UNIVERSITAT POLITÈCNICA DE VALÈNCIA

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Dpto. de Ciencia Animal

Selección por el muestreo mendeliano para aumentar la variabilidad en un programa de mejora genética.

Trabajo Fin de Máster

Máster Universitario en Mejora Genética Animal y Biotecnología de la Reproducción

AUTOR/A: Tabet , Joe Menwer

Tutor/a: Ibáñez Escriche, Noelia

Cotutor/a externo: KRANIS, ANDREAS

CURSO ACADÉMICO: 2021/2022

Master's Degree in Animal Genetics and Breeding and Reproductive Biotechnology

Selecting for Mendelian Sampling term to increase variability in a breeding program.

Master's Thesis

June 2022

Joe-Menwer Tabet

Director:

Dr. Andreas Kranis

Senior research fellow

The Roslin Institute |

University of Edinburgh

Table of Contents

List of figures:

List of tables:

Dedication:

This project is dedicated to my extended family, who is dispersed worldwide; without their help and support, I would not be able to reach the goals I had hoped to achieve in my life.

Thank you, family and friends.

I am not alone, for my father is with me, خليل

Acknowledgment

To all those who made completing this master's possible: my family, who accompanied me throughout my life; teachers and professors throughout my academic training; my friends, who were always present from a distance; my bosses and colleagues.

Thank you to Dr. Andreas Kranis, who made working with him a learning adventure, introducing me to new concepts daily. Moreover, who accompanied me this year and stood by me every step of the way to finish my project.

To the Department of Animal Science, who provided the environment where we could thrive and conduct our research. Thank you to all the members who kept pushing me and kept reminding me what it means to be a scientist and an animal geneticist. Who always made sure I questioned everything and always did my best to find the solution to our problems.

To Dr. Noelia and Dr. Agustìn, who stayed by my side and kept pushing me to become the best I could be, never letting me give up halfway. Thank you for these wonderful two years and for making me feel like I belong not to a team but a family. Forever grateful.

Abstract

Increasing genetic gain while controlling the rate of inbreeding in a population has been most studied by quantitative geneticists and animal breeders. Many selection methods were developed to maximize genetic gain while applying restrictions on the inbreeding rate. Some studies developed quadratic indices (Avendaño et al., 2004); others calculated the gametic variance σ^2_{gamete} to increase the long-term response to selection and limit the increase of the rate of inbreeding to avoid its deleterious consequences (Santos et al., 2019). Mendelian Sampling term, defined as the deviation of the average additive effects of an individual's genes received from both parents from the average effects of genes of both parents, can be used in breeding programs to decrease the inbreeding rate and increase genetic variability in a population.

Including the Mendelian Sampling term in the selection process could present an alternative approach to controlling inbreeding and increasing genetic variability in a population.

In this study, an individual's Mendelian Sampling term (MS) is calculated from a breeding program simulation and from actual data provided by Aviagen and used as a selection criterion for a trait instead of the breeding value. The breeding program simulations were performed using AlphaSimR and aimed to investigate the evolution of a population undergoing selection based on the Mendelian Sampling term and its effect on the evolution of the population. The estimation of the Mendelian Sampling's heritability in real data was performed using REML (Restricted Maximum Likelihood) and Gibbs Sampling to provide evidence into the heritability of the Mendelian Sampling and provide a proof of concept on how to estimate it in a breeding program to possibly create more long-term response to selection and genetic variability in a population.

In addition, to the Mendelian Sampling term, the gametic variance σ^2_{gamete}, which is defined as the variability of all possible gametic breeding values produced by recombination and permutation of each parenteral chromosome, was also calculated and included in the selection process in order to investigate the feasibility of more genetic variability in a population and to increase the long-term response to selection. The gametic variance σ^2_{gamete}, according to Santos et al., 2019, is used to identify individuals who are more likely to produce more variable gametes to preserve the

genetic variability. In this study, the inclusion of both these terms was for the purpose of examining the likelihood of these two components in increasing genetic gain and variability while controlling the inbreeding rate. The accumulated genetic progress was calculated over ten generations of selection.

Using simulations, the breeding program with the Mendelian Sampling term as a selection criterion showed the lowest inbreeding levels and the highest genetic variability levels but presented slower genetic progress compared to selecting the True Breeding value of an individual in a selection program. The Mendelian Sampling term program showed a better genetic gain rate (ΔG) than all the other breeding programs considered in the simulation. When the gametic variance σ^2_{gamete} was included, the breeding programs that contained the Mendelian Sampling term resulted in better accumulated genetic gain than the Mendelian Sampling term alone, allowing for future considerations of using these terms in a breeding program.

Keywords: Mendelian Sampling term, Gametic variance σ^2_{gamete}, Inbreeding rate, Genetic variance, variability, breeding program.

Resumen

El aumento de la ganancia genética mientras se controla la tasa de consanguinidad en una población ha sido más estudiado por genetistas cuantitativos y criadores de animales. Se desarrollaron muchos métodos de selección para maximizar la ganancia genética mientras se aplicaban restricciones a la tasa de consanguinidad. Algunos estudios desarrollaron índices cuadráticos (Avendaño et al., 2004); otros calcularon la varianza gamética σ^2_{gamete} para aumentar la respuesta a largo plazo a la selección y limitar el aumento de la tasa de consanguinidad para evitar sus consecuencias deletéreas (Santos et al., 2019). El término de Muestreo Mendeliano, definido como la desviación de los efectos aditivos promedio de los genes de un individuo recibidos de ambos padres de los efectos promedio de los genes de ambos padres, puede usarse en programas de mejora para disminuir la tasa de consanguinidad y aumentar la variabilidad genética en una población.

Incluir el término de Muestreo Mendeliano en el proceso de selección podría presentar un enfoque alternativo para controlar la consanguinidad y aumentar la variabilidad genética en una población.

En este estudio, el término de Muestreo Mendeliano (MS) de un individuo se calcula a partir de una simulación del programa de mejora y de los datos reales presentados por Aviagen y se utiliza como criterio de selección para un carácter en lugar del breeding value. La simulación del programa de mejora se realizó utilizando AlphaSimR y tuvo como objetivo investigar la evolución de una población en proceso de selección basada en el término de Muestreo Mendeliano y su efecto sobre la evolución de esa población. La estimación de la heredabilidad del Muestreo Mendeliano con datos reales se realizó utilizando REML (Restricted Maximum Likelihood) y Gibbs Sampling para presentar evidencia sobre la heredabilidad del Muestreo Mendeliano y proporcionar una prueba de concepto sobre cómo estimarlo en un programa de mejora para posiblemente crear respuesta a más largo plazo a la selección y la variabilidad genética en una población.

Además del término de Muestreo Mendeliano, también se calculó e incluyó en el proceso de selección la varianza gamética σ^2_{gamete}, que se define como la variabilidad de todos los posibles valores genéticos gaméticos producidos por recombinación y permutación de cada cromosoma parenteral, con el fin de investigar la viabilidad de una mayor variabilidad genética en una población y la posibilidad de aumentar la respuesta a largo plazo a la selección. La varianza gamética σ^2_{gamete}, según Santos et al., 2019, se utiliza para identificar individuos que tienen más probabilidades de producir gametos más variables para preservar la variabilidad genética. En este estudio, la inclusión de estos dos términos fue con el propósito de examinar la probabilidad de que estos dos parámetros aumenten la ganancia y la variabilidad genética mientras se controla la tasa de consanguinidad. El progreso genético acumulado se calculó a lo largo de diez generaciones de selección.

Mediante simulaciones, el programa de mejora con los términos del Muestreo Mendeliano como criterio de selección mostró los niveles más bajos de consanguinidad y los niveles más altos de variabilidad genética, pero presentó un progreso genético más lento en comparación con la selección del valor True Breeding de un individuo en un programa de selección. El programa de término de Muestreo Mendeliano mostró una mejor tasa de ganancia genética (ΔG) que todos los demás programas de mejora considerados en la simulación. Cuando se incluyó la varianza gamética σ^2_{gamete}, los programas de mejora que contenían el término de Muestreo Mendeliano con la varianza gamética σ^2_{gamete}, dieron en una mejor ganancia genética acumulada que el término de Muestreo Mendeliano solo, lo que permitió futuras consideraciones sobre el uso de estos términos en un programa de mejora.

Palabras clave: Muestreo Mendeliano, Varianza gamética σ^2_{gamete}, Tasa de consanguinidad, Varianza genética, variabilidad, programa de mejoramiento.

Résumé

Augmenter le gain génétique tout en contrôlant le taux de consanguinité dans une population a été le plus étudié par les généticiens quantitatifs et les éleveurs d'animaux. De nombreuses méthodes de sélection ont été développées pour maximiser le gain génétique tout en appliquant des restrictions de taux de consanguinité. Certaines études ont développé des indices quadratiques (Avendaño et al., 2004) ; d'autres ont calculé la variance gamétique σ^2_{gamete} pour augmenter la réponse à long terme de la sélection et limiter l'augmentation du taux de consanguinité pour éviter ses conséquences délétères (Santos et al., 2019). Le terme échantillonnage Mendélien, défini comme l'écart entre les effets additifs moyens des gènes d'un individu reçus des deux parents et les effets moyens des gènes des deux parents, peut être utilisé dans les programmes de sélection pour diminuer le taux de consanguinité et augmenter la variabilité génétique dans un population.

L'inclusion du terme échantillonnage Mendélien dans le processus de sélection pourrait présenter une approche alternative pour contrôler la consanguinité et augmenter la variabilité génétique dans une population.

Dans cette étude, le terme d'échantillonnage Mendélien (MS) d'un individu est calculé à partir d'une simulation du programme d'élevage et de données réelles présentées par Aviagen et utilisé comme critère de sélection d'un caractère plutôt que d'une valeur d'élevage. La simulation du programme d'amélioration a été réalisée à l'aide d'AlphaSimR et visait à étudier l'évolution d'une population en cours de sélection basée sur le terme d'échantillonnage Mendélien et son effet sur l'évolution de cette population. L'estimation de l'héritabilité de l'échantillonnage Mendélien sur des données réelles a été réalisée à l'aide du REML (Restricted Maximum Likelihood) et de l'échantillonnage de Gibbs pour présenter des preuves sur l'héritabilité de l'échantillonnage mendélien et fournir une preuve de concept sur la façon de l'estimer dans un programme de sélection pour éventuellement créer une réponse à plus long terme à la sélection et à la variabilité génétique d'une population.

En plus du terme d'échantillonnage Mendélien, la variance gamétique σ^2_{gamete}, qui est définie comme la variabilité de toutes les valeurs d'élevage gamétiques possibles produites par

recombinaison et permutation de chaque chromosome parent, a également été calculée et incluse dans le processus de sélection, afin d'étudier la faisabilité d'une variabilité génétique accrue dans une population et la possibilité d'augmenter la réponse à long terme à la sélection. La variance gamétique σ^2_{gamete}, selon Santos et al., 2019, est utilisée pour identifier les individus les plus susceptibles de produire des gamètes plus variables afin de préserver la variabilité génétique. Dans cette étude, l'inclusion de ces deux termes visait à examiner la probabilité que ces deux paramètres augmentent le gain et la variabilité génétique tout en contrôlant le taux de consanguinité. Le progrès génétique cumulé a été calculé sur dix générations de sélection.

Grâce à des simulations, le programme de sélection avec les termes d'échantillonnage Mendélien comme critères de sélection a montré les niveaux de consanguinité les plus bas et les niveaux de variabilité génétique les plus élevés, mais a présenté des progrès génétiques plus lents par rapport à la sélection de la valeur True Breeding d'un individu dans un programme de sélection. Le programme de termes d'échantillonnage Mendélien a montré un meilleur taux de gain génétique (ΔG) que tous les autres programmes d'élevage pris en compte dans la simulation. Lorsque la variance gamétique σ^2_{gamete} a été incluse, les programmes de sélection contenant le terme d'échantillonnage Mendélien ont entraîné un meilleur gain génétique cumulé que le terme d'échantillonnage Mendélien seul, ce qui a permis de poursuivre l'examen de l'utilisation de ces termes dans un programme de sélection.

Mots clés : Échantillonnage Mendélien, variance gamétique σ^2_{gamete}, taux de consanguinité, variance génétique, variabilité, programme d'élevage.

Introduction

In agriculture, breeders aim to improve the performance of the population based on various criteria; this means selecting the best performing individuals for a particular trait as parents of the next generation. Achieving sustained genetic progress in key traits has been the main objective in breeding programs throughout the years. Selection schemes have significantly improved in choosing the best individuals in a population and increasing the accuracy of evaluation methods used to estimate the breeding values of these selected individuals. For example, BLUP (Best Linear Unbiased Predictor), using either pedigree or genomic information or both, and the introduction of these methods has led to an increase in genetic gain and improvement. However, this increase in genetic improvement in a population has its downside; one is the loss of genetic variability.

In a population where mating is random, all families have an equal opportunity to contribute offspring to the next generation. However, things differ in a population undergoing selection. Superior families will contribute more to the next generation than lower-performing families, which in turn increases inbreeding and leads to loss of genetic variability down the line. Grundy et al., 1998 postulated that genetic improvement in selection programs is usually associated with an increased inbreeding and a loss of genetic variability within the population.

The management of genetic gain and the long-term genetic variability has been one of the main focuses for quantitative geneticists and animal breeders.

True Breeding Value and Mendelian Sampling term:

Let (a_i) be the individual (i) 's breeding value in a population. The law of inheritance entails that an individual with parents (s) and (d) (Sire and Dam) receives one-half of the genes from each parent, so in retrospect, an individual's true breeding value is regarded as the average breeding value of both parents. Since not all offspring receive exactly the same genes from both parents, no progeny is equal to its parent's average. The offspring deviates from the parent's average. That is what we call Mendelian Sampling term. It could be defined as the deviation of the average additive

effects of (*i*) 's genes received from both parents from the average effects of genes of both parents (R.A. Mrode, 2005).

Individual (*i*) breeding value is defined as:

$$a_i = \frac{1}{2} a_s + \frac{1}{2} a_d + ms \quad (1)$$

Where (a_i) is (*i*) 's breeding value, (a_s) and (a_d) are the breeding values of its parents, and (*ms*) is the Mendelian sampling term.

The success of a selection scheme is its ability to accurately estimate an individual's breeding value to correctly identify superior individuals in a population. To accurately predict the breeding value, the method used to estimate the breeding value depends on the information available, such as own information, parental average, progeny performance, pedigree information, or genomic information.

Genetic improvement has been revolutionized with the use of genomic information, and improvement in selection schemes has increased immensely. With the development of genotyping technology and the ability to genotype an animal is becoming cheaper, access to an individual's genomic information has become easier than before. That has substantially increased genetic progress in animal breeding. This has been done by successfully identifying genetically superior animals based on their performance records and based on the genomic information they present. In some breeding programs, genomic selection has doubled the yearly genetic gain compared to other conventional programs (Bijma et al., 2020).

Furthermore, implementing genomic selection in breeding schemes is becoming increasingly popular. Genomic selection puts more pressure on the genome, on more specific regions of the genome that presents a high contribution to the additive effect of a particular trait undergoing selection, and this might increase the risk of loss of favorable alleles and limit the potential of selection on the same trait or even other traits in the future (Wientjes et al., 2022).

Selection schemes have increased concerns regarding the rise of inbreeding and reduced genetic variability, limiting future selection in a population. Various selection strategies have been introduced to control inbreeding and maximize genetic gain or vice versa. Moreover, many of them utilized the Mendelian sampling term as an advantage to lower inbreeding levels and limit the loss

of variability in a population. This resulted in defining the Mendelian Sampling term not only as the deviation from the parent's average.

Bonk et al., 2016, defined Mendelian Sampling as the degree of variability among full-sibs due to the inheritance of random samples of alleles from both parents and postulated that a high Mendelian Sampling variability increases selection opportunities between siblings; in other words identifying the superior sibling within a family. Avendaño et al., 2004 defined Mendelian Sampling as the new variability created each generation. The same study formulated a selection algorithm where the Mendelian Sampling term presented a selective advantage in reducing inbreeding rates and controlling genetic gain. They demonstrated that the success of an optimum breeding program is the utilization of the Mendelian Sampling term information in optimum quadratic indices strategies. These indices could alleviate the weight put on the intrafamily information that led to an increase in genetic gain at the same level of inbreeding. This study relied on the definition of long-term genetic contribution introduced by Thompson et al., 1994. An individual's breeding value can be decomposed into its Mendelian Sampling term plus the average value of its parents, but also postulated that the parent's breeding value could also be decomposed into the same components, and so can their parents; this process can be applied throughout the pedigree. Thus, Thompson et al., 1994, defined an individual's breeding value as the weighted sum of the Mendelian Sampling term of all its ancestors, which led them to say that genetic gain is a function of the Mendelian Sampling term. Based on that definition, Avendaño et al., 2004, presented an algorithm that could maximize genetic gain through a quadratic selection index. These indices' objective is to optimize the contribution of selected candidates to maximize genetic gain and limit rates of inbreeding. This study showed that the Mendelian Sampling term is a selective advantage when using these indices. The quadratic index performed better than the BLUP (Best Linear Unbiased Prediction) selection strategy in a breeding program simulation due to the best available information on the Mendelian Sampling term.

In another study by Grundy et al., 1998, they used Mendelian Sampling term indices to reduce the rate of inbreeding in a selection program. They simulated a breeding program that used these indices instead of BLUP's estimated breeding value over several generations. They argued that using the Mendelian Sampling term would change the weight given to the family information, decreasing inbreeding rates. Compared with the BLUP-based selection strategy, the Mendelian

Sampling term indices reduced the inbreeding rate by 24%, presenting a higher variance than with BLUP (Grundy et al., 1998). Including the Mendelian Sampling terms into the equation of candidate selection has been studied for many years. It has shown that it reduces inbreeding and the loss of genetic variability. Better estimations of the Mendelian Sampling terms increases the differentiation between full sibs within a family, reducing the co-selection of siblings in a population, which leads to a decrease in the inbreeding rate (Daetwyler et al., 2007).

Starting from equation (1), the calculation of the Mendelian sampling term is as follows:

$$ms_i = ai - \frac{1}{2}(a_s + a_d) \qquad (2)$$

where (ms_i) is the Mendelian Sampling term of the individual (i), (a_i) its breeding value and $\frac{1}{2}(a_s + a_d)$ the parent's average of the individual (*i*). From equation (2), we can calculate the Mendelian Sampling variance $\sigma^2_{ms} = \frac{1}{2}\sigma^2_a$, when inbreeding is equal to zero (F=0). In a breeding program, when more selection is performed, higher the inbreeding increases, and the Mendelian Sampling variance decreases, as shown here:

$$\sigma^2_{ms} = \frac{1}{2}\sigma^2_a$$

with inbreeding F = 0, when F ≠ 0, the Mendelian sampling variance decreases and becomes equal to

$$\sigma^2_{ms} = \frac{1}{2}\sigma^2_a\,[1 - 0.5(Fs + Fd)],$$

With (*Fs*) and (*Fd*), the inbreeding coefficients of animal's (*i*) sire and dam (R.A. Mrode, 2005). The heritability of the Mendelian Sampling term when F = 0 is $h^2 = \sigma^2_{ms} / \sigma^2_p$ and since $\sigma^2_{ms} = \frac{1}{2}\sigma^2_a$ the heritability of the Mendelian Sampling term, when inbreeding is zero, is $h^2 = \frac{1}{2}\sigma^2_a / \sigma^2_p$ and this decreases with more selection and with increased inbreeding. So, for example, an $h^2 = 0.5$ should have a heritability of the Mendelian Sampling term of 0.25 when F=0, or a genetic variance of 0.30 should have a mendelian sampling variance of 0.15 when F=0.

Quantitatively we can say that the Mendelian Sampling term is heritable, and the heritability depends on the trait's heritability. Still, since it is heritable, it can undergo selection and might be taken advantage of in a breeding program. Estimation of the Mendelian Sampling term highly depends on the method of estimating breeding values and the estimation's accuracy. In this study, we aim to use a simulation program (AlphaSimR), so we can investigate the evolution of a population undergoing selection of candidates based on their best Mendelian Sampling term instead of breeding values for the possibility of increasing the levels of genetic variability within a selected population; more on that in the section on Material and Methods.

Gametic Variance:

The variance of gametic diversity σ^2_{gamete} is the variability of all possible gametic values produced by recombination and permutation of each parenteral chromosome (Santos et al., 2019). It means that only heterozygous loci of an individual contribute to $\sigma^2_{gamete,}$ signifying that to calculate the value of σ^2_{gamete} for a specific individual, only heterozygous loci and their effects are considered.

Santos et al., 2019 calculated the σ^2_{gamete} using a binomial variance; in other words, at a single biallelic locus (*k*) of animal (*i*) with allele substitution effect (α_k), the gametic variance formula is

$$\sigma^2_k = np(1-p)\ \alpha_k^2,$$

where n = 1 the number of alleles transmitted to gamete loci, p =0.5 the probability of transmitting an allele to a gamete.

Considering this formula for all QTLs in the genome, the gametic variance becomes

$$\sigma^2_{gamete} = \sum_{k=1}^{N} \sigma^2_k + 2 \sum_{k=1}^{N} \sum_{j=k+1}^{N} \sigma_{jk}, \qquad (3)$$

where (*k*) and (*j*) are two loci on the same chromosome and are supposed to be inherited together, σ_{jk} is the covariance between the loci (*k*) and (*j*), and (N) is the number of heterozygous loci. Assuming that all loci are independent means that they segregate independently, the gametic

variance becomes: $\sigma^2_{gamete} = [\alpha_1 \ldots \alpha_j]\ \mathbf{M}\ [\alpha_1 \ldots \alpha_j]'$, and the matrix **M** takes the following shape when all loci are independent

$$\mathbf{M} = \begin{bmatrix} 0.25 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 0.25 \end{bmatrix}$$

According to the same study, the variance can be used to identify animals with higher probabilities of producing progenies with more variability and provide better future gametes. Enhanced genetic gains have been noted when the selection process included individuals presenting the best gametes in a population (Goiffon et al., 2017). Santos et al., 2019 used this variance and included it in a selection program to improve the genetic progress. As a result, the group obtained better genetic improvement when σ^2_{gamete} was included in a selection index compared to the genomic selection, with a 3-8% increase in genetic progress. There was also an increase in the frequency of rare favorable alleles in the population.

Selection indices:

In this project, a population underwent four different selection schemes. Each one was evaluated by genetic progress, additive genetic variance, inbreeding, and QTL frequencies.

The first two selection schemes are straightforward: animals with the highest true breeding value in one and the highest Mendelian sampling term in another are selected as parents for the next generation. As for the other two schemes, along with the true breeding value or the Mendelian Sampling term, each individual's gametic variance σ^2_{gamete} value is added to the decision-making process. The two terms were summed up, and the candidates for the next generation were selected based on that value. The purpose of including the gametic variance σ^2_{gamete} was to provide more weight to the heterozygosity in the population. In order to increase the variability of the population, it is assumed that including the gametic variance σ^2_{gamete} would increase the chance of preserving some QTLs that might otherwise be lost in the conventional selection schemes. Taking this

variance into consideration will add a value for each individual that will be seen as their heterozygous breeding value ability; that is, for each individual's genotype, we would calculate the value of their heterozygous loci multiplied by the QTL effect of that locus, which would lead assumingly to a higher variability level and less inbreeding in a population.

Objective

This project aims to study the evolution of a population undergoing selection via different selection criteria. AlphasimR, a breeding R-package simulator, is used to simulate a breeding program that will undergo different criteria to select candidates for the next generation; the populations will undergo ten generations of selection.

There will be four different selection processes, and each one will be evaluated in the same manner.

The selection processes will be as follows:

1- Selecting individuals based on the highest true breeding value
2- Selecting individuals based on the highest true Mendelian Sampling value
3- Selecting individuals based on the highest value from the sum of the true breeding value and the gametic variance σ^2_{gamete}
4- Selecting individuals based on the highest value from the sum of the true Mendelian Sampling value and the gametic variance σ^2_{gamete}

The evaluation of each breeding program and the analysis of each population's development will be done by assessing the genetic trend. Since we can obtain true additive genetic values from the populations, genetic progress or gain will be calculated from the mean genetic values of each generation. The additive genetic variance will also be calculated from the genetic values obtained from the simulation.

Investigating these strategies presents a new perspective on controlling the increase of inbreeding within a population and unlocking the population's genetic variability potential, consequently reducing the loss of favorable alleles at low frequencies within a population and providing better long-term selection.

Since the Mendelian Sampling term is a selection criterion, the variance of Mendelian Sampling values will also be evaluated, along with calculating the pedigree inbreeding coefficient. Moreover, the frequency of QTLs is calculated and compared in each population.

Furthermore, the heritability of the Mendelian Sampling term will be estimated through Restricted Maximum Likelihood (REML) and Gibbs Sampling from simulated data obtained from the simulations and also from actual data from a breeding program shared by Aviagen. Ltd.

Estimating the heritability and investigating the selection of the Mendelian Sampling term in a breeding program provides a starting point for implementing the Mendelian Sampling term as a selectable trait and the opportunity to be integrated into breeding programs as a selection criterion.

Material and Methods

Breeding program simulations:

The investigation of strategies based on controlling inbreeding and consequently loss in genetic variance were put forward. Breeding program simulations were used to model populations with discrete generations undergoing selection, using the R-package AlphaSimR.

AlphaSimR is an animal and plant breeding program simulator (Faux et al., 2016). This program simulates genomic data, a population genome, and a specific user criteria trait. The genetic variance, phenotypic variance, number of loci controlling the trait, and its effect are user-specified and controlled.

An infinitesimal additive model was simulated for a single trait with heritability equal to 0.35; both sexes were recorded for the trait. The true breeding value (TBV) of individuals in the base population was obtained from a normal distribution with a mean of zero and an additive genetic variance σ^2_a =0.35; AlphaSimR allows the simulation of genotypes and QTL effects. These effects are normally distributed with a mean equal to that of the trait, zero, and a variance equal to the additive genetic variance σ^2_a**.** Consequently, the True breeding values are the sum of the QTL effects of each individual's genotype. The breeding value in subsequent generations corresponds to the value of the interaction of both gametes or haplotypes passed on from each parent to create the new genotype of the progeny.

The genetic variance is set equal to the heritability of the trait, $\sigma^2_a = h^2 = 0.35$; therefore, the phenotypic values of the individuals were obtained by adding a normally distributed environmental component with a mean equal to zero and a variance of $\sigma^2_e = 1 - \sigma^2_a$ (with $\sigma^2_p = 1$).

The Mendelian sampling term values for everyone were calculated as the difference between the progeny's True breeding value and its parent's True breeding value average, as shown in the equation (2),

$$ms_i = a_i - \frac{1}{2}(a_s + a_d) \qquad (2)$$

where (ms_i) is the Mendelian Sampling term of the individual (i), (a_i) its True breeding value and $\frac{1}{2}(a_s + a_d)$ the parent's average of the individual (i).

It must be noted that the Mendelian Sampling term in the base population is equal to the individuals' true breeding values.

It is assumed that the Mendelian Sampling term is taken from a normal distribution with a mean of zero and a variance that, according to R.A. Mrode, 2005, is equal to

$$\sigma^2_{ms} = 0.5\,[1\text{-}0.5\,(F_s + F_d)]\,\sigma^2_a \quad \text{with } \sigma^2_a = h^2 \text{ therefore,}$$

$$\sigma^2_{ms} = 0.5\,[1\text{-}0.5\,(F_s + F_d)]\,h^2$$

with F_s and F_d are the inbreeding coefficients of the sire and dam.

It is important to note that when F = 0 (there is no inbreeding), the Mendelian Sampling variance is equal to half of the genetic value $\sigma^2_{ms} = 0.5\,\sigma^2_a$.

Selection is made through standard truncation selection, where a fixed number of females and males are selected as candidates for the next generation. The base population is composed of 300 females and 30 males. The selection of the candidates is based on the highest selection index implemented in that selection scheme, whether it was their true breeding value or the Mendelian Sampling value.

A total of seven replicates were repeated for each selection scheme, where mating was done at random. Each mating produced three animals, and each generation was a discrete generation. The populations remained at a constant size of 2000 animals per generation with an equal frequency of each sex. At each generation and for each of the seven replicates, the genetic mean, the genetic

variance, the Mendelian Sampling variance, and the inbreeding coefficients were calculated for the animals born in that generation.

Table 1: Summary of simulation parameters

Parameter Summary	
Genome Parameter	**Value**
# Of Chromosomes	10
# Of Segregating sites	4000/chromosome
Mutation rate	2.5×10^{-08}
# Of QTLs	100/chromosome
QTL position	Random (Uniform distribution)
Trait parameters	
# Of traits	1
Heritability h^2	0.35
Phenotypic variance σ^2_p	1
Genetic Variance σ^2_a	0.35
Environmental Variance σ^2_e	0.65
Sex limited trait	No
Selection Parameters	
# Of Generations	10
Reference Population	Generation 0
# Of individuals in the founder population	8000 (50% females)
Discrete Generation	Yes
# Of individuals per Generation	2000 (50% Females)
# Male candidates	30 (1.5%)
# Female candidates	300 (15%)
# Offspring per female	3
Selection Criterion	True Breeding Value Mendelian Sampling True Breeding Value + σ^2_{gamete}
	Mendelian Sampling + σ^2_{gamete}

The mutation rate was fixed at 2.5×10^{-08}. The total segregating sites in the genome were 4000 sites per chromosome with ten chromosomes as a total. Each chromosome contained 100 QTLs, with a total of QTLs equal to 1000, that are randomly distributed uniformly.

There are four selection schemes based on the different criteria that will be further explained in the selection schemes section. Each scenario was repeated seven times.

Table 1 summarizes all the parameters set to perform all the simulations.

Gametic Variance:

The gametic variability was proposed by Santos et al., 2019 to evaluate the potential haploid breeding values to improve genetic gain and variability within a population. Several studies proposed new selection criteria to identify individuals with the best variable gametes to enhance genetic gain and reduce genetic loss (Bijma et al., 2020; Bonk et al., 2016; Goiffon et al., 2017). As a result, these studies proved that selecting the best gametes is beneficial for breeding programs because it shows a better genetic gain rate. (Note that genetic gain rate is calculated as the difference in genetic mean between two generations; this is mentioned to clarify that a population on a particular selection criterion might have a higher genetic mean or cumulative genetic gain but a lower genetic gain rate between two generations, which means that the progress from generation to another is decreasing in comparison to another population, and that is due to genetic variability loss from generation to another. Meanwhile, a higher genetic gain rate means there is still the possibility of genetic progress within a population due to more genetic variability.)

The gametic variance σ^2_{gamete} is the variability of gametic values produced by permutation and recombination of each parental chromosome; in other words, it is the value of the variability that a gamete can produce by each parent. It is measured by taking the sum of the effects of the heterozygous loci from each parental genotype. In this way, the heterozygous profile of each parent is given a value and, therefore, can be selected based on its value. The heterozygous profile of an individual is the indicator of the variability this individual can pass to its progeny, and the gametic variance is a measurement of that probability. The gametic variance σ^2_{gamete} is a tool to identify individuals that produce more variable gametes than others. In this project, the differences in

gametic variability are quantified by simulating a set of virtual QTLs with their effects (randomly distributed), generating genotypes for each individual within the simulation, and calculating the sum of the heterozygous loci effects. Recombination and linkage maps are not available for this study, so the loci are assumed to segregate independently. Each has the same probability of being passed on to the progeny, and the covariance between 2 loci is zero.

The calculation of the gametic variance, as shown before in equation (3), is,

$$\sigma^2_{\text{gamete}} = \sum_{k=1}^{N} \sigma^2_{\text{k}} + 2 \sum_{k=1}^{N} \sum_{j=k+1}^{N} \sigma_{\text{jk}}, \qquad (3)$$

where (*k*) and (*j*) are two loci on the same chromosome

σ_{jk} is the covariance between the loci (*k*) and (*j*)

(N) is the number of heterozygous loci

In matrix form, the gametic variance becomes:

$$\sigma^2_{\text{gamete}} = [\alpha_1 \dots \alpha_j]\ \mathbf{M}\ [\alpha_1 \dots \alpha_j]'$$

with α, the QTL effect of the heterozygous loci, and the matrix **M** takes the following shape when all loci are independent

$$\mathbf{M} = \begin{bmatrix} 0.25 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 0.25 \end{bmatrix}$$

The application of the gametic variance in selection programs is included in selection indices that will be further explained in the following section.

Selection index:

For this study, four selection schemes were proposed to preserve genetic variability and gain within a population, decreasing inbreeding while minimizing loss in genetic improvement.

The first breeding program is based on selecting candidates with the highest True breeding values (TBV). This program will be the model to which the three other programs will be compared. Selecting for True breeding value provides the reference values regarding inbreeding, genetic variance, genetic gain, and genetic mean.

The second breeding program selects candidates based on the individual's highest Mendelian Sampling values (MS). As seen before in other studies, the objective of this program is that when included in a selection index, it results in lower inbreeding rates, hence better levels of genetic variability. Here, the Mendelian Sampling will be considered a selectable trait, in addition, to an attempt to decrease inbreeding levels. In this program, calculation of the variance of Mendelian Sampling term and estimation of its heritability is included. As mentioned before, the Mendelian Sampling variance is expected to equal one-half of the genetic variance when inbreeding is equal to zero (Base Population-level) and, upon selection, starts to decrease due to the accumulation of inbreeding. As for the Mendelian Sampling term heritability, it is estimated to be at least equal to half of the trait's heritability (the estimation procedure of h^2 is described in the section Heritability estimation). This index provides the first step into considering the Mendelian Sampling value of an individual as a selective advantage when choosing candidates in a breeding program.

The third breeding program (SI_{TBV}) is a selection index based on the True breeding values and gametic variance. Candidates will be selected based on the highest value of the index. This index aims to preserve genetic improvement as much as possible while preventing genetic loss. This index aims to prevent the loss of rare favorable alleles that add value to genetic variability within a population while minimizing the loss of genetic improvement measured by the average genetic values and reducing the decrease of the genetic variance. Santos et al., 2019 proposed a similar index in their study. As a result, it showed better genetic gain and more variability within the population than the conventional selection schemes using BLUP and genomic selection. Adding the gametic variance σ^2_{gamete} counts for the individual with the best variable gametes available to pass on to their progeny. Along with the highest True breeding, this index grants the possibility to select the best performers within the population and the best individuals with the most variable gametes available to be inherited onto the next generation since selecting the highest True breeding value in a population represents the highest possible values obtained concerning genetic values and best performers. The expected performance of this index is to minimize the loss of genetic

improvement (expressed as the genetic mean of a population per generation) while preserving a certain amount of variability within a population. In the form of an equation, the selection index $SI_{TBV,}$

$$SI_{TBV(i)} = a_i + \sigma^2_{\text{gamete (i)}} \qquad (4)$$

The fourth breeding program (SI_{MS}) is similar to the third with one distinction; the highest Mendelian Sampling terms are used instead of the True breeding value plus the gametic variance in this index. This index focuses on increasing variability within a population as much as possible. This index presents the opportunity to further reduce inbreeding in a population. This selection scheme is aimed to better maintain genetic variability in a breeding program. Moreover, this scheme is expected to give a higher genetic gain rate (as explained earlier) since it provides a strategy to preserve rare favorable alleles that, with selection, will increase their frequencies and add value to the genetic variability and genetic progress in future generations. This index in the form of an equation is as follows:

$$SI_{MS(i)} = ms_{(i)} + \sigma^2_{\text{gamete (i)}} \qquad (5)$$

Based on the study's objective, all the different breeding programs' performance will be evaluated in the same manner to investigate the possibility of increasing genetic gain, genetic variance, and reducing inbreeding rates while minimizing the loss in genetic progress interpreted as the genetic mean of the population.

Each breeding program will be assessed according to the following criteria:

Response to the selection or the cumulative genetic gain, which will be calculated from the mean of the genetic values of the individuals in each generation, is done to investigate the effect of each scheme on the selection response of the populations. In addition, the genetic gain rate will be calculated as the differential between two generations.

Genetic variance and Mendelian Sampling variance per generation are calculated from the breeding values and Mendelian Sampling terms of each population to assess the variability for each one.

Inbreeding coefficients and inbreeding rates are calculated from pedigree information obtained from the RENUMF90 program, part of the BLUPF90 programs.

Heritability estimation:

In this project and part of the investigation into the possibility of selecting the Mendelian Sampling term as an advantage for less inbreeding and more genetic variability in a breeding program. The Mendelian sampling term (h^2_{MS}) 's heritability was estimated from two sources.

First, from the simulations of the different breeding programs performed, the Mendelian Sampling term was calculated as the difference between the individual's True breeding program and its parent's True breeding average values. With these values, the heritability of the Mendelian Sampling term was estimated with two statistical methods, Restricted Maximum Likelihood (REML) and Gibbs Sampling.

Second, from actual data shared by Aviagen, these data are breeding values estimated from body weight phenotypes in broilers. Aviagen estimated the breeding values from 1 879 680 phenotypes of broiler body weight using a pedigree-based estimation method (BLUP), with sex and pen number as fixed effects. The parents whose breeding values were estimated had at least 30 offspring each.

After estimating the breeding values, 26 186 estimated breeding values corresponded to the candidates chosen for selection in the breeding program. From these 26 186 data, the Mendelian Sampling term was calculated for individuals with information about their breeding value and the breeding values of both their parents. The calculation was done in the same manner as with the simulated data, as the difference between the individual's breeding value and the average of their parents. In the end, 24 015 animals had estimated Mendelian Sampling term values ready for analysis.

The heritability estimation was done using programs from the BLUPF90 family programs, specifically, Airemlf90, which runs the Restricted Maximum Likelihood algorithm, and Gibbsf90, Markov Chain Monte Carlo algorithm Gibbs Sampling. BLUPF90 programs are a family of programs used for mixed model computations (Misztal et al., 2015). The variance components estimation was performed using these programs using a univariate and bivariate model analysis.

The univariate model used was as follows:

$$\mathbf{y_{MS}} = \mathbf{1}\boldsymbol{\mu} + \boldsymbol{Za} + \boldsymbol{e} \qquad (6)$$

Where $\mathbf{y_{MS}}$ is the vector of the values of Mendelian Sampling from the estimated breeding values, $(\boldsymbol{\mu})$ is the estimated mean for Mendelian Sampling, $(\mathbf{a})$ is the vector for the random effect, and $(\boldsymbol{e})$ is the vector of estimated associated error. $(\boldsymbol{Z})$ is the incidence matrix associated with the vector $(\mathbf{a})$.

Heritability estimation was performed for the phenotypes of body weight provided by Aviagen for validation of the heritability of the character. Heritability was estimated initially from the 1 000 000 data by Aviagen was equal to 0.35. The model used for the estimation included one fixed effect with 6007 levels:

$$\mathbf{y_{BWT}} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \qquad (7)$$

Where $\mathbf{y_{BWT}}$ is the vector of the phenotypes for body weight, $(\mathbf{b})$.is the vector for the fixed effect (a super factor capturing the combined effect of sex, hatch, mating group, and pen), $(\mathbf{a})$ is the vector for the random effect, and $(\boldsymbol{e})$ is the vector of estimated associated error. $(\boldsymbol{Z})$ Moreover, $(\mathbf{X})$ are incidence matrices related to vectors $(\mathbf{a})$ and $(\mathbf{b})$.

The bivariate model is as follows:

$$\mathbf{y_{MS}} = \mathbf{1}\boldsymbol{\mu} + \mathbf{Z_1 a_1} + \mathbf{e_1} \qquad (8)$$
$$\mathbf{y_{BWT}} = \boldsymbol{Xb} + \mathbf{Z_2 a_2} + \mathbf{e_2}$$

Where $\mathbf{y_{MS}}$ and $\mathbf{y_{BWT}}$ are the vectors for Mendelian Sampling and the phenotype for body weight, $\boldsymbol{\mu_1}$ is the estimated mean for Mendelian Sampling, $(\mathbf{b})$ is the vector of fixed effects as described in equation (7), $(\mathbf{a_1})$ and $(\mathbf{a_2})$ are the vectors for the random effects, and $(\mathbf{e_1})$ and $(\mathbf{e_2})$ are the vectors of estimated associated error. $(\mathbf{Z_1})$ and $(\mathbf{Z_2})$ are incidence matrices related to the vectors $(\mathbf{a_1})$ and $(\mathbf{a_2})$, and finally, $(\mathbf{X})$ is the incidence matrix.

After calculating Mendelian Sampling values from the estimated breeding values provided by Aviagen, three files are needed to run the BLUPF90 programs; the data file contains all the

information required for the model of choice; in the case of the univariate model, the Mendelian Sampling term. The file had the animal ID, the values of Mendelian Sampling, and other population information. The pedigree is required, and a parameter file that groups all the information to be renumbered by RENUMF90. The program series only accepts numerical values, and RENUMF90 does the renumbering process to analyze the data with the other programs (Masuda, 2019). Three output files are generated, where renf90.par file is the parameter file used for both, Airemlf90 and Gibbsf90 programs is variance component estimation. The files for RENUMF90 were modified accordingly to proceed with the bivariate analysis of the model. The file Renf90.inb contains the results of the inbreeding coefficient with the population mean calculated from the pedigree.

For the Bayesian analysis performed with Gibbsf90, 1 000 000 iterations were done to estimate the Variance and heritability, and the Highest Probability Density (HPD) was calculated.

Results

In this project, it was sought to use simulations through AlphaSimR to evaluate the performance of a breeding program when considering the Mendelian Sampling term of an individual as the selection criteria and the Mendelian Sampling term when used alongside the gametic variance of an individual $\sigma^2_{\text{gamete (i).}}$

The following is a presentation of the results of these simulations, where genetic mean, genetic variance, Mendelian Sampling variance, inbreeding, and the number of QTL loci within a population were analyzed.

In order to evaluate the performance of each of these selection schemes, each population was compared with a population where the True Breeding value of an individual was adopted as the selection criterion.

Genetic Gain:

The accumulated genetic response and the rate of genetic gain were calculated as the mean of the True Breeding values of each population.

Figures 1 and 2 show the genetic trend in each population. It can be noted that all four schemes resulted in an increase in genetic gain, and there was overall genetic progress.

As expected, the population with the True Breeding value as the selection criterion presented the highest cumulative genetic response.

Figure 1: The average cumulative genetic response of the simulated breeding schemes for (TBV) True breeding value as selection criteria and (MS) Mendelian Sampling term as selection criteria.

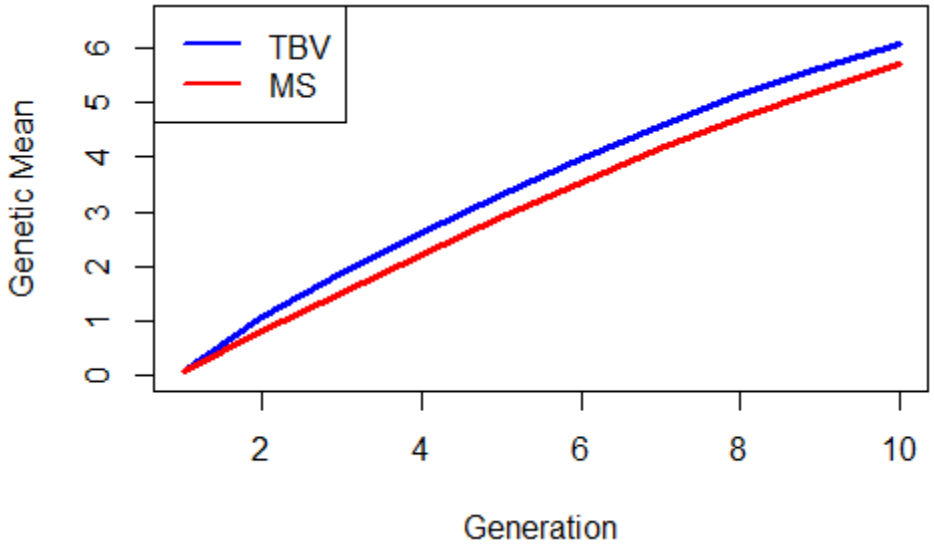
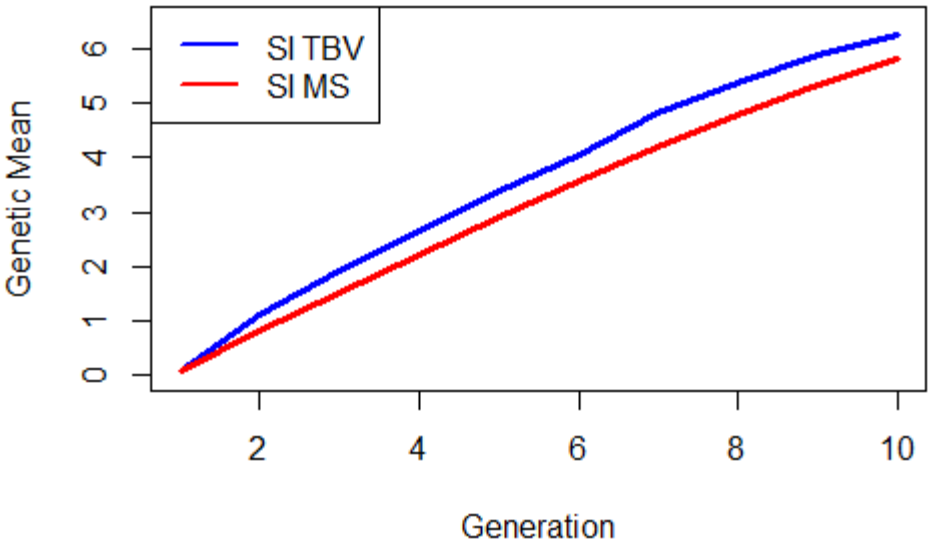


Figure 2: The average cumulative genetic response of the simulated breeding schemes for (SI TBV) True breeding value with gametic variance as selection criteria and (SI MS) Mendelian Sampling term with gametic variance as selection criteria.



The percentages of the rates of genetic gain (**ΔG**) per generation in selection schemes compared with the True breeding value (TBV) program are found in table 2. As well as for the accumulated genetic responses from ten generations of selection.

Table 2: Genetic progress and standard error observed after ten generations of selection in simulated breeding schemes. The selection methods performed in these schemes were: (TBV) True breeding value, (MS) Mendelian Sampling term, (SI TBV) True breeding value with Gametic variance σ^2_{gamete}, (SI_{MS}) Mendelian Sampling term with Gametic Variance σ^2_{gamete}.

Selection scheme	G_{10}	s.e	% Deviated from TBV	ΔG^1 9-10	%Deviated from ΔG 9-10[2]
TBV	6.08	± 0.04		0.44	
MS	5.69	± 0.05	-6.43	0.485	8.5
SI TBV	6.25	± 0.04	2.79	0.461	4.2
SI MS	5.81	± 0.067	-4.56	0.477	6.8

1: The difference between genetic means of the last two generations of selection (9-10). 2: The percentage of genetic gain deviated from the true breeding value program of the three other breeding programs.

It was expected that selecting the True Breeding value would result in the highest levels of genetic gain, which was the case. When selecting the Mendelian Sampling term, there was a 6.43% loss in cumulative genetic gain; this result was in accordance with Grundy et al., 1998. When the gametic variance σ^2_{gamete} was added to the selection scheme, combined with the True Breeding value, it gave rise to a ~ 3% increase in genetic gain compared to the True Breeding value alone.

It is worth noting that when the gametic variance σ^2_{gamete} was added to the Mendelian Sampling term, the loss in cumulative genetic gain decreased by 4.56% less than the Mendelian Sampling term alone. The gametic variance σ^2_{gamete} in the simulation added more genetic progress to both selection programs.

Overall, the difference between the simulated selection schemes was small; there was expected less genetic progress when True breeding values were not included in the selection scheme.

In a later generation, even though the rate of genetic gain (ΔG) decreased over time, the rates in all the selection schemes were higher than in the True Breeding value (TBV) scheme, and that indicated there was a more available response to selection than (TBV) program. The Mendelian Sampling term (MS) program showed a difference of nearly ~9% in the rate of genetic gain (ΔG), followed by the (SI_{MS}) program with nearly ~7%. These programs provided the possibility of a more long-term response to selection than (TBV), where the rate decreased faster than the others.

Additive Genetic Variance σ^2_a:

Like genetic gain, the Additive Genetic variance (σ^2_a) was calculated similarly from the true breeding values of individuals of each population. Figures 3 and 4 show the trend of the additive genetic variance throughout all the populations. It was anticipated that there would be an overall decrease in variability due to selection. The following figures (3) and (4) present how the genetic variance has decreased. It is clear that with the True Breeding value (TBV) program, the genetic variance was the lowest due to the high pressure on selecting loci with favorable effects, leading to a faster change in allele frequencies and faster loss in variability.

Figure 3: Additive Genetic Variance σ^2_a observed after ten generations of selection in simulated breeding schemes. The selection methods in these schemes were: (TBV) True breeding values as a selection criterion, (MS) Mendelian Sampling term as the selection criterion.

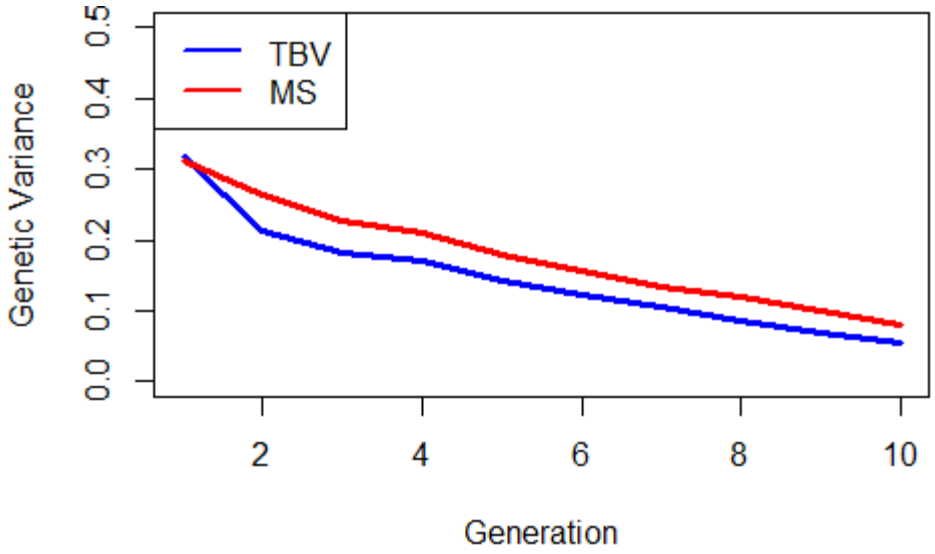


Figure 4: Additive Genetic Variance σ^2_a observed after ten generations of selection in simulated breeding schemes. The selection methods in these schemes were: (SI TBV) True breeding values with the gametic variance σ^2_{gamete} as a selection criterion, (SI_{MS}) Mendelian Sampling term with the gametic variance σ^2_{gamete} as the selection criterion.

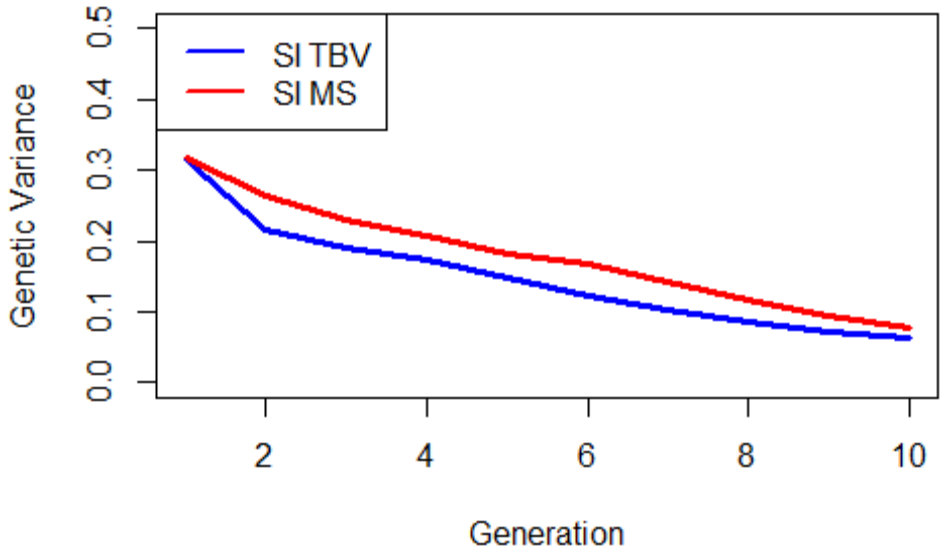


Table 3 presents the average additive genetic variances after ten generations of selection.

There were higher additive genetic variance (σ^2_a) levels in all breeding programs than in the True Breeding value (TBV) program. When selecting the Mendelian Sampling term (MS), there was nearly ~ 46% more genetic variability than (TBV); this result was consistent with Grundy et al., 1998.

Table 3: Additive Genetic variance and their standard error after ten generations of selection. The selection methods performed were: (TBV) True breeding value, (MS) Mendelian Sampling term, (SI TBV) True breeding value with Gametic variance σ^2_{gamete}, (SI_{MS}) Mendelian Sampling term with Gametic Variance σ^2_{gamete}.

Selection scheme	Gen 10	s.e	% From TBV
TBV	0.055	± 0.002	
MS	0.08	± 0.003	45.45
SI TBV	0.062	± 0.001	12.7
SI MS	0.077	± 0.002	40

The loss of genetic variability was the slowest in the programs that included the Mendelian Sampling term in the selection process. The change of allele frequency and the fixation of specific alleles in the populations was slower with Mendelian Sampling (MS) compared to the True Breeding value (TBV); this can be seen in table 6.

Overall, the simulated breeding programs showed better sustainability of genetic variability and more available variability for future generations.

The Mendelian Sampling Variance:

Thompson et al., 1994, termed the Mendelian Sampling variance as the new genetic variation created each generation. After calculating the Mendelian Sampling term from equation (2), the Mendelian Sampling variation was calculated in this study. As mentioned before, in an outbred population where mating is random (i.e., inbreeding F = 0), the Mendelian Sampling variance is equal to half the genetic variance of the population $\sigma^2_{ms} = 0.5\ \sigma^2_a$; in figures 5 and 6, the Mendelian Sampling variance at Generation 0 is equal to 0.17, which is half of the genetic variance σ^2_a at Generation 0. As selection continues in the populations, the Mendelian Sampling variance decreases. Because inbreeding is no longer equal to zero, the Mendelian Sampling variance is equal to $\sigma^2_{ms} = 0.5\ [1\text{-}0.5\ (F_s + F_d)]\ \sigma^2_{a,}$ where (F_s) and (F_d) are the inbreeding coefficients of the individuals (*i*) parents.

These figures have shown that the inclination of the variance in all populations is decreasing. The True Breeding value (TBV) showed the lowest levels of Mendelian Sampling variance compared to other populations. These figures show that the simulated population with the True Breeding value in the selection process has the least genetic variation created each generation. This decrease could be due to the increased inbreeding and loss of alleles leading to a decrease in variability.

Figure 5: Mendelian Sampling Variance σ^2_{ms} observed after ten generations of selection in simulated breeding schemes. The selection methods in these schemes were: (TBV) True breeding values as a selection criterion, (MS) Mendelian Sampling term as the selection criterion.

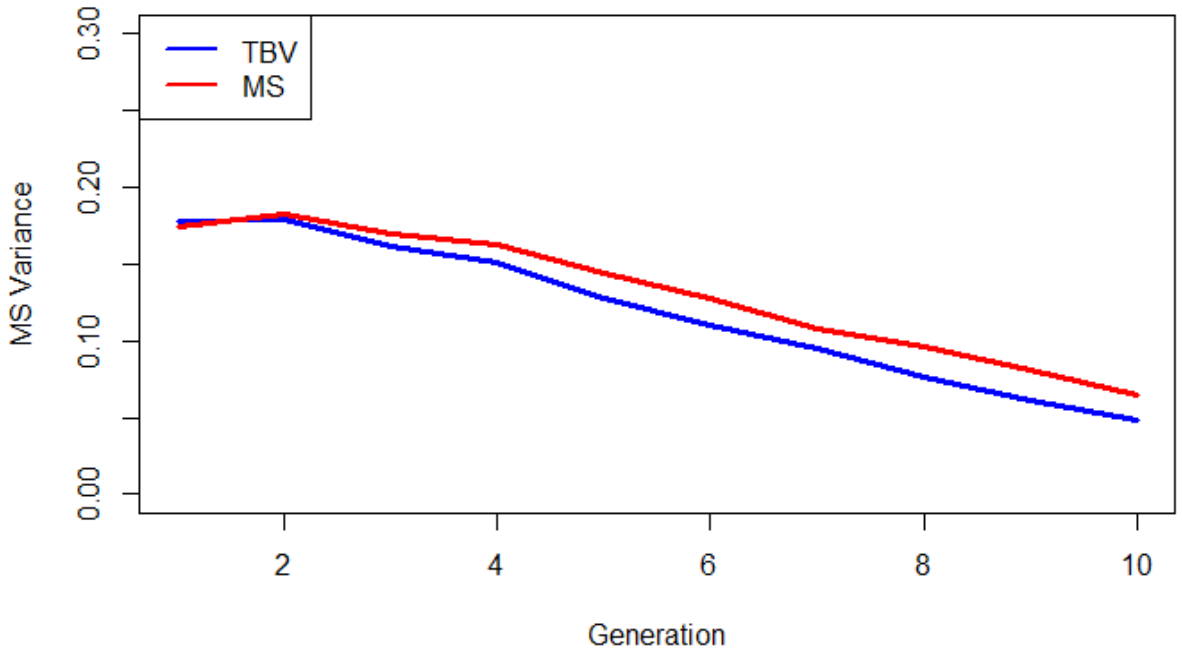


Figure 6: Mendelian Sampling Variance σ^2_{ms} observed after ten generations of selection in simulated breeding schemes. The selection methods in these schemes were: (SI TBV) True breeding values with the gametic variance σ^2_{gamete} as a selection criterion, (SI_{MS}) Mendelian Sampling term with the gametic variance σ^2_{gamete} as the selection criterion

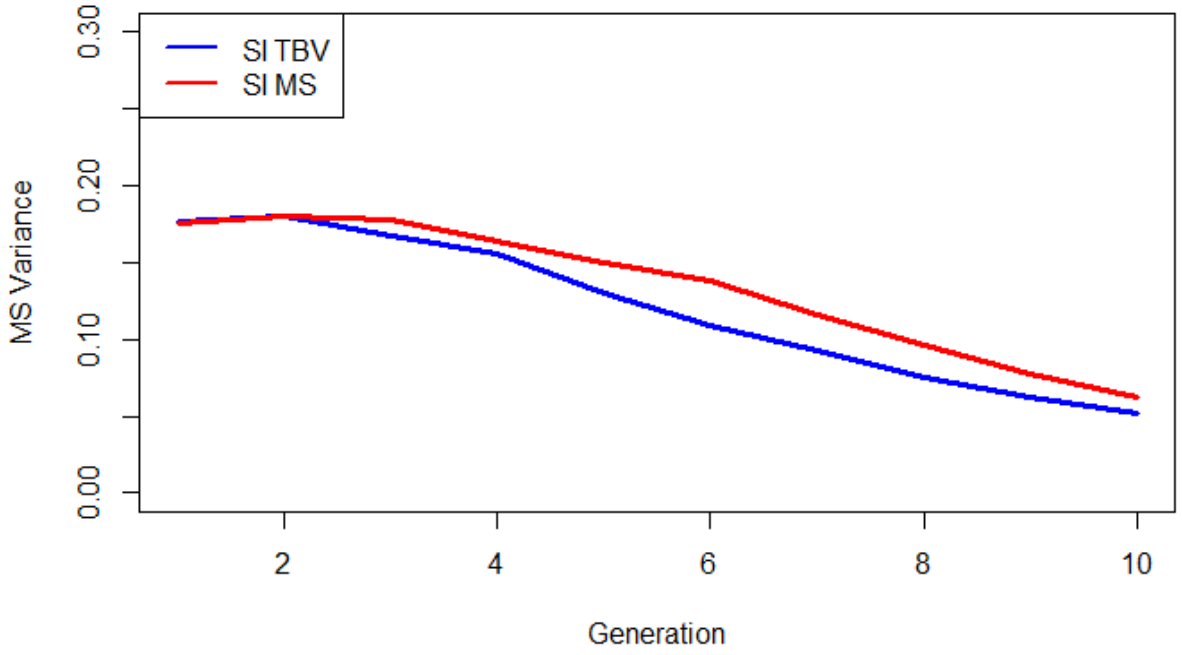


Table 4: Mendelian Sampling variance and their standard error after ten generations of selection. The selection methods performed were: (TBV) True breeding value, (MS) Mendelian Sampling term, (SI TBV) True breeding value with Gametic variance σ^2_{gamete}, (SI_{MS}) Mendelian Sampling term with Gametic Variance σ^2_{gamete}

Selection scheme	Gen 10	s.e	% From TBV
TBV	0.048	± 0.001	
MS	0.065	± 0.002	34
SI TBV	0.051	± 0.001	6.62
SI MS	0.061	± 0.001	27

In table 4, it is clear that the Mendelian Sampling term (MS) program showed the highest levels of Mendelian Sampling term variance, followed by the program where Mendelian Sampling and Gametic variance σ^2_{gamete} were the selection criteria (SI_{MS}). The Mendelian Sampling term (MS) program gave 34% more variance than the True Breeding value (TBV) program, followed by the Mendelian Sampling with the Gametic variance σ^2_{gamete} (SI_{MS}) with 27% more variance and True Breeding value with Gametic variance σ^2_{gamete} (SI_{TBV}) program with ~ 7% more than the True Breeding value program.

Inbreeding:

In this project, inbreeding (F) and the rate of inbreeding (**ΔF**) were calculated from the pedigree per generation of each breeding program. It is expected that the True Breeding value (TBV) program would result in the highest inbreeding and rate of inbreeding (**ΔF**) among the four programs since selecting for (TBV) applies more pressure on selected regions of the genome with high additive effects, and that would lead to a faster loss in rare alleles and increase of homozygosity between individuals.

In figures 7 and 8, the inbreeding coefficient of all the simulated populations over ten generations of selection showed an increase in the inbreeding, which was expected. However, as mentioned earlier and now seen clearly in the figures, the True Breeding value (TBV program) presented the highest levels of inbreeding among them all. The Mendelian Sampling term (MS) program presented one of the least levels of inbreeding.

Figure 7: Inbreeding coefficient (F) observed after ten generations of selection in simulated breeding schemes. The selection methods in these schemes were: (TBV) True breeding values as a selection criterion, (MS) Mendelian Sampling term as the selection criterion.

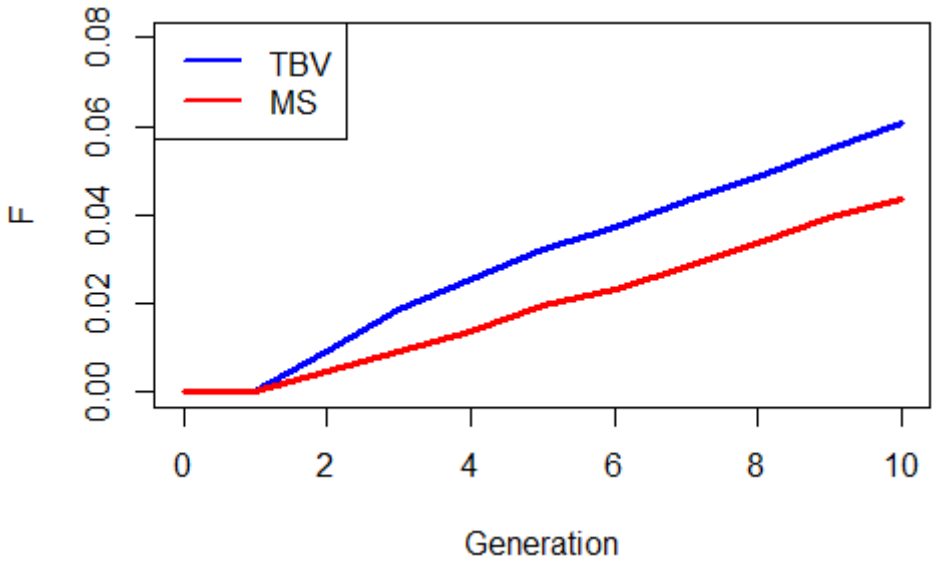
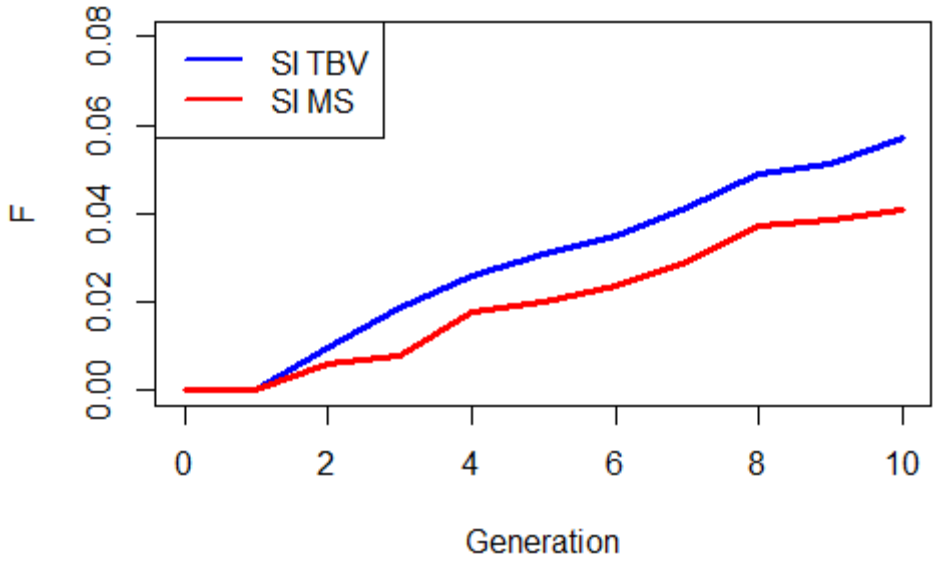


Figure 8: Mendelian Sampling Variance σ^2_{ms} observed after ten generations of selection in simulated breeding schemes. The selection methods in these schemes were: (SI_{TBV}) True breeding values with the gametic variance σ^2_{gamete} as a selection criterion, (SI_{MS}) Mendelian Sampling term with the gametic variance σ^2_{gamete} as the selection criterion



In table 5, the Mendelian Sampling term (MS) program and (SI_{MS}) program showed the least levels of inbreeding, with nearly ~33% less than the (TBV) program and with (SI_{TBV}) having ~12% less inbreeding than (TBV) program.

Table 5: Inbreeding and their standard error and the rate of inbreeding after ten generations of selection. The selection methods performed were: (TBV) True breeding value, (MS) Mendelian Sampling term, (SI TBV) True breeding value with Gametic variance σ^2_{gamete}, (SI MS) Mendelian Sampling term with Gametic Variance σ^2_{gamete}

Selection scheme	Gen 10	s.e	%From TBV	ΔF	%From TBV
TBV	0.065	± 0.003		0.0055	
MS	0.043	± 0.0004	-33.27	0.0043	-22.14
SI TBV	0.057	± 0.002	-12.57	0.0052	-5.35
SI MS	0.042	± 0.0007	-34.46	0.004	-27.24

As for the rate of inbreeding (**ΔF**), the (SI_{MS}) program had the lowest level of inbreeding rate (~27%) compared to the (TBV) program, followed by (MS) program with a ~22% less rate than the (TBV) and (SI_{TBV}) program with nearly ~5% less inbreeding rate.

QTL frequencies:

In this study, the evolution of the frequency of QTLs in each was calculated to investigate the effect of each breeding strategy on the QTL frequencies and the number of segregating QTLs in the simulated population. The frequencies were calculated after a primary quality control consisting of Minor allele frequency > 0.05 (MAF >0.05) was performed to analyze the results.

It is assumed that where there are higher levels of variability, there would be less loss of QTLs and a higher number of QTLs with increased frequencies.

Table 6 shows the results of the number of QTLs in each simulated population, where the average number of QTLs lost after ten generations are represented below,

Table 6: Average number of QTLs lost after ten generations of selection in each breeding scheme.

Selection Scheme	Gen 0	After QC	Gen 10	Lost QTLs
TBV	1000	467	344	123
MS	1000	467	369	98
SI TBV	1000	467	432	35
SI MS	1000	467	443	24

The True Breeding value (TBV) program was the highest in this study, with 123 QTLs lost at generation 10. (SI_{MS}) showed the least loss of QTLs with 24 compared to the other programs, followed by (SI_{TBV}) with 35 QTLs lost. As for the Mendelian Sampling term (MS) program, there were 98 QTLs lost at generation 10.

The (SI_{MS}) program increased the rare allele frequency in the population. It was the most successful in preserving the QTLs, leading to better variability and a better potential for future response selection.

h^2_{ms} estimations:

Table 7 shows the results of the two models (Univariate and Bivariate) from equations (6), (7), and (8), where Aviagen provided the data for analysis.

Table 7: Univariate analysis of the Bodyweight and the calculated Mendelian Sampling term, with variance components estimation, using REML and Gibbs Sampling from data shared by Aviagen.

REML	**h^2**	**[3]Vg**	**[4]Ve**
MS (1)[1]	0.05	1.252	19.808
BWT (2)[2]	0.28	46.8	0.9

Gibbs Sampling	**h^2**	**[5]HPD**	**Vg**	**Ve**
MS (1)	0.075	0.034 – 0.08	1.61	19.68
BWT (2)	0.28	0.24-0.31	32.427	83.45

[1]MS: Mendelian Sampling term; [2]BWT: Bodyweight; [3]Vg: Genetic Variance;[4] Ve: Residual Variance; [5]HPD: Highest probability density

The heritability of the initial character, the Bodyweight in broilers, is equal to 0.28 with an HPD of 0.24-0.31. Both methods led to similar estimations. As mentioned earlier, the initial estimate of the heritability of body weight for the population is 0.35, calculated from 1 879 680 data points.

Regarding the estimation of the heritability of Mendelian Sampling, it was worth noting that there has been no estimation of the Mendelian Sampling term heritability from actual data; there is no bibliographical information to rely upon in the assessments. It was assumed that when there is zero inbreeding, the Mendelian Sampling term variance would be equal to half of the genetic variance, meaning that at the reference population level, where inbreeding is zero (F=0), the heritability of the Mendelian Sampling term would be equal to half of the heritability of the initial character, and when selection is performed inbreeding increases and both variances (Genetic and the Mendelian Sampling term) would start to decrease.

The heritability estimated was relatively low, at 0.07; that indicated that most variation is not genetic, but there was genetic variation, nonetheless.

As for the simulated data, the same procedure was performed for each model; the heritability of Mendelian Sampling was estimated from 20 000 simulated data points to compare the result with the actual data above. The results are shown in Table 7 for both the original simulated trait and the calculated Mendelian Sampling term from the simulated True Breeding values.

Table 8: Univariate analysis of data from the simulated breeding program (TBV), variance components estimation, and heritability estimation.

REML	**h^2**	**[3]Vg**	**[4]Ve**
[1]MS	0.09 ± 0.0032	0.05 ± 0.0007	0.5 ± 0.008
[2]Trait	0.31 ± 0.002	0.31 ± 0.002	0.79 ± 0.07

Gibbs Sampling	**h^2**	**HPD**	**Vg**	**Ve**
MS	0.1 ± 0.034	0.05 - 0.13	0.045 ± 0.03	0.4 ± 0.12
Trait	0.3 ± 0.09	0.28 - 0.32	0.45 ± 0.1	1.05 ± 0.022

[1]MS: Mendelian Sampling term; [2]Trait: Simulated trait with AlphaSimR; [3]Vg: Genetic Variance;[4] Ve: Residual Variance

Similar results were obtained from the real data analyzed, which provided a preliminary indication that the heritability of the Mendelian Sampling term might be around 0.1.

Notably, many traits in animal breeding present low heritability, like reproductive traits (e.g., litter size, fertility) and health traits (e.g., disease resistance, mastitis). The results indicated genetic

control over the Mendelian Sampling term, allowing for genetic progress with the Mendelian Sampling term as a selectable trait, even though a limited h^2 of the trait restricts the rate of genetic progress achieved. This estimation provided an initial concept into the possibility of selecting for the progeny's superiority over its parent average.

In table 9, the bivariate analysis results for the variance components estimation. The model used was the same as equation (8) described in the Materials and Methods section.

Table 9: Bivariate analysis of the Bodyweight phenotypes and the calculated Mendelian Sampling term, with variance components estimation with REML and Gibbs Sampling from data shared by Aviagen.

REML	**h^2_{ms}**	**h^2_{BWT}**	**Vg**	**Ve**	**[1]rg**		
MS/BWT	0.07 ± 0.0001	0.27 ± 0.0003	1.5 / 32	19.6 / 84	0.65		
Gibbs Sampling	**h^2_{ms}**	**HPD**	**h^2_{BWT}**	**HPD**	**Vg**	**Ve**	**rg**
MS/BWT	0.068	0.045 - 0.09	0.27	0.24 - 0.3	1.39 / 31.4	19 / 84.15	0.58 (0.47 - 0.65)

[1]Rg: Genetic Correlation between Mendelian Sampling and Bodyweight

The heritability of the Mendelian Sampling term from the model (8) presented a heritability similar to that of the univariate analysis, of 0.07. In addition, the correlation between the phenotypes and the calculated Mendelian Sampling term is 0.65.

Discussion

In this study, the effects of selecting the Mendelian Sampling term and gametic variance σ^2_{gamete} combined in a breeding program were investigated via simulations. The evaluation of each breeding program was based on the analysis of the cumulative genetic gain and rate of genetic gain, calculated as the mean of the genetic values in each generation and the difference in means of two consecutive generations.

The additive genetic variance, the Mendelian Sampling variance, inbreeding rate from pedigree, and the change in allele frequencies of QTLs were also evaluated and compared in each program.

Results showed a general increase in genetic gain after ten generations of selection and a decrease in additive genetic variance due to selection as expected from the theory of quantitative genetics. These evaluations aimed to identify which program led to the highest levels of genetic variability, which means the highest additive genetic variance while limiting the loss in genetic gain and controlling inbreeding.

The True Breeding Value breeding program had the highest cumulative genetic gain compared to the other programs. This result was expected since the selected candidates have the program's highest breeding values. That meant that the genetic gain of the other simulated programs was compared to the breeding program of the True Breeding value since it represents the highest possible cumulative genetic gain that could be achieved between all the breeding programs.

Selection resulted in an increase in the frequency of QTLs with the highest favorable effect, which led to an increase in cumulative genetic gain. The Mendelian Sampling term program and the selection scheme of Mendelian Sampling term and gametic Variance σ^2_{gamete} (SI_{MS}) showed the lowest levels of genetic gain after ten generations of selection. Also, it is worth noting that the program of the True Breeding value and the gametic variance σ^2_{gamete} (SI_{TBV}) combined led to the same levels of genetic gain while presenting a higher amount of additive genetic value. It was approximately 3% higher than the True Breeding program.

As for the rate of genetic gain ($\mathbf{\Delta G}_{\mathbf{9\text{-}10}}$), the highest value was found with the Mendelian Sampling program (MS), with 8.5% more ($\mathbf{\Delta G}_{\mathbf{9\text{-}10}}$) than the True Breeding value program, and

the two other programs also showed higher rates (~ 5% and 7%). This suggests that compared to the True Breeding program, these breeding programs present more long-term selection responses.

This result was consistent with Grundy et al., 1998, who used the Mendelian Sampling term in their selection process. The indices gave less genetic gain but a higher rate of genetic gain ($\mathbf{\Delta G}$).

This higher rate of genetic gain between the True Breeding program and the other breeding programs was related to the closeness in the relationship between the candidates selected for the next generation. The True Breeding Value exerted more pressure on choosing more related relatives in the selection process, increasing the cumulative genetic gain but restricting the genetic gain rate from one generation to another.

The Mendelian Sampling term represents the possibility of selecting the best full sibling within a family, which leads to less pressure on selecting candidates more related to their parents, giving rise to fewer candidates that resemble each other as parents of the next generation. It can be noticed that when combining the true breeding value with the gametic variance σ^2_{gamete}, the response was positive. There was no loss in genetic gain, and there was an increase in the levels of genetic variability, which created more opportunities for long-term selection. This breeding program (SI_{TBV}) allowed for the selection of candidates with the highest true breeding values but also allowed for the selection of candidates with the highest heterozygous loci effect, which after ten generations of selection has increased the frequency of rare favorable alleles and led to no loss in genetic gain and a higher level of genetic variability than the True Breeding value program.

There was expected to be an overall decrease in genetic variability within a population due to selection. The decrease was also noticed in the Mendelian Sampling variance. Based on calculations earlier from the simulated data, the evolution of the Mendelian Sampling variance was a function of the additive genetic variance σ^2_a and inbreeding (F). As the additive genetic variance σ^2_a decreases and inbreeding (F) increases, the Mendelian sampling variance σ^2_{ms} decreases.

The most significant drop in additive genetic variance σ^2_a was observed in the first generations of selection. That is due to an effective selection of candidates with the highest breeding values, which means directly selecting the QTLs with the highest effects. The additive genetic variance σ^2_a was almost similar in all the selection schemes in the early stage. However, it would be reduced the most in the most efficient selection method, which in this case is the True Breeding value program. This reduction in genetic variability is due to selection that changed allele frequencies, leading to the loss of alleles and the fixation of others. Another motive for reducing variability was the Bulmer effect due to negative covariances between loci (Bulmer, 1970).

The Mendelian Sampling term program presented the highest value of the additive genetic variance σ^2_a, followed by (SI_{MS}) and (SI_{TBV}), which indicated that the Mendelian Sampling program maintained the most genetic variability in the simulated population allowing for the opportunity for long-term response to selection.

Thompson et al., 1994, postulated that the Mendelian Sampling term in a selection index could significantly improve the long-term genetic response since a considerable amount of inbreeding (F) is avoided, and a higher level of additive genetic variance σ^2_a is maintained. This is seen in this study, where the Mendelian Sampling term program led to the lowest inbreeding (F) levels at generation ten. Both the Mendelian Sampling term and (SI_{MS}) program led to the lowest rate of inbreeding ($\mathbf{\Delta F}$).

In breeding programs, the long-term consequence of genetic variability reduction is usually ignored (Daetwyler et al., 2007), but high ($\mathbf{\Delta F}$) usually has a more immediate effect giving rise to inbreeding depression and an increase of recessive deleterious alleles as selection progresses. In a later generation, the impact of inbreeding becomes an increasing factor in reducing additive genetic variance. So, an association between increased genetic gain and increased inbreeding (F) while the additive variance σ^2_a decreases has been observed. One of the purposes of these simulations in this study is to investigate the possibility of weakening that association by reducing the decrease of genetic variability due to selection while limiting the loss of genetic gain and controlling inbreeding. It is worth noting that selecting the Mendelian Sampling term, in other words, for the superior sibling to its parent average, led to a decreased inbreeding rate

($\mathbf{\Delta F}$). That is due to a better differentiation between full siblings and a reduction in the co-selection of relatives (Chen et al., 2009).

Including the gametic variance, σ^2_{gamete} has led to an increase in the cumulative genetic gain, a reduction in the loss of additive genetic variance σ^2_a, and fewer rates of inbreeding ($\mathbf{\Delta F}$). The reason is that the gametic variance σ^2_{gamete} offered the possibility of adding more weight to the heterozygous loci, which led to more genetic variability within a population where individuals with the highest heterozygous effect were selected.

It should be noted that the two simulated programs that included the gametic variance σ^2_{gamete} tended to increase the frequency of the rare favorable alleles in the population since they showed the smallest loss in the number of QTLs (35 for SI_{TBV} and 24 for SI_{MS}), which led to an increase in genetic variability while controlling the loss in genetic gain.

These results show that the True Breeding value program focuses more on a genome subset that rapidly changes the allele frequencies. In contrast, the other programs spread the selection pressure more evenly across the genome. This indicates that the long-term response can be improved by modifying the selection pressure on a QTL as its allele frequency changes (Wientjes et al., 2022).

Furthermore, the results show that selecting the Mendelian Sampling term or even including the term in the selection process has led to a more sustainable breeding program in terms of maintaining future genetic fain while controlling the rate of inbreeding within a population. Selecting the Mendelian Sampling term meant choosing candidates according to their independent and unique superiority or inferiority with respect to the parenteral average. It was proven empirically that to maximize candidate contribution and therefore maximize genetic gain while controlling inbreeding (F), Mendelian Sampling proved to be an advantage when using quadratic indices to optimize contribution (Avendaño et al., 2004). These indices generally manage genetic gain and inbreeding in a breeding program.

Thompson et al., 1994, proved that genetic gain is related to the covariance between long-term genetic contribution and the Mendelian Sampling term.

From the simulated breeding programs, it has been shown that including the Mendelian Sampling term in the selection scheme could be beneficial.

It should be noted that the results analyzed were from simulated breeding programs, where the true breeding value and not the estimated breeding values were used to interpret the results of each simulation. It should be acknowledged that in practical situations, the estimated breeding value would be used, and the calculation of the Mendelian Sampling term will depend on how well the accuracies of these estimations are. That said, the impact of the Mendelian Sampling term on the genetic gain, genetic variance, inbreeding, and the number of segregating loci in practical situations might be less or more emphasized. Caution is needed when using the estimated Mendelian Sampling term since it depends on the accuracy of the estimated breeding value of the animal and its parents.

Finally, from the data provided by Aviagen, the Mendelian Sampling term calculation of each member in the population and the estimation of the heritability (h^2_{ms}) were done by REML and Gibbs Sampling (h^2_{ms} = 0.05-0.07). These estimations provide preliminary evidence regarding the heritability of the Mendelian Sampling term from actual and simulated data, which has shown to be in function of the additive genetic variance from previous calculations.

There have not been any previous studies regarding estimating the heritability of the Mendelian Sampling term with real data. However, this study provided preliminary findings that the Mendelian Sampling term is heritable and could be considered a selection criterion in future breeding programs for the possibility of decreasing the loss of genetic variability and the control inbreeding to improve the long-term response to selection.

The breeding values brought forth by Aviagen were estimated from more than one million data points. The breeding values used for the calculations in the simulated breeding programs were the true breeding values and not estimated as the real data. It should be brought to light that two types of data were used and led to similar estimations; no certain conclusions can be deduced from the difference between the two types of data, but it is worth noting concerning the similarity of the two estimations.

Future implications:

The results showed better additive genetic variance and variability when using the Mendelian Sampling term or the gametic variance σ^2_{gamete}. In terms of relevance to breeding programs, the results would be used as a component to increase the genetic variability in a population. However, emphasis on the accuracy of estimation of breeding values of animals should be noted since the Mendelian Sampling term, and its impact are highly dependent on the accuracy of estimation of these values. Other considerations could be made, such as the Mendelian Sampling term, since selecting for it showed increased genetic variability in a simulated breeding program. It could be regarded as a quantifiable component for other factors responsible for creating variability, for example, recombination rate. The recombination rate is known to increase genetic variability through crossing-overs that create new combinations of alleles, which means more opportunities for selection (Gonen et al., 2017).

Moreover, an investigation of the relationship between recombination events and the Mendelian Sampling term could be put forth to explore the possibility that the Mendelian Sampling term might capture part of the variability created by recombination and could be a proxy for selection for higher recombination events in order to increase genetic variability in a population.

It is worth noting that the estimated heritability of the recombination rate is around 0.05-0.07 (Johnsson et al., 2021). In this study, preliminary estimations of the Mendelians were conducted and resulted in a low heritability of 0.07 as well.

In a study conducted by Sosa-Madrid et al. 2022, they estimated the additive genetic variance from data of a poultry breeding program. The results showed constant levels of genetic variance over twenty generations of selection. It is essential to mention that the data analyzed were from a large subset of data from a population selected for various traits. In the results analyzed from real data, the genetic correlation between the Mendelian Sampling term and body weight was equal to 0.65, which might provide evidence or explanation for the consistent levels of additive genetic variance in the population over a considerable period of time. This is just an assumption, and more investigation into the idea that the Mendelian Sampling term might explain part of the phenomenon of the unchanged levels of genetic variance; of course, multiple other factors are in play, but the Mendelian Sampling term might be one of them.

Finally, since the Mendelian Sampling term and the gametic variance σ^2_{gamete} have increased the genetic variability in the simulated population. These two components might present the opportunity for identifying animals for high-density genotyping when the purpose is to obtain more genetic diversity between the genotyped animals.

For example, in a breeding company, before selecting the candidates to be genotyped, a preselection is performed to increase the chance of choosing the best animals available for the training set of a population. Usually, the preselection process relies on the highest estimated breeding values from pedigree information or selecting the least related animals in a population. Choosing these animals based on the former would lead to selecting the best animals in the population, and the latter might not include the best animals. In both cases, it is not selecting the most diverse animals. It was assumed from this study that Mendelian Sampling terms and gametic variance σ^2_{gamete} maintain higher levels of genetic variability, meaning it selects the individuals that present more variability in a population. Hence, there is a possibility that the Mendelian Sampling term and the gametic variance σ^2_{gamete} could be used to identify the most variable individuals in a population and be utilized in selecting candidates for genotyping, representing a sample of the populations' diverse genes.

Table 100: Summary of the results of the simulated breeding programs evaluation criteria

Breeding program evaluation criteria	Simulation results	Likely mechanism
Rate of genetic gain	- The overall increase in genetic gain - The true Breeding value program presented the highest gain - The genetic gain rate was the highest in the Mendelian Sampling term program, followed by the SI MS. In addition to the Mendelian Sampling, the selection process involved the gametic variance. - The True Breeding value selection program saw the lowest rate of genetic gain.	- Increase genetic gain due to selection, efficient selection of QTLs with the highest effect. - The True Breeding value program exerted more pressure on selecting candidates that resembled each other; the Mendelian Sampling term better differentiated the superior sibling for selection which led to less co-selection of relatives
Additive genetic variance	- A general decrease in additive genetic variance - A significant drop in additive genetic variance in the first few generations - The lowest decrease was observed in the Mendelian Sampling term program	- Change in allele frequencies due to selection, leading to loss of alleles and fixation of others. - Effective selection of candidates with the highest breeding values. - There was reduced additive genetic variance due to the Bulmer effect. - There is less pressure on selecting individuals that resemble, leading to fewer candidates that are relatives and increased variability in a population.
Inbreeding	- The minor inbreeding levels were seen where the Mendelian Sampling term was included in the selection process.	- Better differentiation between full siblings led to a decrease in the co-selection of relatives
Number of QTLs	- The lowest loss in QTL number was in the SI TBV and SI MS simulations.	- The gametic variance increased the frequency of the rare favorable alleles, which decreased the loss of alleles in a population and increased genetic variability.

Conclusion

- A common approach in animal breeding is to focus on selection strategies to increase genetic gain, which can lead to an increase in inbreeding. Even though the simulations conducted in this study presented less progress in the short term, it showed that considering the variability of future generations leads to a higher long-term response to selection and lower rates of inbreeding.
- Mendelian Sampling term provided a new approach for the conversation of long-term response to selection by presenting higher levels of genetic variability and lower inbreeding in a simulated population.
- The heritability of the Mendelian Sampling term is in relation to the additive genetic variance and inbreeding; the estimated heritability is 0.05-0.07; this conveys that there is an opportunity of including the Mendelian Sampling term in the selection process of a breeding program in order to increase the variability of a population.
- The gametic variance allowed to select individuals with the most variable QTL profile—selecting the candidates with the highest heterozygous loci effect led to a decrease in genetic gain loss and an increase in additive genetic variance. Santos et al., 2019, provided evidence on the feasibility of estimation and application of the gametic diversity in a breeding program. It can be obtained from genomic models where SNP effects are estimated and used to control genetic diversity.
- This study proposed formulas to emphasize the genetic variability in a population to increase long-term progress and preserve genetic variability with less reduction of short-term progress.
- The proposed breeding programs (SI_{MS} and SI_{TBV}) are easy to obtain and apply. With greater genetic diversity and better genetic progress preservation, these programs offer an opportunity to minimize additive genetic variance loss and better control inbreeding. The accuracy of estimation of the breeding values is a very important component in the impact of the Mendelian Sampling term and gametic variance.
- There was more genetic variation in the population and slower loss in allele frequency throughout the selection process by selecting the progeny superior to its parent's average.

- The simulated population presented less inbreeding than the population selected for the highest true breeding value.

References

Avendaño, S., Woolliams, J. A., & Villanueva, B. (2004). Mendelian sampling terms as a selective advantage in optimum breeding schemes with restrictions on the rate of inbreeding. *Genetical Research*, *83*(1), 55–64. https://doi.org/10.1017/S0016672303006566

Bijma, P., Wientjes, Y. C. J., & Calus, M. P. L. (2020). Breeding Top Genotypes and Accelerating Response to Gametic Variance. *Genetics*, *214*(January), 91–107.

Bonk, S., Reichelt, M., Teuscher, F., Segelke, D., & Reinsch, N. (2016). Mendelian sampling covariability of marker effects and genetic values. *Genetics Selection Evolution*, *48*(1), 1–11. https://doi.org/10.1186/s12711-016-0214-0

Bulmer, M. G. (1971). The effect of Selection on Genetic Variability. *The American Naturalist;105:201–11*. http://www.journals.uchicago.edu/t-and-c

Chen, J., Ytournel, F., Erbe, M., Pimentel, E. C. G., & Simianer, H. (2009). Use of Mendelian Sampling Terms in Genomic Models. *January*, 0–3.

Daetwyler, H. D., Villanueva, B., Bijma, P., & Woolliams, J. A. (2007). Inbreeding in genome-wide selection. *Journal of Animal Breedings and Genetics*, *124*, 369–376.

Faux, A., Gorjanc, G., Gaynor, R. C., Battagin, M., Edwards, S. M., Wilson, D. L., Hearne, S. J., Gonen, S., & Hickey, J. M. (2016). AlphaSim: Software for Breeding Program Simulation. *The Plant Genome*, *9*(3). https://doi.org/10.3835/plantgenome2016.02.0013

Goiffon, M., Kusmec, A., Wang, L., Hu, G., & Schnable, P. S. (2017). Improving response in genomic selection with a population-based selection strategy: Optimal population value selection. *Genetics*, *206*(3), 1675–1682. https://doi.org/10.1534/genetics.116.197103

Gonen, S., Battagin, M., Johnston, S. E., Gorjanc, G., & Hickey, J. M. (2017). The potential of shifting recombination hotspots to increase genetic gain in livestock breeding. *Genetics Selection Evolution*, *49*(1), 1–12. https://doi.org/10.1186/s12711-017-0330-5

Grundy, B., Luo, Z. W., Villanueva, B., & Woolliams, T. A. (1998). The use of Mendelian indices to reduce the rate of inbreeding in selection programs. *Journal of Animal Breeding and Genetics*, *115*(1), 39–51. https://doi.org/10.1111/j.1439-0388.1998.tb00326.x

Johnsson, M., Whalen, A., Ros-Freixedes, R., Gorjanc, G., Chen, C. Y., Herring, W. O., de Koning, D. J., & Hickey, J. M. (2021). Genetic variation in recombination rate in the pig. *Genetics Selection Evolution*, *53*(1), 1–19. https://doi.org/10.1186/s12711-021-00643-0

Masuda, Y. (2019). Introduction to BLUPF90 suite programs:Standard Edition. The *University of Georgia. Http://Nce.Ads.Uga.Edu/Wiki/Doku.Php?Id=documentation*, *September* 1–185.

Misztal, I., Tsuruta, S., Lourenco, D., Aguilar, I., Legarra, A., & Vitezica, Z. (2015). BLUPF90 family of programs. *University of Georgia, Athens, USA*, 1–125.

R.A. Mrode, P. (2005). Linear models for the prediction of Animal Breeding values. In *CABI Publishing* (Second edi). CABI Publishing.

Santos, D. J. A., Cole, J. B., Lawlor, T. J., VanRaden, P. M., Tonhati, H., & Ma, L. (2019). Variance of gametic diversity and its application in selection programs. *Journal of Dairy Science*, *102*(6), 5279–5294. https://doi.org/10.3168/jds.2018-15971

Sosa-Madrid, S., Kranis, A.(2022). Estimating additive genetic variance over time in chicken breeding programs. *Proceedings of the 73rd Annual Meeting of EAAP*.

Thompson, J. A. W. and R. (1994). A theory of Genetic contributions. *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, Vol. 19*, *19*, 127–134. https://doi.org/10.4324/9781315514697-10

Wientjes, Y. C. J., Bijma, P., Calus, M. P. L., Zwaan, B. J., Vitezica, Z. G., & van den Heuvel, J. (2022). The long-term effects of genomic selection: 1. Response to selection, additive genetic variance, and genetic architecture. *Genetics Selection Evolution*, *54*(1), 1–21. https://doi.org/10.1186/s12711-022-00709-7