, P. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi:10.3168/jds.2007-0980
- VanRaden, P., Van Tassell, C., Wiggans, G., Sonstegard, T., Schnabel, R., Taylor, J., et al. (2009). Invited review: Reliability of genomic predictions for north American holstein bulls. *J. Dairy Sci.* 92, 16–24. doi:10.3168/jds.2008-1514
- Wientjes, Y. C. J., Bijma, P., Vandenplas, J., and Calus, M. P. L. (2017). Multi-population genomic relationships for estimating current genetic variances within and genetic correlations between populations. *Genetics* 207, 503–515. doi:10.1534/genetics.117.300152
- Wolc, A., Kranis, A., Arango, J., Settar, P., Fulton, J., O’Sullivan, N., et al. (2016). Implementation of genomic selection in the poultry industry. *Anim. Front.* 6, 23–31. doi:10.2527/af.2016-0004