A Computational Model to Understand the Response of Methane Emission to Genomic and Microbial Selection in Cattle

Nateela Shoukat

Independent project in Animal Science, EXO870 • 30 hp
Swedish University of Agricultural Sciences, SLU
Faculty of Veterinary Medicine and Animal Science/ Department of Animal Breeding and Genetics
European Master’s in Animal Breeding and Genetics/ Master’s programme in Animal Science
Uppsala, Sweden, 2025

EUROPEAN MASTER IN ANIMAL BREEDING AND GENETICS

A Computational Model to Understand the Response of Methane Emission to Genomic and Microbial Selection in Cattle

Nateela Shoukat

09/2024

Main supervisor: Martin Johnsson (Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences)

Co-supervisor: Armin Schmitt (Division of Breeding Informatics, Georg-August- Universität Göttingen)

Co-supervisor: Jan Ingemar Ohlsson, SLU, Department of Animal Breeding and Genetics, Uppsala, Sweden

Co-funded by the
Erasmus+ Programme
of the European Union

A Computational Model to Understand the Response of Methane Emission to Genomic and Microbial Selection in Cattle

En beräkningsmodell för att förstå svaret av metanutsläpp på genomiskt och mikrobiellt urval hos nötkreatur

Nateela Shoukat

Supervisor:	**Martin Johnsson, PhD, SLU, Department of Animal Breeding and Genetics, Uppsala, Sweden**
Assistant supervisor:	Jan Ingemar Ohlsson, SLU, Department of Animal Breeding and Genetics, Uppsala, Sweden
Assistant supervisor:	Armin Schmitt, PhD, Georg-August-Universität Göttingen, Division of Breeding Informatics
Examiner:	Tomas Klingström, PhD, SLU, Department of Animal Breeding and Genetics, Uppsala, Sweden

Credits:	30 credits
Level:	Second cycle, A2E
Course title:	Independent project in animal science
Course code:	EXO870
Programme/education:	European Master in Animal Breeding and Genetics/ Master's programme in Animal Science
Course coordinating dept:	Department of Animal Breeding and Genetics
Place of publication:	Uppsala, Sweden
Year of publication:	2025
Keywords:	Rumen Microbiome, AlphaSim R, Heritability, Microbiability, Additive Variance, Genomic Selection, Methane Emission

Swedish University of Agricultural Sciences
Faculty of Veterinary Medicine and Animal Science Department
Department of Animal Breeding and Genetics

Abstract

Methane emissions from cattle, influenced by the interaction between their genetics and the rumen microbiome, pose a significant environmental challenge. This study developed a computational model to investigate how genetic and microbial factors affect methane production. Utilizing AlphaSimR, we simulated methane traits under different additive genetic variances and microbial methane effects. We employed Bayesian models to predict methane emissions and to assess heritability and microbiability across various scenarios.

The results indicated that lower microbial methane production effects consistently led to reduced emissions, and heritability values increased with increasing genetic variance. These simulated results emphasize the host genetics' stronger impact on high additive variance, while microbiability values decreased under these conditions. The model achieved high prediction accuracy, ranging from 85% to 91%, and demonstrated that combining genomic and microbial selection effectively lowered methane emissions over ten generations.

These findings provide valuable insights for breeding programs to minimize methane emissions and address environmental sustainability in livestock production.

Keywords: Rumen Microbiome, AlphaSim R, Heritability, Microbiability, Additive Variance, Genomic Selection, Methane Emission

Table of contents

List of tables

List of figures

1. Introduction

Enteric methane, excreted by livestock, significantly contributes to greenhouse gas emissions and causes global warming. Methane emission in ruminants is a complex phenomenon involving host genotype, rumen microbiome, and animal feed (Mahala et al. 2022). In 2017, the Animal Production and Health Division of the Food and Agriculture Organization (FAO) reported that cattle are responsible for about two-thirds of livestock methane emissions, mainly due to enteric fermentation, accounting for roughly 30% of the world's total methane emissions (AGA 2017). Methane is the second most prevalent greenhouse gas after carbon dioxide, and it is more effective at trapping heat (Milich 1999). The level of methane is rising by about 1% each year (Heilig 1994), contributing to climate change, such as rising sea levels, higher temperatures, unpredictable rainfall, heat waves, and severe droughts. Therefore, finding ways to reduce methane emissions is crucial.

Rumen microbes produce methane as a byproduct of digestion in ruminant animals, leading to an energy loss that impacts both the animal's efficiency and the environment. Methanogens, i.e., organisms emitting methane such as *Methanobrevibacter ruminantium*, which is dominant in the cattle rumen, *Methanobacterium formicicum*, and *Methanosarcina barkeri* play a key role in maintaining efficient fermentation in the rumen but result in an energy loss of 2–12% as methane emissions (Khobragade et al. 2024). Broucek also highlights in his study that the methane generated through the microbial fermentation of hydrolyzed carbohydrates represents an energy loss for the host, with the extent of this loss influenced by factors such as energy consumption, feed quality, rumen microbes, growth rate, host genetics, and the environmental conditions (Broucek 2014).

Several researchers hypothesized that methane emission variation in cows is determined by animal genomics and the rumen microbiome (Difford et al. 2018; Wallace et al. 2019). Difford and colleagues also hypothesized that the rumen microbiome composition is influenced by host genetics. Cattle additive genetics impact only 12% of archaeal OTU (Operational taxonomic unit) abundances, which means that cow genetics play a role in determining the composition and abundance of these microbes (Difford et al. 2018).

It is also important to understand how specific microbial populations contribute to methane emissions in livestock as it can help to develop strategies to reduce livestock methane production. Ramayo-Caldas and colleagues identified 86 hydrogen-producing bacteria in cattle, which accounts for 24% of methane variation, as well as the host's genetics contributing around 14%, which explains the interaction between host genetics and rumen microbiome. This gives an insight into how microbial markers can identify high- or low-emitting cattle (Ramayo-Caldas et al. 2020). Moreover, the genetic markers in cattle accounted for 7-12% of the variation in the abundance of bacteria and archaea, including methanogens like *Methanobrevibacter* and *Methanosphaera*, which showed a strong genetic influence (Zhang et al. 2020).

Understanding the interaction between host genome and microbiome information will improve the ability to predict selection for a trait, such as methane emission from cattle. However, identifying the relationship between the host genome and microbiome is challenging due to highly variable microbial data, estimating microbiability, which refers to the trait variance due to microbes, and the complex interaction between microbiome, phenotype, and genotype (Pérez-Enciso et al. 2021). Pérez-Enciso and colleagues overcame these challenges with a novel simulation strategy based on real microbiome and genotype data, and they used Bayesian RKHS and Bayes C to estimate variance components by modeling the contributions of genetic, microbial, and interaction effects. Their results suggest that combining microbiome data with genomic data could increase the prediction accuracy of the phenotype by about 50%, but there are difficulties regarding collecting consistent microbiome data over time and identifying specific causative microbes (Pérez-Enciso et al. 2021).

To better understand how the rumen microbiome influences methane emissions, it is essential first to describe some key genetic terms. From the perspective of quantitative genetics, conventional heritability describes the proportion of phenotypic variance, which is explained by the additive genetic effect of a population, and microbiome heritability is a variation in microbial abundance that is attributable to the host's genetics (Opstal & Bordenstein 2015). The effect of the microbiome on a trait can be quantified by estimating the microbiability, which is the proportion of phenotypic variance of the trait that is explained by between-animal differences in the microbial community (Aliakbari et al. 2022).

The aims of this study are to:

1. Develop a simulation model to evaluate the effectiveness of selecting cattle for reduced methane emissions, a trait influenced by the abundance of different rumen microbes.
2. Measure the accuracy of predictions, heritability, and microbiability values depending on different levels of additive genetic variance for methane emissions and the microbial methane effect.

3. Simulate the long-term response to selection for methane emissions over multiple generations for different scenarios of additive variance and microbial methane effect.

To achieve these objectives, AlphaSim R will be utilized to simulate a host population and model methane emissions as influenced by both host genetics and microbial factors. The methane trait will be modelled with a heritability of 0.4, while the Operational Taxonomic Units (OTUs) will have varying heritabilities specified for each OTU set within the simulation. For example, some OTUs may receive higher heritabilities, such as 0.725 for certain groups, while others may have lower heritabilities, such as 0.087.

A microbial matrix will represent the abundance of OTUs across individuals. This microbial matrix will be normalized to ensure that the relative abundance of each OTU is accurately accounted for in every individual. The normalized microbial data will then be combined with transformed phenotypic data that reflect the genetic and environmental effects on microbial composition.

The microbial methane effect will be simulated under three different variance scenarios: low, medium, and high. These scenarios will represent different levels of microbial influence on methane emissions. Different levels of genetic variance will also be applied to the methane trait to capture varying degrees of genetic influence on methane emissions. Following this, the methane trait will be simulated for these levels of additive variance under different microbial methane effect scenarios.

The BGLR method will then be employed to estimate genomic prediction. Subsequently, the response to selection will be simulated across several generations by selecting predicted phenotypes from the Bayesian method in offspring. Individuals with the most favourable methane phenotypes in each generation will be selected to form the next generation. This process will allow for the long-term tracking of the response to selection.

After running this simulation, we expect to gain insights into the influence of host genetics and the rumen microbiome on methane emissions. The analysis of varying genetic variances for the methane trait and the three microbial methane effect scenarios (low, medium, and high variance) will enhance our understanding of how genetic and microbial factors drive methane emissions and effect heritability and microbiability values, which ultimately can help design a breeding program to reduce methane emissions in cattle.

2. Literature Review

2.1 Methane production in cattle

Enteric methane is a colorless and odorless gas produced as a byproduct of fermentation in the digestive tract of ruminants. As described by Murray and colleagues, 87% of methane is produced in the rumen, and the remaining 13% is produced in the hindgut during digestion. Most of the methane is released through burping, and the remaining is generated in the large intestine and is mostly evacuated through the lungs and also a smaller amount through the anus (Murray et al. 1976).

No single microbial species is responsible for the complete degradation of the substances in the rumen. In contrast, a complex plethora of microorganisms participates in complete substrate catabolism into the end products of fermentation.

2.2 Rumen microbial community

The rumen microbiota significantly affects the ruminants' health and digestive efficiency. All three domains of life, which include Bacteria, Archaea, and Eukaryotes, are present in the rumen (Woese et al. 1990). The rumen microbiome is primarily composed of bacteria, which are considered the most stable species across various ruminant species. They play a symbiotic role by providing essential metabolic functions for the host protozoa, which vary more widely in presence and abundance, and archaea are the key players in contributing to methane production (Wallace et al. 2019).

All known methanogens belong to the archaea kingdom. Methanogens are classified into seven orders: *Methanobacteriales*, *Methanococcales*, *Methanomicrobiales*, *Methanosarcinales*, *Methanocellales*, *Methanopyrales*, and *Methanomassiliicoccales* (Vanwonterghem et al. 2016). Methanogenic microbes belong to the phylum Euryarchaeota, which includes diverse groups of genera. The most dominant genus is *Methanobrevibacter*, which includes *M. ruminantium*, *M. smithii*, and *M. gottschalkii*. Another genus, *Methanosphaera*, includes species like *M. stadtmanae* and *M. cuniculi* (Aryee et al. 2023). The genus *Methanomassiliicoccus* and *Methanospirillum* are also found within the order Methanomicrobiales in the rumen. This includes species such as *M. hungatei* and *M. lacunae*, which contribute to the community's methane production (Aryee et al. 2023). In conclusion, these Euryarchaeotic methanogens form a diverse ruminal ecosystem contributing to methane emissions.

2.3 Heritability of methane emission

In animal breeding, this genetic parameter predicts the response to selection, representing the difference in phenotypic means between a population and its offspring. The values of heritability range from 0 to 1; the higher the value, the higher the proportion of the variation in phenotype is explained by genetic variance, which means faster response in the targeted trait of interest (Oldenbroek & Waaij 2014). Methane emission is a heritable trait. One of the estimates of methane emission heritability is 0.21 with a standard error of 0.06 obtained by Lassen and Løvendahl from Holstein dairy cows (Lassen & Løvendahl 2016). Similar heritability estimates like 0.23 were also reported by (Manzanilla-Pech et al. 2016) and 0.19±0.09 by (Difford et al. 2018). Kamalanathan and colleagues found methane emission heritability estimates of 0.16 for methane production (MeP), 0.27 for methane yield (MeY), and 0.21 for methane intensity (MeI) (Kamalanathan et al. 2023). In another study, the heritability estimates for methane emissions in cattle ranged from 0.18 to 0.33, indicating a moderate genetic influence on methane traits (Van Breukelen et al. 2023). The average estimates of heritability for methane emissions ranged from 0.05 to 0.45, as reviewed by (Lassen & Difford 2020).

Methanogenic archaea produce methane, as described by Hook and colleagues (Hook et al. 2010). Methane production is a heritable trait, so the composition of methanogenic archaea and associated microbiomes regulating methane emissions is also heritable.

2.4 Microbiome heritability

One of the studies by Weimer and colleagues found that a cow's ruminal bacterial community and physiological traits, like pH and volatile fatty acid concentration, quickly restored even after the total exchange of rumen content, showing strong host specificity (Weimer et al. 2010). In another study, the host can return the rumen composition to its original state even after approximately 95% rumen content exchange between cows (Weimer et al. 2017). Furthermore, Roehe and colleagues have found that the cow's genetics can influence the rumen microbe and the amount of methane produced by rumen microbes, suggesting that the host's genetics play a key role in regulating the trait (Roehe et al. 2016). Moreover, Difford and colleagues found heritabilities for the relative abundance of rumen bacteria and rumen archaea, reaching 0.4 and 0.3, respectively (Difford et al. 2018). Several studies reported high heritability estimates for ruminal microorganisms; for instance, (Sasson et al. 2017) reported heritability estimates larger than 0.7 for 22 bacterial Operational Taxonomic Units (OTU) in dairy cattle. A study done in 2020 used a genome-wide association study to estimate the heritability of the rumen microbiome in cattle and identified that maximum heritability estimates for OTU, family, and phylum levels were 0.820, 0.722, and 0.722, respectively (Abbas et al. 2020). They also found that the taxa belonging to phyla Euryarchaeota and family *Methanobacteriaceae* had heritability estimates greater than 0.5 (Abbas et al. 2020). In a recent study, the heritability estimates for the microbiome's functional capacity, including methane metabolism and ribosomal synthesis, range from 0.13 to 0.61, and the heritability of the relative abundance of individual rumen microbes ranges from 0.08 to 0.48 (Worku 2024).

2.5 Host genetics and microbial community

A recent study has also emphasized the important connection between the composition of the bacteria in the rumen and the amount of methane released, finding that certain types of microbes, such as *Succinivibrionaceae* and *Methanosphaera*, contribute as much as 24% to the variation in methane production, while host genetics account for about 14% of the variation (Ramayo-Caldas et al. 2020). Using a linear mixed model approach, Difford and colleagues found that the combined microbial abundance and additive genetic effects are responsible for almost 34% of the total phenotypic variation in CH_4 emissions (Difford et al. 2018). A study by Saborío-Montero and colleagues focused on the role of certain methanogenic microorganisms like *Methanobrevibacter* and *Methanosphaera*, which contribute significantly to methane production, accounting for up to 40% of the variation between individual cattle. These researchers found a strong genetic link between the rumen's host and microbial community, with *Methanobrevibacter* and *Methanosphaera* showing positive correlations with methane emissions (Saborío-Montero et al. 2021).

2.6 Microbiability

Microbiability is a useful quantitative measurement that quantifies the proportion of phenotypic variance in a trait attributed to differences in the microbial community between animals. It was first proposed by Ross and colleagues using a microbial relationship matrix from meta-genomics sequencing data and predicting the prediction accuracy of 0.466 for the methane emission phenotype of cows (Ross et al. 2013). In another study, the proportion of variation in methane emission by rumen microbiome in cows is estimated at 0.15 ± 0.08 (estimate ± S.E) using a mixed model approach (Difford et al. 2018).

2.7 Bayesian Generalized Linear Regression

This type of regression is common in genetic studies where phenotypes are regressed on many predictors. Regressing such traits, which are influenced by many predictor variables, poses challenges such as the high dimensionality of predictors and interactions between genetic and environmental factors (Pérez & De Los Campos 2014). Bayesian methods are particularly effective in this domain, combining shrinkage, variable selection, and flexible modeling to achieve reliable predictions and inference (de los Campos et al. 2013). As Pérez and De Los Campos explained, the Bayesian framework provides a consistent and unified approach to these challenges, offering a range of parametric and semi-parametric models (Pérez & De Los Campos 2014). Parametric methods assume a predefined model like a normal distribution, semiparametric methods combine fixed models with flexible components based on data, and nonparametric methods impose no assumptions about the data distribution; each requires different levels of prior knowledge (Howard et al. 2014).

The BGLR package, developed by Paulino Pérez and Gustavo de los Campos, is an essential tool for genomic prediction that implements various Bayesian regression models (Pérez & De Los Campos 2014). One is Bayesian Ridge Regression (BRR), which shrinks coefficients uniformly and makes it well-suited for studies where all predictors contribute equally. Another model, Bayesian LASSO (BL), employs double-exponential priors that enable size-dependent shrinkage, allowing more significant effects to stand out while controlling smaller ones. Furthermore, Reproducing Kernel Hilbert Spaces (RKHS) Regression extends this framework to semi-parametric modeling, using Gaussian processes to capture non-linear relationships. Moreover, Perez and Paulino described that BayesA uses scaled-t priors to induce more significant shrinkage on more minor effects while accommodating markers with larger variances, including hyperparameters such as degrees of freedom and scale, which uses Gamma distributions to balance flexibility and precision (Perez n.d.). In contrast, BayesC introduces a mixture model that combines a point mass at zero for null effects with a Gaussian slab for non-null effects, effectively incorporating variable selection, including parameters like the proportion of non-zero effects modelled with a Beta distribution and variance parameters to refine the estimation process (Perez n.d.).

By integrating these methods into a unified framework, the BGLR package provides researchers with a versatile and efficient platform for genomic analysis (Pérez & De Los Campos 2014).

2.8 Genomic selection for methane emission

To predict methane emissions, many researchers have focused on integrating rumen microbial and genomic data to select cows that produce less methane. For example, a study by Saborio-Montero and colleagues employed a novel Principal Component Analysis (PCA) approach to consolidate the diverse microbial community into simplified phenotypic variables. These variables exhibited high heritability and significant genetic correlations with methane emissions, representing a promising strategy for targeting the rumen metagenome and reducing methane production through genetic selection (Saborío-Montero et al. 2021).

Another study by Roehe and colleagues in 2016 used the sire progeny groups to demonstrate that the host's genetics influence methane emissions and have significant differences based on microbial abundance and emissions across groups. Their findings suggest that the relative abundance of rumen microbes can predict traits like methane production and feed efficiency, which targets breeding strategies to reduce emissions while improving livestock productivity (Roehe et al. 2016).

In the paper by Pérez-Enciso et al., the authors integrated microbial and genomic data into simulation models to predict complex traits. The authors simulated scenarios where microbial abundances either independently affected traits or were partially controlled by the host genome. They also discussed the challenges they faced, including the complexity of genome-microbiome interactions, biases in estimating microbiability, and difficulties in identifying specific microbial taxa influencing traits. However, incorporating microbial data improved the accuracy of trait predictions despite these hurdles (Pérez-Enciso et al. 2021).

2.9 Simulation models in breeding programs

Simulation models test different breeding strategies without the time and cost of real-world experiments, and based on the predicted outcomes of these simulations, breeders can optimize their breeding programs. A study by Hassanpour and colleagues demonstrated the effectiveness of stochastic simulations in identifying optimal breeding strategies for dairy cattle to optimize breeding programs by simulating multiple scenarios for high productivity and genetic diversity (Hassanpour et al. 2023). Their study applies stochastic simulation to optimize cattle breeding, which aligns with my research focus. While we use different tools, namely MoBPS in their study and AlphaSim R in mine, both approaches rely on stochastic modeling to evaluate breeding programs.

AlphaSim R emerged as a flexible tool widely used for simulating plant and animal breeding programs. Its user-friendly design also allows for integrating other R packages for advanced analysis (Gaynor et al. 2021). It uses stochastic simulations to model complex breeding strategies. Stochastic simulations utilize randomness to model variability in biological and breeding processes, making them valuable for evaluating genetic gain and selection strategies under uncertainty. Deterministic simulations, on the other hand, are based on complex formulas to give results without standard errors and are usually fast. AlphaSim R allows users to simulate a wide range of breeding operations, including crossing, selection, and genomic prediction, with functionality for detailed modeling of traits influenced by additive, dominance, and epistatic effects (Gaynor et al. 2021).

2.10 Integrating microbial data into the simulation model

To understand how host genetics interact with the microbiome with a complex trait like methane emission in cattle, incorporating microbial data into the simulation model can enhance this analysis. By adding microbial abundance, specifically OTUs as traits in AlphaSim R, the simulation can effectively capture the intricate relationship between host genetics and the rumen microbiome. Utilizing literature data on microbial OTU heritabilities (Difford et al. 2018), their abundances (Tapio et al. 2017). Their connections to methane emissions will help define the simulation’s parameters. Integrating microbial data into AlphaSim R simulations presents a novel approach to understanding and mitigating methane emissions through genomic and microbial selection in cattle.

3. Methodology

This study develops a computational framework to analyse the response of methane emissions in cattle to genomic and microbial factors. The methodology incorporates genomic simulations by using AlphaSim R, microbial modeling, and predictive Bayesian modeling to assess heritability and microbiability values under varying microbial methane effect and genetic variance scenarios and then the response of selection over generations. The whole process was conducted in several steps as described below:

3.1 Generation of founder genome

The first step in the simulation study is to create a founder population that would serve as the genetic basis for subsequent generations. This was achieved using the runMacs function in AlphaSimR; 500 individuals (founders) were simulated across 30 chromosomes. Each chromosome contained 10,000 segregating sites, ensuring adequate genetic diversity for later analyses. The sexes of the individuals in the founder population were assigned using the "yes_sys" condition to maintain a 1:1 male-to-female ratio. A simulation parameter framework was established to manage key settings, including genetic variation, trait inheritance, and selection processes. This ensured consistency across generations and structured data for genomic selection. A SNP chip dataset with 1,000 SNPs per chromosome was added into the simulation and was utilized in subsequent genomic selection procedures.

3.2 Parameters

Quantitative Trait Loci (QTLs) are assigned to govern specific traits in the population. In AlphaSim R, QTLs are randomly distributed across the genome. The effects of the QTLs are initially sampled from a standard normal distribution and then scaled to achieve the specified genetic variance for the trait in the founder population (Gaynor et al. 2021). Methane emission, a key trait under study, is controlled by 96 QTLs distributed across the cattle genome, according to the cattle QTL database. These QTLs are classified under the trait class ‘production’ and the trait type ‘energy efficiency’ related explicitly to methane emissions (https://www.animalgenome.org/QTLdb/cattle).

Similarly, 30 Operational Taxonomic Units (OTUs) are modelled, each controlled by 10 QTLs, which shows that microbial abundance in the rumen is also influenced by host genetics. The choice of 30 OTUs specifically related to methane emission and belonging to *Euroarcheaota* is according to Abbas and colleagues (Abbas et al. 2020). Furthermore, specific heritability values are assigned to OTUs, varying from high to low contributions, following the same study by Abbas and Colleagues to simulate the heterogeneous effects of microbial taxa on methane production (Abbas et al. 2020).

3.3 Simulation of the traits

Methane emission and OTU (Operational Taxonomic Units) were simulated as traits. The methane emissions trait was modelled with a heritability of 0.19, which will vary according to the different values of additive variance and the OTU traits represented microbial contributions with varying heritabilities, such that 31 traits were simulated overall. The additive genetic variance for the methane emission trait is set at 0.5, 1, and 2 to study how the difference in genetic variance influences the methane heritability. The additive genetic variance for the OTU trait is set as 0.25 to include the host genetic effect on microbial abundances.

For the OTU traits, 30 OTUs were simulated, each controlled by 10 QTLs. The heritabilities of the OTUs were predefined based on a previous study: 20 OTUs exhibited moderate to low heritability, ranging from $h^2 = 0.145$ to $h^2 = 0.086$, while the remaining 10 OTUs showed high heritability ($h^2 = 0.725$) (Difford et al. 2018).

3.4 Generation of microbial data

3.4.1 OTU abundance matrix

The relative abundances of 30 OTUs represent microbial contributions to methane emissions. Tapio and colleagues calculated the relative abundances of Archae under different diets, and the average of these values was taken across all the diets for the OTU abundance matrix (Tapio et al. 2017). These values are then normalized by the scaling process, in which row normalization ensures that the sum of the values in each row equals 1. This makes the data comparable across individuals and reflects their proportional contributions.

3.4.2 Integration of host genetics and environmental factors

To integrate the host genetic and environmental effects of the host into the OTU matrix by transforming phenotypic data derived from the simulated population. The phenotype matrix from the population object, which contains quantitative traits for each individual, such as methane emission or microbial abundances, was extracted. Next, the transformation 2^{Pheno} applies an exponential scaling to the phenotypic values. This operation ensures that higher phenotypes, indicative of stronger genetic influences, are exponentially emphasized. This transformed phenotype matrix is then multiplied with the OTU abundance matrix to produce a matrix that incorporates both host genetic and environmental effects on OTU abundance, followed by normalization, where each row of the matrix is divided by its sum to ensure that microbial abundances remain proportionate across individuals.

3.5 Methane emission modelling with both microbial and host effects

The phenotype (Y), which is methane emission, is modelled as a combination of genetic, microbial, and environmental effects using the following equation:

$$Y = G + M + E$$

Y is methane emission, G is the host genetic effect on methane, M is the microbial effect on methane, and E is the environmental effect. The microbial impact on the trait is further decomposed as:

$$M = (G_m + \ E_m\ ) * y_m$$

G_m is the host genetic influence on microbial OTU abundance. Em is the environmental influence on OTU abundance, which affects methane, and y_m is the microbial effect on methane emission.

G is the host genetic effect on methane emission taken from the pop object in AlphaSim R. These values represent the observed phenotype, which is methane emission influenced by host genetics. However, to model the methane effect of microbes, the gamma distribution is applied using parameters used by Pérez-Enciso and colleagues (Pérez-Enciso et al. 2021). The gamma-distributed values of (y_m)

simulate the varying influence of microbial taxa on methane emissions. The normalized OTU matrix with host genetic and environmental effect is multiplied by y_m, with each OTU column adjusted proportionally to its gamma value, resulting in the new matrix called OTU_y_m. The row sums of this matrix give the microbial effects M, which quantify each individual's cumulative microbial contribution to methane emissions. Next, the environmental effect E is incorporated as random values drawn from a normal distribution with a mean of 0 and a standard deviation of 1. Finally, after incorporating all the effects, the following equation is applied:

$$Y = G + \{(G_m + E_m) * y_m\} + E$$

The methane effect (Y), the heritability (h^2), and microbiability (b^2) were calculated to quantify the respective contributions of host genetics and microbial influences on the total phenotypic variance. Heritability (h^2) was calculated as:

$$h^2 = \frac{var_G}{var_Y}$$

Microbiability (b^2) was calculated as:

$$b^2 = \frac{var_M}{var_Y}$$

Var_G represents genetic variance, Var_M denotes microbial variance, and Var_Y signifies methane effect variance, which is the sum of Var_G, Var_M, and Var_E.

3.6 BayesC and BayesA models for predicting methane emissions based on host genetics and microbial contributions

Two separate models were employed to predict methane emissions using BayesC and BayesA. The response variable, Y, representing the total methane effect calculated earlier, was used as the dependent variable. Genetic contributions were encoded in a matrix G, representing SNP genotype data for all individuals. The normalized OTU matrix represented microbial contributions, capturing the combined genetic and environmental effects on microbial abundance. Two predictive models were constructed within the BGLR framework. The genetic model utilized BayesC, tailored for SNP-based genetic variance, while the microbial model employed BayesA, designed to handle gamma-distributed microbial effects with specified shape and rate parameters. These models were defined in the ETA list, an argument in the function that specifies the structure of the model by outlining how genetic and microbial factors are incorporated and assigning appropriate priors to each effect, which defines the structure of the model. The ETA list facilitated the inclusion of both genetic and microbial sources of variation, enabling the fitting of the models to predict methane emissions.

3.7 Prediction accuracy

Predicted phenotypes are compared to observed values to assess the model's accuracy. The correlation between predicted and observed methane emissions is a key metric for evaluating the model's accuracy in predicting methane emission levels based on host genetic and microbial factors.

3.8 Breeding program simulation

3.8.1 Selection criteria

The predicted phenotypes, which are the predicted methane effect calculated from the Bayes C and Bayes A prediction model, are incorporated into the EBV s column to use them in the selection process. Breeding simulations begin with selecting individuals based on their predicted phenotypes, guiding the selection of sires (25 males) and dams (250 females) for the next generation. Regarding the selection of fewer sires than dames, it is important to note that in livestock species, a single male can successfully mate with multiple females. This makes it biologically and practically efficient to have a lower number of sires compared to dams.

3.8.2 Generational progression

The breeding program is simulated over ten generations. In each generation, offspring are produced through random mating between selected sires and dams. Using updated genetic and microbial data from running the model and BGLR for every subsequent generation, the phenotypes of these offspring are predicted. This process ensures that breeding decisions are informed by the latest genetic data, promoting improved traits in future offspring.

3.8.3 Simulation of additive variance and microbial methane effect scenarios

Finally, a loop structure is added to run nine scenarios for three levels of additive variance: 0.50, 1, and 2; and three gamma distribution scenarios for microbial methane effects named low, medium, and high, to represent different microbial effects. For each variance value, simulations are performed for all three gamma distributions. The 'Low' scenario has a skewed distribution, with most values concentrated at lower levels, while the 'Medium' and 'High' scenarios exhibit a distinct shift to more even and higher distributions. This helps us explore how varying microbial contributions impact methane effects, heritability values, microbiability values, and the selection of response.

4. Results

4.1 Methane emission at varying scenarios

Simulations show that host genetic variance and the magnitude of the causal effect of microbial composition both affect methane emissions predicted by the model. Methane emissions were evaluated under three levels of methane additive variance (0.5, 1, 2) and three gamma distribution scenarios for microbial methane effects (low, medium, and high), representing different microbial impacts. These scenarios help identify key conditions associated with low methane emissions. The density distributions shown in Figure 1 highlight how varying scenarios of genetic additive variance and microbial effects influence methane output.

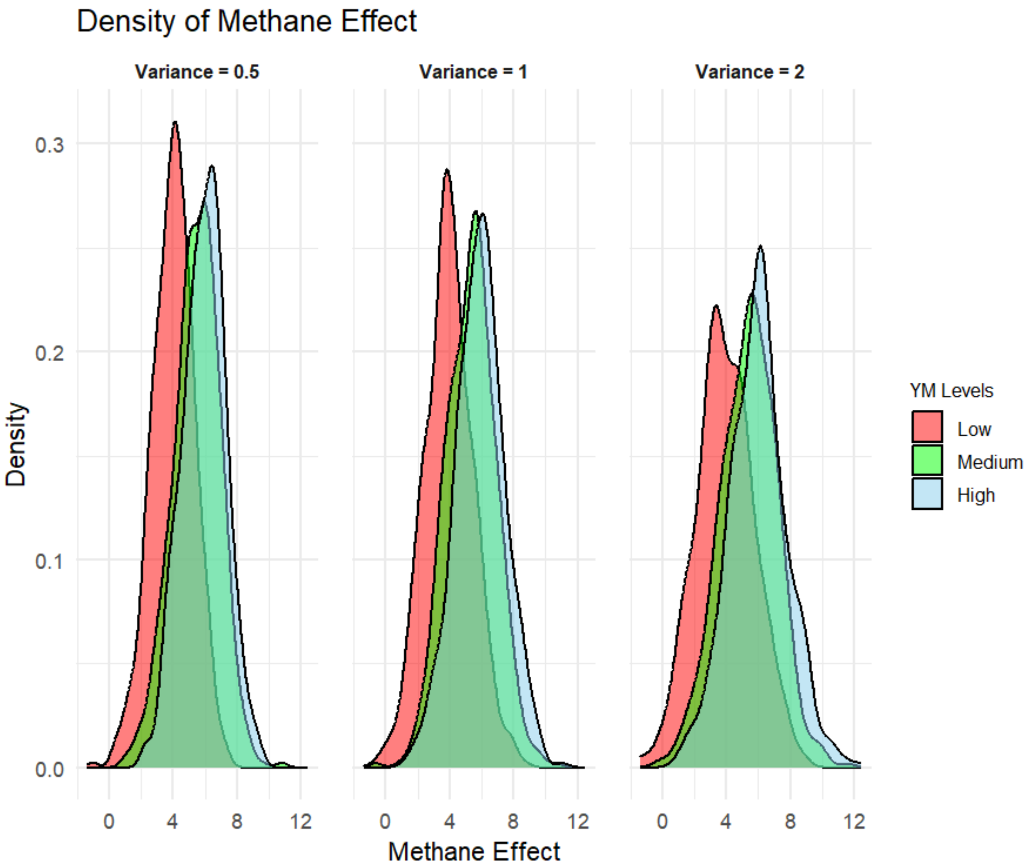

Figure 1. "Density" presents the density distributions of methane effects under different scenarios, categorized by the additive variance (var) and microbial methane effect levels y_m. The x-axis represents the methane effect, while the y-axis shows the density of observations. The plot consists of three density distributions corresponding to var values of 05, 1, and 2, displayed from left to right.

At the low value of additive variance 0.5 in the low y_m scenario, the methane effect density curve is left-skewed, reaching peaks at lower values compared to medium and high scenarios (y_m). In the medium and the high y_m scenarios, there are slightly broader distributions, and their peaks also show higher values than those in the low y_m scenario, which shows moderate increases in methane effects, indicating that emissions begin to rise with increasing y_m.

With an additive variance of 1, the low y_m scenario continues to produce the lowest methane emissions. Its density distribution peak is lower than at an additive variance of 0.5, likely due to the increasing genetic influence. In contrast, the medium and high y_m scenarios show slightly higher density distributions and peak at greater methane effect values. However, the overall methane effect in the low y_m scenario remains minimal.

At an additive variance of 2, the low ym scenario’s distribution decreases even further, maintaining a low methane effect peak. Meanwhile, the medium and high y_m scenarios continue to broaden in distribution and peak at higher methane effect values than the low y_m scenario.

In conclusion, scenarios with low methane emissions are associated with the low y_m scenario across all the additive variance levels, and the density distribution shows a downward trend as the y_m scenarios and the additive variance increase. The medium and high y_m scenarios shift their peak of methane effect to higher values as y_m increases and decreases in density as the additive variance increases.

4.2 Heritability and microbiability values at varying scenarios

The results highlight how genetic and microbial factors influence methane emissions across different scenarios. Tables 1, 2, and 3 summarize the heritability and microbiability values observed for methane emissions across nine scenarios, incorporating three values for additive variance and then for each value, three levels of microbial effect on methane. After running the methane emission model that accounts for microbial and host effects in each scenario, the calculated values are presented in Table 1.

Table 1. Heritability and Microbiability values at 0.5, 1, and 2 additive variances.

Additive variance Scenario	y_m scenario	Heritability	Microbiability
0.50	Low	0.313	0.103
0.50	Medium	0.247	0.212
0.50	High	0.257	0.228
1	Low	0.454	0.072
1	Medium	0.360	0.174
1	High	0.427	0.164
2	Low	0.677	0.051
2	Medium	0.634	0.133
2	High	0.571	0.123

The results in Table 1 show low heritability values observed in the scenario with additive variance set at 0.50, ranging from 0.247 to 0.313. As the microbial effect increases from low to high, microbiability rises noticeably from 0.103 to 0.228. There is only a slight difference in the values of heritability and microbiability between the medium and high y_m scenarios. However, both values contrast more than the low y_m scenario. The low y_m scenario has higher heritability values and lower microbiability values than the other scenarios.

In scenarios where the additive variance is set at 1, both heritability and microbiability values follow a trend similar to when the additive variance is set at 0.5. Specifically, heritability values increase from 0.360 to 0.454, while microbiability values rise from 0.070 to 0.174 as the y_m effect increases. Additionally, heritability values tend to rise with a higher additive variance, while microbiability values decrease when compared to those observed with an additive variance of 0.5. This relationship suggests that microbiability tends to be lower when heritability is higher. This may be due to the y_m values being drawn from distributions with fixed ranges, even as the overall variance of the methane trait expands.

Furthermore, with the additive variance set at 2, both heritability and microbiability values increase, rising from 0.571 to 0.677 and from 0.051 to 0.133, respectively. Heritability values at an additive variance of 2 are much higher than in other scenarios. In contrast, microbiability values decrease substantially as the additive variance increases. This trend suggests that the methane trait is now more influenced by host genetics rather than the rumen microbiome.

Overall, heritability values increase as the additive variance rises, with the lowest values at 0.50 and the highest at 2. These results also show that high heritability and low microbiability values are observed at low microbial methane effect conditions regardless of the additive variance value. Additionally, as microbial influences strengthen, microbiability values rise, highlighting the microbiome's key role in phenotypic variance.

4.3 Prediction accuracy

To evaluate the model's accuracy in predicting methane emissions based on both host genetics and microbial factors, we calculated the correlation coefficients between the methane effects (which include host genetic and microbial influences) and the predicted methane values from the BGLR model across various scenarios, as presented in Table 2.

Table 2. Correlation coefficient values for different scenarios

Additive variance	Microbial methane effect	Correlation coefficient
0.5	Low	0.909
0.5	Medium	0.847
0.5	High	0.783
1	Low	0.839
1	Medium	0.872
1	High	0.806
2	Low	0.964
2	Medium	0.934
2	High	0.910

Overall, the correlation values are high, ranging from 0.783 to 0.964, which indicates that the model is effective in predicting methane emissions in most cases. The accuracy tends to be higher when the additive variance is significant, especially under conditions of low microbial methane effects. Conversely, the correlation is weaker when microbial methane effects are high or the additive variance is low. This suggests that the complexity of microbial interactions may slightly reduce prediction accuracy in these situations.

4.4 Response of selection across generations for predicted phenotypes

The response of selection based on predicted phenotypes over ten generations was simulated. Each subsequent generation is selected based on the predicted phenotypes of the previous generation, which reflect the anticipated methane effects. For each scenario, the overall mean of the ten generations was calculated, and the results for the response of selection are shown in Figure 2.

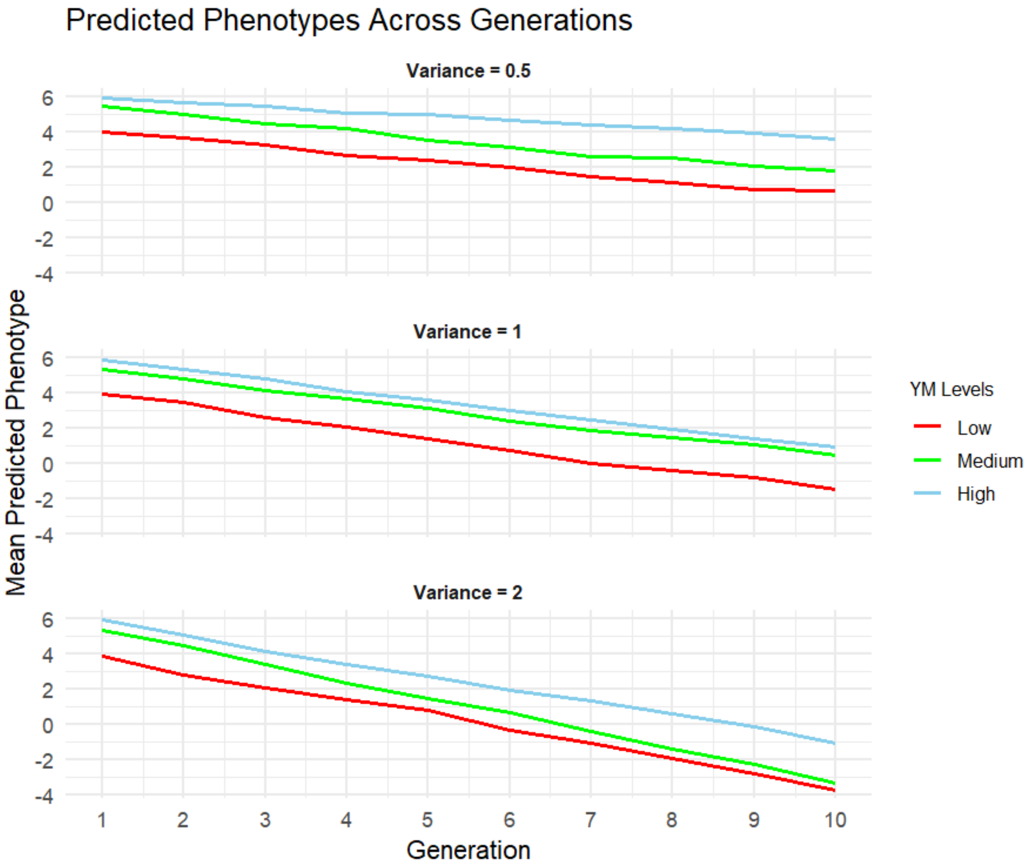

Figure 2. This plot shows the mean predicted phenotypes across 10 generations for three values of additive variance (0.5, 1, 2) each under three scenarios for microbial methane effect y_m (low, medium, high). The x-axis represents the generation, while the y-axis shows the mean of predicted phenotypes across these generations. The plot consists of three panels corresponding to var values of 0.50, 1, and 2, displayed from top to bottom.

In the top panel, where the variance is set at 0.5, the predicted phenotypes show a gradual decline over the 10 generations. The rate of decline varies depending on the microbial methane effect (y_m) scenarios: phenotypes decrease most rapidly under the "low" y_m scenario, moderately under the "medium" y_m scenario, and least rapidly under the "high" y_m scenario. The differences between the scenarios remain relatively small, indicating limited variation when variance is low.

In the middle panel, where the genetic variance is set at 1, the decline in predicted phenotypes is steeper than in the top panel, which means that as variance increases, the rate of reduction in predicted phenotypes also intensifies. The separation between the three y_m scenarios becomes more evident, showing more significant variation as generations progress. However, the y_m scenarios show the same trend as with variance 0.5. However, there is a more intensified decrease in mean predicted phenotypes, with the "low" y_m scenario exhibiting the steepest decline and the "high" y_m scenario showing the slowest decline.

The bottom panel, with a variance of 2, demonstrates the most pronounced changes. The decline in predicted phenotypes is steepest across all y_m scenarios. This panel highlights that greater variance allows more extreme candidates to be selected, intensifying this decline in predicted phenotypes. The "low" y_m scenario continues to show the fastest decline. In contrast, the "high" y_m scenario shows the slowest, with the "medium" y_m scenario in between the low and the high y_m scenarios.

Overall, the predicted phenotypes decrease over 10 generations, with the rate of decline becoming steeper as variance increases. Higher variance enhances the strength of selection, leading to faster reductions in phenotypes. Notably, some phenotypic values drop below zero in later generations, especially when the additive variance value is high. Despite this, the results demonstrate that higher genetic variance and the low y_m scenario accelerate the decline in predicted phenotypes, while higher y_m effects slightly slow it down.

5. Discussion

Using a computational model, this study investigated the response of methane emissions in cattle to genomic and microbial selection. The results demonstrate that host genetics and the rumen microbiome influence methane emissions. This study examined how different levels of additive variance (0.5, 1, 2) and varying microbial methane effects (y_m: low, medium, high) influenced methane emissions in cattle. At low additive variance (0.5), the low y_m scenario resulted in the least methane emissions. As additive variance increased to 1 and then 2, the microbial methane effect was still at the level of reduced methane emissions. Heritability and microbiability values varied significantly across scenarios. At 0.5 additive variance, heritability ranged from 0.247 to 0.313, while microbiability increased from 0.103 (low y_m) to 0.228 (high y_m). As additive variance goes from 1 to 2, heritability increased to a range of 0.360 to 0.677, while microbiability declined. This indicates a stronger genetic influence on methane emissions as additive variance increases alongside the reduction in microbiability, which indicates less influence from microbial communities. Prediction accuracy remained high, with correlation coefficients between the methane effect from the methane model, which accounts for both host genetics and microbial effect, and the methane-predicted values from the BGLR ranging from 0.85 to 0.91. Over ten generations, predicted phenotypes decreased with the steepest decline under low y_m and the most gentle decline under high y_m across all additive variance levels but a more pronounced decline at high additive variance levels.

Overall, the simulation results showed the significance of maintaining low microbial methane effects to reduce methane emissions, especially under higher additive variance conditions. Cattle breeding programs could mitigate livestock's environmental impact through genetic and microbial management strategies.

5.1 The role of microbial methane effect (y_m) in reducing methane emission

This study's findings highlight microbial methane's impact on determining methane emissions in cattle. The term "microbial methane effect" refers to the influence of methanogens in the rumen on total methane emissions. In this study, we simulated y_m using a gamma distribution to represent three levels of microbial impact: low, medium, and high. The results consistently showed that lower microbial effects corresponded to reduced methane emissions, regardless of the levels of additive genetic variance.

Notably, scenarios with high additive genetic variance (variance = 2) demonstrated a strong influence of host genetics on methane production. However, this high genetic variance is unlikely to occur naturally and was included to explore the upper limits of genetic control over methane emissions. The scenario with high genetic variance is simulated and does not represent real-world averages. However, they provide valuable theoretical insights into how host genetics can dominate methane traits in extreme cases. This relationship suggests that targeting microbial methane effects with genetic strategies could effectively mitigate emissions.

Research efforts have demonstrated potential in directly addressing methane effects via microbiome management. Feed additives, such as 3-nitrooxypropanol (3-NOP), have been shown to suppress methanogenic pathways, significantly reducing methane emissions (Hristov et al. 2015). Dietary modifications, such as using oils, tannins, or fiber-digesting enzymes, have also been explored to reduce the activity or population of methanogens in the rumen (Beauchemin et al. 2008). These interventions demonstrate how manipulating the rumen microbiome can reduce methane emissions in practical livestock systems.

Combining microbiome management with genetic selection could achieve even greater reductions. For example, breeding programs targeting cattle with traits associated with lower methane production, alongside strategies that optimize rumen microbial communities, offer a dual approach to tackling emissions. For example, Difford and colleagues highlighted the potential of genomic selection with microbiome-informed strategies, such as investigating microbial taxa with low methanogenic activity (Difford et al. 2018). In 2019, Wallace and colleagues further emphasized that host-determined, heritable microbes contribute significantly to methane emissions, supporting the concept of microbiome-led breeding programs (Wallace et al. 2019).

In conclusion, effective methane emissions management requires a balanced strategy. This strategy should integrate genomic selection with microbiome management techniques. It is important to select cattle with traits associated with lower methane emissions while implementing interventions such as feed additives, probiotics, or dietary adjustments targeting methanogens. Additionally, these strategies must not negatively impact other important traits, such as fertility, productivity, and overall animal health, which are essential for sustainable breeding programs (Lassen & Difford 2020).

5.2 Heritability and microbiability of methane emissions

Methane emissions are affected by both genetic factors (heritability) and microbial contributions (microbiability). Understanding the relationship between these components is crucial for creating effective breeding strategies. In this study, heritability values for methane emissions ranged from 0.247 to 0.677, with higher values associated with increasing the value of additive variance. This indicates that host genetics play a more significant role under these conditions. These findings are consistent with studies by (Lassen & Løvendahl 2016) and (Difford et al. 2018), which reported heritability estimates between 0.16 and 0.45 for methane emissions in dairy cattle.

Microbiability showed an inverse relationship with additive variance, ranging from 0.051 to 0.228. Higher microbiability values occurred at lower additive variance, suggesting that microbial effects are stronger when genetic contributions are weaker. In high heritability scenarios, host genetics mainly influence methane traits, although microbial contributions still affect methane emissions.

The interaction between heritability and microbiability highlights the necessity of a balanced breeding strategy incorporating genomic and microbial data. Higher values for heritability mean a quicker response to selection. While the heritability values increase, microbiability values decrease, so it is important to understand the balance between them to have an effective and balanced breeding program. This indicates the need to combine genomic and microbial data in a balanced way to improve breeding selection strategies. For instance, targeting specific microbial taxa with lower methanogenic activity, such as *Methanobrevibacter* species, can enhance reductions when combined with genomic selection (Wallace et al. 2019; Abbas et al. 2020).

The maintenance of this balance in real-world conditions relies on the complex interactions among diet, environment, and the rumen microbiome, with host genetics being significant at higher heritability levels, while microbiability also plays a relevant role. For example, (Roehe et al. 2016) showed that host genetics and environmental factors affect the abundance of key rumen microbes, indicating that microbiability can still play a significant role even in scenarios with high heritability.

5.3 Long-term response to selection

The simulation results showed that the response to selection reveals critical insights into how additive genetic variance and microbial methane effects (y_m) influence methane emissions over successive generations. The predicted phenotypes for methane emissions decline with successive generations, and the rate of decline is faster under scenarios of higher additive variance and lower y_m. However, negatively predicted phenotype values in later generations are also observed, especially when additive variance is high. This might be due to the additive genetic model used in the simulation, which simulates without explicit biological constraints. The simulation tool (AlphaSim R) assumes that additive genetic and microbial effects contribute linearly to the phenotypic values, and it does not impose a lower limit or biological threshold for methane emissions. As selection increasingly favours individuals with lower methane emissions, the predicted values could drop below zero, which is biologically unrealistic as methane emissions cannot be negative. This situation exemplifies the intense selection pressure from high additive variance and low microbial effects. The increased genetic variance enhances the chances of identifying extreme individuals within the population while leading to reduced predicted phenotypes. In addition, low microbial methane effects further empower the selection's capability to lower methane emissions, generating a synergistic effect that drives phenotypes to implausibly low levels.

To address this, the model could include constraints that set a minimum threshold for methane production based on biological requirements. This threshold could be derived from experimental data on the lowest observed emissions in low-methane cattle, ensuring that predictions remain realistic. Additionally, incorporating non-linear models that better capture the diminishing returns of selecting for extremely low methane emissions would align the simulations with biological processes.

Integrating more comprehensive datasets, such as methane measurements from diverse breeds, environments, and diets, would help fine-tune these thresholds. Physiological insights, like the minimum methanogenic activity necessary for rumen function, could also provide a scientific basis for defining these lower limits. These adjustments would make the simulations more reflective of real-world scenarios while preserving their ability to explore long-term selection impacts.

Despite this limitation, the trends observed in response to selection demonstrate that combining genetic variance with low y_m scenarios is a highly effective approach for mitigating methane emissions in cattle. In terms of genomic selection strategies, the Bayesian model used in this study effectively combined SNP markers and microbial abundances, providing accurate predictions. However, other models could also be explored. For example, linear mixed models or machine learning approaches could offer alternative ways to analyse data and predict methane emissions and are particularly useful for capturing non-linear interactions between SNP effects and microbial composition or for dealing with complex real-world datasets (Howard et al. 2014; Pérez-Enciso et al. 2021).

The outcomes might differ significantly if selection were based solely on microbial composition or SNP effects. Selecting only on microbial composition could result in quicker but potentially less sustainable reductions in methane emissions, as microbial communities are dynamic and influenced by external factors such as diet and environment (Difford et al. 2018). Conversely, selecting only on SNP effects would likely yield slower but more stable long-term improvements, as genetic changes are permanent and less affected by environmental variability. Comparing these strategies in future studies would provide valuable insights into their relative efficiency and trade-offs.

5.4 Comparative analysis with a recent study

The results of this study can be compared with those of the study by Cristina Casto-Rebollo and colleagues (Casto-Rebollo et al. 2024).They used the HoloSim R simulation framework. They also emphasized the significance of host genetics and the use of microbiota in influencing phenotypic traits. Both studies highlight the contribution of additive genetic variance and microbial effects, particularly in contexts that explore the microbiota's effect on the phenotypic response to selection. However, these studies differ in scope and methodology. This simulation study focuses specifically on methane emissions and studies how varying levels of y_m and additive genetic variance influence heritability, microbiability, and methane emissions.

In comparison, Casto-Rebollo and colleagues simulated different scenarios of microbial heritability. They also included the effects of the maternal and housing environment and the symbiosis between microbial species to see how host genetics and microbial interactions shape the overall microbial composition and affect the phenotype, and they explored the coevolution of genomes and microbiomes under different selection pressures (Casto-Rebollo et al. 2024).

In this simulation study, the negative predicted phenotypic values, particularly evident under high additive variance, reflect a limitation of additive genetic models that lack biological constraints. Extreme selection pressure can push phenotypes

into unrealistic ranges, such as negative methane emissions, which are not biologically plausible. They simulated phenotypic values by considering symbiosis effects and environmental microbiota contributions (Casto-Rebollo et al. 2024). This highlights the importance of integrating biologically realistic parameters into simulation models to enhance their applicability to real-world breeding programs.
However, both studies highlight the importance of microbial contributions to phenotypic variance.

5.5 Limitations of the simulation study

This study showed how the rumen microbiome and host genetics influence methane emissions, but several limitations should be considered. One significant limitation is the simplified modeling of microbial effects, where the y_m parameter serves as a scalar for microbial contributions. This simulation study does not consider any interactions among microbial species, such as symbiosis or competition, which ultimately affect microbial communities. Incorporating these interactions in future models could lead to a better understanding of the microbiome's role in methane emissions.

This simulation study may not entirely reflect real-world variability in microbial and genetic influence on the phenotype. Many other factors can affect the phenotype, such as environmental effects, maternal effects on microbial communities, vertical transmission of the microbiome, changes in diet, management practices, and temporal changes in the microbiome. All these effects could enhance the model's applicability. For example, accounting for shifts in the rumen microbiome during early life stages or dietary transitions could improve the accuracy of assessed microbial contributions to methane emissions (Meale et al. 2021).

The additive variance-based genetic framework also has its constraints as it does not account for non-additive genetic effects such as epistasis or dominance, which may also play a role in complex traits like methane emissions. Including these effects in future modeling efforts could provide a more comprehensive perspective on host genetics.

Host genetic variation significantly influences the rumen microbiome, including the abundance of specific microbial taxa associated with methane production, suggesting the presence of gene-by-microbe interactions in cattle (Roehe et al. 2016). This implies that non-additive effects could shape microbial communities in ways that additive models do not fully capture. Similarly, microbial composition may exhibit non-additive patterns, as interactions between microbial species can influence the overall community structure and its contributions to methane production (Martínez-Álvaro et al. 2020).

Incorporating non-additive effects into genomic selection in the context of methane emissions could provide a more complete picture of how host genetics and microbial factors work together. For example, certain SNPs might only impact methane traits in the presence of specific microbial taxa, or vice versa, creating epistatic or synergistic interactions between genes and microbes. Accounting for these interactions could improve the predictive power of selection models and help identify cattle with optimal combinations of genetic and microbial traits for methane reduction.

As we know, microbial communities are dynamic and change with different diets and environmental conditions (Hristov et al. 2015; Smith et al. 2022). There may also be a difference in microbial communities in different breeds of cattle, and there may be a maternal effect that can influence the early microbial environment of cattle and can affect their abundances (Meale et al. 2021)This study is generalized, so prediction reliability needs to be improved; future studies can focus on the dynamics of microbial communities under different conditions.

Additionally, this study focused on OTUs known to contribute to methane emissions. In a real-world scenario, many more OTUs would be measured, including those with little or no impact on the trait. This added complexity could dilute the predictive power of microbial-based selection models. Advanced feature selection methods, such as LASSO regression or dimensionality reduction techniques like Principal Component Analysis (PCA), could help identify the most relevant OTUs and improve prediction accuracy in real-world datasets (Howard et al. 2014; Pérez & De Los Campos 2014). Including these methods in the selection process might make the model more robust and better suited for practical applications. Addressing these limitations can further enhance the model's utility for breeding programs aimed at sustainability and climate change mitigation.

5.6 Future Research and Implications for breeding practices

As this study focused on methane emissions, the same framework could be extended to other important traits, such as feed efficiency and disease resistance, which are also influenced by genetics and the microbiome. The model used in this study provides valuable insights into how host genetics and microbial interactions shape traits. Future research could study microbial taxa's specific in shaping these traits. For instance, low-producing methanogens help identify microbes that enhance feed efficiency, guiding the development of probiotics or dietary supplements to reduce methane emissions and improve livestock productivity.

Integrating microbial data into selection decisions in breeding practices could significantly reduce methane emissions. This study helps align the cattle breeding program with climate goals. However, making this practical requires advancements in microbiome sampling and involvement in breeding evaluations. Developing simple, cost-effective tools for field use could bridge the gap between research and real-world applications like measuring methane emissions or sampling rumen microbiome communities. Developing cost-effective tools for field use could make microbial data integration more practical for breeding programs. Instead of expensive metagenomic sequencing, targeted approaches like 16S rRNA sequencing can reduce costs while identifying key microbial taxa. Portable biosensors that detect microbial metabolites or marker genes associated with methane production are another promising alternative.

It is also important to ensure that this focus does not negatively affect other important traits, such as production, fertility, or other health traits. Ethical issues, such as genetic diversity or intensive selection for reduced methane emission, may arise. This could make populations more vulnerable to disease outbreaks and environmental changes. Thus, it is important to maintain a balance between reducing methane and preserving the population's overall genetic health and diversity.

Additionally, we need to support and incentivize small-scale farmers, who often emphasize production traits over environmental considerations, to adopt these practices. One important ethical issue is that not everyone may have the same access to advanced breeding tools. For example, smaller farms or those with limited resources might struggle to utilize these technologies. Future policies must tackle these obstacles by promoting inclusivity and ensuring that new tools are accessible and affordable for all farmers, regardless of their size or resources.

Future research could consider the challenges and ethical implications before transitioning this approach from theory to practice, benefiting the industry and society.

6. Conclusion

Using a computational model, this study explored how cattle genetics and the rumen microbiome influence methane emissions. By simulating different genetic and microbial scenarios, we found that minimizing microbial methane effects is key to reducing emissions, especially in populations with higher genetic variance. The study also demonstrated that genomic and microbial data improves prediction accuracy and can effectively help design breeding programs over multiple generations. While the model is highly accurate, it does not fully account for real-world complexities, like environmental changes or microbial dynamics, which future research should address. This simulation study provides valuable insights for developing sustainable cattle breeding strategies to mitigate the environmental impact of livestock.

References

Abbas, W., Howard, J.T., Paz, H.A., Hales, K.E., Wells, J.E., Kuehn, L.A., Erickson, G.E., Spangler, M.L. & Fernando, S.C. (2020). Influence of host genetics in shaping the rumen bacterial community in beef cattle. *Scientific Reports*, 10 (1), 15101. https://doi.org/10.1038/s41598-020-72011-9

AGA (2017). *Livestock solutions for climate change*. FAO ; https://openknowledge.fao.org/handle/20.500.14283/i8098en [2024-09-12]

Aliakbari, A., Zemb, O., Cauquil, L., Barilly, C., Billon, Y. & Gilbert, H. (2022). Microbiability and microbiome-wide association analyses of feed efficiency and performance traits in pigs. *Genetics Selection Evolution*, 54 (1), 29. https://doi.org/10.1186/s12711-022-00717-7

Aryee, G., Luecke, S.M., Dahlen, C.R., Swanson, K.C. & Amat, S. (2023). Holistic View and Novel Perspective on Ruminal and Extra-Gastrointestinal Methanogens in Cattle. *Microorganisms*, 11 (11), 2746. https://doi.org/10.3390/microorganisms11112746

Beauchemin, K.A., Kreuzer, M., O'Mara, F. & McAllister, T.A. (2008). Nutritional management for enteric methane abatement: a review. *Australian Journal of Experimental Agriculture*, 48 (2), 21–27. https://doi.org/10.1071/EA07199

Broucek, J. (2014). Production of Methane Emissions from Ruminant Husbandry: A Review. *Journal of Environmental Protection*, 05 (15), 1482–1493. https://doi.org/10.4236/jep.2014.515141

de los Campos, G., Hickey, J.M., Pong-Wong, R., Daetwyler, H.D. & Calus, M.P.L. (2013). Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. *Genetics*, 193 (2), 327–345. https://doi.org/10.1534/genetics.112.143313

Casto-Rebollo, C., Pocrnic, I., Gorjanc, G. & Ibáñez-Escriche, N. (2024). HoloSimR: a comprehensive framework for simulating breeding programs including the hologenome scenario. https://doi.org/10.21203/rs.3.rs-5409986/v1

Difford, G.F., Plichta, D.R., Løvendahl, P., Lassen, J., Noel, S.J., Højberg, O., Wright, A.-D.G., Zhu, Z., Kristensen, L., Nielsen, H.B., Guldbrandtsen, B. & Sahana, G. (2018). Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS genetics*, 14 (10), e1007580. https://doi.org/10.1371/journal.pgen.1007580

Gaynor, R.C., Gorjanc, G. & Hickey, J.M. (2021). AlphaSimR: an R package for breeding program simulations. *G3 Genes|Genomes|Genetics*, 11 (2), jkaa017. https://doi.org/10.1093/g3journal/jkaa017

Hassanpour, A., Geibel, J., Simianer, H. & Pook, T. (2023). Optimization of breeding program design through stochastic simulation with kernel regression. De Los Campos, G. (ed.) (De Los Campos, G., ed.) *G3: Genes, Genomes, Genetics*, 13 (12), jkad217. https://doi.org/10.1093/g3journal/jkad217

Heilig, G.K. (1994). The greenhouse gas methane (CH4): Sources and sinks, the impact of population growth, possible interventions. *Population and Environment*, 16 (2), 109–137. https://doi.org/10.1007/BF02208779

Hook, S.E., Wright, A.-D.G. & McBride, B.W. (2010). Methanogens: Methane Producers of the Rumen and Mitigation Strategies. *Archaea*, 2010 (1), 945785. https://doi.org/10.1155/2010/945785

Howard, R., Carriquiry, A.L. & Beavis, W.D. (2014). Parametric and Nonparametric Statistical Methods for Genomic Selection of Traits with Additive and Epistatic Genetic Architectures. *G3 Genes|Genomes|Genetics*, 4 (6), 1027–1046. https://doi.org/10.1534/g3.114.010298

Hristov, A.N., Oh, J., Giallongo, F., Frederick, T.W., Harper, M.T., Weeks, H.L., Branco, A.F., Moate, P.J., Deighton, M.H., Williams, S.R.O., Kindermann, M. & Duval, S. (2015). An inhibitor persistently decreased enteric methane emission from dairy cows with no negative effect on milk production. *Proceedings of the National Academy of Sciences*, 112 (34), 10663–10668. https://doi.org/10.1073/pnas.1504124112

Kamalanathan, S., Houlahan, K., Miglior, F., Chud, T.C.S., Seymour, D.J., Hailemariam, D., Plastow, G., de Oliveira, H.R., Baes, C.F. & Schenkel, F.S. (2023). Genetic Analysis of Methane Emission Traits in Holstein Dairy Cattle. *Animals : an Open Access Journal from MDPI*, 13 (8), 1308. https://doi.org/10.3390/ani13081308

Khobragade, T., Damor, A., Pandya, M., Ramana, D., Majeli, A. & Yeptho, M. (2024). Rumen Ecosystem.

Lassen, J. & Difford, G.F. (2020). Review: Genetic and genomic selection as a methane mitigation strategy in dairy cattle. *Animal: An International Journal of Animal Bioscience*, 14 (S3), s473–s483. https://doi.org/10.1017/S1751731120001561

Lassen, J. & Løvendahl, P. (2016). Heritability estimates for enteric methane emissions from Holstein cattle measured using noninvasive methods. *Journal of Dairy Science*, 99 (3), 1959–1967. https://doi.org/10.3168/jds.2015-10012

Mahala, S., Kala, A. & Kumar, A. (2022). Host genetics associated with gut microbiota and methane emission in cattle. *Molecular Biology Reports*, 49 (8), 8153–8161. https://doi.org/10.1007/s11033-022-07718-1

Manzanilla-Pech, C.I.V., De Haas, Y., Hayes, B.J., Veerkamp, R.F., Khansefid, M., Donoghue, K.A., Arthur, P.F. & Pryce, J.E. (2016). Genomewide association study of methane emissions in Angus beef cattle with validation in dairy cattle. *Journal of Animal Science*, 94 (10), 4151–4166. https://doi.org/10.2527/jas.2016-0431

Martínez-Álvaro, M., Auffret, M.D., Stewart, R.D., Dewhurst, R.J., Duthie, C.-A., Rooke, J.A., Wallace, R.J., Shih, B., Freeman, T.C., Watson, M. & Roehe, R. (2020). Identification of Complex Rumen Microbiome Interaction Within Diverse Functional Niches as Mechanisms Affecting the Variation of Methane Emissions in Bovine. *Frontiers in Microbiology*, 11. https://doi.org/10.3389/fmicb.2020.00659

Meale, S.J., Popova, M., Saro, C., Martin, C., Bernard, A., Lagree, M., Yáñez-Ruiz, D.R., Boudra, H., Duval, S. & Morgavi, D.P. (2021). Early life dietary intervention in dairy calves results in a long-term reduction in methane emissions. *Scientific Reports*, 11 (1), 3003. https://doi.org/10.1038/s41598-021-82084-9

Milich, L. (1999). The role of methane in global warming: where might mitigation strategies be focused? *Global Environmental Change*,

Murray, R.M., Bryant, A.M. & Leng, R.A. (1976). Rates of production of methane in the rumen and large intestine of sheep. *British Journal of Nutrition*, 36 (1), 1–14. https://doi.org/10.1079/BJN19760053

Oldenbroek, K. & Waaij, L. van der (2014). Animal breeding and genetics for BSc students: Textbook animal breeding. https://research.wur.nl/en/publications/animal-breeding-and-genetics-for-bsc-students-textbook-animal-bre [2025-03-06]

Opstal, E.J.V. & Bordenstein, S.R. (2015). Rethinking heritability of the microbiome. *Science*, 349 (6253), 1172–1173. https://doi.org/10.1126/science.aab3958

Perez, P. (n.d.). BGLR: A Statistical Package for Whole Genome Regression and Prediction.

Pérez, P. & De Los Campos, G. (2014). Genome-Wide Regression and Prediction with the BGLR Statistical Package. *Genetics*, 198 (2), 483–495. https://doi.org/10.1534/genetics.114.164442

Pérez-Enciso, M., Zingaretti, L.M., Ramayo-Caldas, Y. & de los Campos, G. (2021). Opportunities and limits of combining microbiome and genome data for complex trait prediction. *Genetics Selection Evolution*, 53 (1), 65. https://doi.org/10.1186/s12711-021-00658-7

Ramayo-Caldas, Y., Zingaretti, L., Popova, M., Estellé, J., Bernard, A., Pons, N., Bellot, P., Mach, N., Rau, A., Roume, H., Perez-Enciso, M., Faverdin, P., Edouard, N., Ehrlich, D., Morgavi, D.P. & Renand, G. (2020). Identification of rumen microbial biomarkers linked to methane emission in Holstein dairy cows. *Journal of Animal Breeding and Genetics*, 137 (1), 49–59. https://doi.org/10.1111/jbg.12427

Roehe, R., Dewhurst, R.J., Duthie, C.-A., Rooke, J.A., McKain, N., Ross, D.W., Hyslop, J.J., Waterhouse, A., Freeman, T.C., Watson, M. & Wallace, R.J. (2016). Bovine Host Genetic Variation Influences Rumen Microbial Methane Production with Best Selection Criterion for Low Methane Emitting and Efficiently Feed Converting Hosts Based on Metagenomic Gene Abundance. Leeb, T. (ed.) (Leeb, T., ed.) *PLOS Genetics*, 12 (2), e1005846. https://doi.org/10.1371/journal.pgen.1005846

Ross, E.M., Moate, P.J., Marett, L.C., Cocks, B.G. & Hayes, B.J. (2013). Metagenomic Predictions: From Microbiome to Complex Health and Environmental Phenotypes in Humans and Cattle. White, B.A. (ed.) (White, B. A., ed.) *PLoS ONE*, 8 (9), e73056. https://doi.org/10.1371/journal.pone.0073056

Saborío-Montero, A., López-García, A., Gutiérrez-Rivas, M., Atxaerandio, R., Goiri, I., García-Rodriguez, A., Jiménez-Montero, J.A., González, C., Tamames, J., Puente-Sánchez, F., Varona, L., Serrano, M., Ovilo, C. & González-Recio, O. (2021). A dimensional reduction approach to modulate the core ruminal microbiome associated with methane emissions via selective breeding. *Journal of Dairy Science*, 104 (7), 8135–8151. https://doi.org/10.3168/jds.2020-20005

Sasson, G., Kruger Ben-Shabat, S., Seroussi, E., Doron-Faigenboim, A., Shterzer, N., Yaacoby, S., Berg Miller, M.E., White, B.A., Halperin, E. & Mizrahi, I. (2017). Heritable Bovine Rumen Bacteria Are Phylogenetically Related and Correlated with the Cow's Capacity To Harvest Energy from Its Feed. *mBio*, 8 (4), e00703-17. https://doi.org/10.1128/mBio.00703-17

Smith, P.E., Kelly, A.K., Kenny, D.A. & Waters, S.M. (2022). Differences in the Composition of the Rumen Microbiota of Finishing Beef Cattle Divergently Ranked for Residual Methane Emissions. *Frontiers in Microbiology*, 13, 855565. https://doi.org/10.3389/fmicb.2022.855565

Tapio, I., Fischer, D., Blasco, L., Tapio, M., Wallace, R.J., Bayat, A.R., Ventto, L., Kahala, M., Negussie, E., Shingfield, K.J. & Vilkki, J. (2017). Taxon abundance, diversity, co-occurrence and network analysis of the ruminal microbiota in response to dietary changes in dairy cows. Smidt, H. (ed.) (Smidt, H., ed.) *PLOS ONE*, 12 (7), e0180260. https://doi.org/10.1371/journal.pone.0180260

Van Breukelen, A.E., Aldridge, M.N., Veerkamp, R.F., Koning, L., Sebek, L.B. & De Haas, Y. (2023). Heritability and genetic correlations between enteric methane production and concentration recorded by GreenFeed and sniffers on dairy cows. *Journal of Dairy Science*, 106 (6), 4121–4132. https://doi.org/10.3168/jds.2022-22735

Vanwonterghem, I., Evans, P.N., Parks, D.H., Jensen, P.D., Woodcroft, B.J., Hugenholtz, P. & Tyson, G.W. (2016). Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nature Microbiology*, 1 (12), 1–9. https://doi.org/10.1038/nmicrobiol.2016.170

Wallace, R.J., Sasson, G., Garnsworthy, P.C., Tapio, I., Gregson, E., Bani, P., Huhtanen, P., Bayat, A.R., Strozzi, F., Biscarini, F., Snelling, T.J., Saunders, N., Potterton, S.L., Craigon, J., Minuti, A., Trevisi, E., Callegari, M.L., Cappelli, F.P., Cabezas-Garcia, E.H., Vilkki, J., Pinares-Patino, C., Fliegerová, K.O., Mrázek, J., Sechovcová, H., Kopečný, J., Bonin, A., Boyer, F., Taberlet, P., Kokou, F., Halperin, E., Williams, J.L., Shingfield, K.J. & Mizrahi, I. (2019). A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. *Science Advances*, 5 (7), eaav8391. https://doi.org/10.1126/sciadv.aav8391

Weimer, P.J., Cox, M.S., Vieira de Paula, T., Lin, M., Hall, M.B. & Suen, G. (2017). Transient changes in milk production efficiency and bacterial community composition resulting from near-total exchange of ruminal contents between high- and low-efficiency Holstein cows. *Journal of Dairy Science*, 100 (9), 7165–7182. https://doi.org/10.3168/jds.2017-12746

Weimer, P.J., Stevenson, D.M., Mantovani, H.C. & Man, S.L.C. (2010). Host specificity of the ruminal bacterial community in the dairy cow following near-total exchange of ruminal contents1. *Journal of Dairy Science*, 93 (12), 5902–5912. https://doi.org/10.3168/jds.2010-3500

Woese, C.R., Kandler, O. & Wheelis, M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*, 87 (12), 4576. https://doi.org/10.1073/pnas.87.12.4576

Worku, D. (2024). Unraveling the genetic basis of methane emission in dairy cattle: a comprehensive exploration and breeding approach to lower methane emissions. *Animal Biotechnology*, 35 (1), 2362677. https://doi.org/10.1080/10495398.2024.2362677

Zhang, Q., Difford, G., Sahana, G., Løvendahl, P., Lassen, J., Lund, M.S., Guldbrandtsen, B. & Janss, L. (2020). Bayesian modeling reveals host genetics associated with rumen microbiota jointly influence methane emission in dairy cows. *The ISME Journal*, 14 (8), 2019–2033. https://doi.org/10.1038/s41396-020-0663-x

Popular science summary

Cattle play a crucial role in agriculture but are also significant contributors to the emission of methane, which is a greenhouse gas that affects climate change. Methane is produced during digestion by microbes living in the cow's stomach (the rumen). This study examined how a cow's genetics and the types of microbes in its rumen affect methane production.

We tested various scenarios using AlphaSim R simulations to determine how genetic traits and microbial activity contribute to methane emissions. Our findings revealed that cattle with fewer methane-producing microbes and selective breeding based on their genetic makeup could significantly reduce methane emissions over time.

Additionally, this simulation study showed that genetic factors substantially influence methane traits when combined with low microbial effects. These findings provide a framework for sustainable breeding practices that reduce methane emissions and support livestock production, aligning with global efforts to combat climate change while maintaining productivity.

Acknowledgements

First and foremost, I want to express my deepest gratitude to Allah Almighty for His grace, wisdom, and strength, which enabled me to complete this work successfully.

I am profoundly thankful to my main supervisor, Dr. Martin Johnsson, for his invaluable guidance, encouragement, and intellectual support throughout this study. I also extend my heartfelt thanks to my co-supervisors, Professor Armin Schmitt and Professor Jan Ingemar Ohlsson, for their expert advice and thoughtful insights, greatly enriching this research.

I am deeply grateful to my parents, whose unwavering support and encouragement have been my foundation. To my mother: your constant love and prayers have always anchored me, and to my father: your belief in my abilities has inspired me to strive for excellence. A special thank you goes to my daughter, Zimr Daneen, who has been a source of endless joy and motivation. Your innocent laughter and presence have reminded me of the purpose and fulfillment behind this journey.

Finally, I extend my gratitude to the EMABG scholarship board for their financial support, which made this academic journey possible. To my colleagues in the EMABG program, thank you for your support.

.

Publishing and archiving

Approved students' theses at SLU can be published online. As a student you own the copyright to your work and in such cases, you need to approve the publication. In connection with your approval of publication, SLU will process your personal data (name) to make the work searchable on the internet. You can revoke your consent at any time by contacting the library.

Even if you choose not to publish the work or if you revoke your approval, the thesis will be archived digitally according to archive legislation.

You will find links to SLU's publication agreement and SLU's processing of personal data and your rights on this page:

- https://libanswers.slu.se/en/faq/228318

☒ YES, I, Nateela Shoukat, have read and agree to the agreement for publication and the personal data processing that takes place in connection with this

☐ NO, I/we do not give my/our permission to publish the full text of this work. However, the work will be uploaded for archiving and the metadata and summary will be visible and searchable.