Inbreeding and Genetic Gain in the Presence of Random Mating and Mate Allocation Using Genomic-Pedigree Relationships in Chickens: A Simulation Study

V. Nikoonejad Fard¹, and H. Mehrabani Yeganeh¹*

ABSTRACT

In this simulation study, Mate Allocation (MA) strategy using combined genomic-pedigree information was compared with Random Mating (RM) aiming at controlling the level of inbreeding (ΔF) with minimum impacts on the amounts of Genetic Gain (ΔG) in poultry breeding programs. Five equally-sized subpopulations of chickens (P₁ to P₅) were simulated. A genome encompassing five chromosomes involving 15,000 bi-allelic markers was defined for each bird. Potentially, 500 QTL impacted a trait, which had a heritability of 0.1. Only pedigree information was assumed to be available in P₁ while the percent of genotyped birds were 10% in P₂, 20% in P₃, and 50% in P₄ and P₅. Estimated Breeding Values (EBVs) were computed using the traditional approach (PBLUP) and the Single-Step method (SSGBLUP). In P₅, early predictions were applied to estimate GEBVs. Comparisons were made based on the reductions in ΔF and changes in ΔG between two mating scenarios and two evaluation methods within and across subpopulations, respectively. After seven generations, MA resulted in 20 to 30% less ΔF within subpopulations compared with RM with negligible impacts on ΔG. Furthermore, in both mating scenarios, SSGBLUP brought about 11 to 61% less ΔF compared to PBLUP across subpopulations. Results indicated that the benefits of using combined genomic-pedigree relationships could be more than improving the accuracy of EBVs through the SSGBLUP as they can also be used in mating designs to restrict ΔF with a minimum impact on ΔG. Also, this study verified that SSGBLUP could bring about lower ΔF compared with PBLUP.

Keywords: Estimated breeding values, Mating design, PBLUP, SSGBLUP.

INTRODUCTION

Thanks to the considerable advancements in genomic technologies during the past decades, the traditional strategies in livestock breeding have been progressively replaced with novel approaches that take advantage of genomic information. Nowadays, the genomic selection, (Meuwissen et al., 2001), is broadly accepted as a premier method in breeding programs, especially in dairy cattle owing to its remarkable potential to boost the rates of genetic response (ΔG) through improving the accuracy of evaluations and decreasing the generation interval (Hayes et al., 2009; Schaeffer, 2006). However, although its potential advantages in poultry breeding seem to be promising, it has not been applied in large scales up until now, mainly due to the genotyping costs (Preisinger, 2012; Wolc et al., 2016).

Experiences have shown that genomic selection could improve the accuracy of evaluation, especially in low heritable traits. Moreover, it allows for direct selection in sex-limited traits (i.e., egg production), sex-influenced traits (i.e., body weight), traits...
that take a long time to obtain information (i.e., advanced age performance), and traits that are difficult or expensive to measure (i.e., feed efficiency) (Muir, 2007). Besides, genomic selection enables breeders to reduce the size of the breeding programs and restrain performance testing compared with that of the Pedigree-based method (PBLUP) (Wolc et al., 2015).

From another perspective, mating plans were traditionally designed using pedigree information of selection candidates; nonetheless, better results can be achieved utilizing genomic information. Models that use genomic information are comparatively more capable in capturing and controlling the level of inbreeding (ΔF), preserving the genetic diversity, restricting the frequency of harmful recessive defects, finding out the specific genotypes related to disease resistance, and handling the mating according to the breeding objectives (Daetwyler et al., 2009; Pryce et al., 2012). Therefore, using the proportionate economic strategies to exploit these advantages in current poultry breeding schemes can lead to further progress, reduce the regular rearing costs, and probably counterbalance the investments of genotyping.

In this regard, one way would be to partially genotype the breeding populations for the trait of interests and incorporate this information into pedigree data and phenotypic records. It would be an appropriate strategy, especially in improving the low heritable traits that have shown trivial rates of genetic progress through traditional approaches. Additionally, it is not often feasible to genotype all individuals in breeding populations due to a variety of reasons such as culling, slaughtering, foreign animals, or even high costs of genotyping. Misztal et al. (2009) proposed a unified approach (SSGBLUP) that can estimate Genomic Breeding Values (GEBVs) for both genotyped and un-genotyped animals at the same time through a combined genomic-pedigree relationship matrix (H) (Christensen and Lund, 2010; Legarra et al., 2009). Compared to the traditional multi-step approach, SSGBLUP can bring about higher accuracies for both genotyped and un-genotyped animals (Christensen et al., 2012) as well as less bias in evaluations (Vitezica et al., 2011). The method has also demonstrated its potential to improve the current chicken breeding programs through increasing the accuracy of evaluation and reducing the generation interval (Aleme et al., 2016; Chen et al., 2011; Lourenco et al., 2015). However, the benefits of utilizing the information of H can be more than improving the accuracy of EBVs. As genomic relationships between genotyped individuals are propagated to all relatives, the elements of this matrix represent more accurate estimations of the relationships. Therefore, it can also be used to design more precise mating plans.

This study aimed to examine the hypothesis that the information of H can also be used in a Mate Allocation strategy (MA) to reduce the level of ΔF compared to the commonly used Random Mating scheme (RM). Besides, the accuracy of evaluations, rates of ΔG, and levels of ΔF realized by SSGBLUP have been compared with that of PBLUP assuming different percentages of birds have been genotyped in the medium-scale simulated populations of chickens.

**MATERIALS AND METHODS**

A genome encompassing five chromosomes with equally 1.5 Morgan length (750 cM in total) and 15000 evenly distributed bi-allelic markers with equal allele frequencies in the first generation of the historical population was simulated for each bird. This marker density is used to mimic nearly the usage of the 60k genotyping panel. Potentially, 500 Quantitative Trait Loci (QTL) impacted a trait, which had a heritability of 0.1. For the sake of simplicity, no fixed effect was simulated, and QTL effects were sampled from a normal distribution, assuming that
only the additive effects explain the total genetic variance. Then, a historical population with an effective population size of 100 with equal sex ratio was randomly mated for 50 generations (g). During these generations, the inheritance of the SNPs follows Mendel’s law, where the Haldane mapping function was used to simulate recombination. The population was expanded to 4,000 offspring in g= 51. RM was continued for two additional generations building up an adequate pedigree. At g= 54, the population was expanded into five equally-sized subpopulations (P₁ to P₅), each of which included 4000 offspring. This population size was maintained the same for all of the next generations.

Evaluations were performed in subpopulations through the seven target generations (g= 54 to 60) assuming that only pedigree information is provided for P₁ while the percent of genotyped birds are 10% in P₂, 20% in P₃, and 50% in P₄ and P₅. Use of genomic information in evaluations was started at g= 54. Birds were randomly genotyped with equal proportion in both sexes such that the genomic information for at least one bird from each full-sib family was assumed to be available. Needless to say that genotyping birds regardless of their sexes help in improving predictions (Lourenco et al., 2015). Pedigree information for all individuals and one phenotypic record for each bird were also assumed to be available. True Breeding Values (TBVs) were calculated as:

\[
TBV_i = \sum_{j=1}^{n} X_{ij} b_j.
\]

Where, n is the number of QTL, X_{ij} is the number of major alleles that individuals i carries in locus j, and b_j is the effect of j\textsuperscript{th} locus. For each individual, a phenotype was simulated via the summation of TBV with a random number extracted from a distribution with N(0, \sigma_j^2), where \sigma_j is the square root of residual variance. In each generation, up to two ancestral generations of phenotypic records were used in evaluations because using the phenotypes of more distant ancestors, especially in poultry breeding, does not have any significant impact on the accuracy of evaluations (Mehrabani-Yeganeh et al., 1999). The whole pedigree information (all information from g= 52) was used in constructing corresponding pedigree relationship matrices and calculating AF. Besides, to investigate the possibility of decreasing the generation interval, the selection was made in P₅ based on early predictions in newly born birds without using their own phenotypes throughout the generations.

Breeding values of animals were estimated using two different single-trait models. The first model, which was used as a benchmark in P₁, was a regular mixed model using the Pedigree relationship matrix (PBLUP): \(y = 1\mu+Zu+e\), where y is the vector of phenotypes, 1 is a vector of ones, \(\mu\) is the overall mean, Z is a design matrix allocating phenotypes to random effects, u is a vector of estimated breeding values with the distribution of N(0, \(A\sigma_u^2\)), where A is a numerator pedigree-based relationship matrix and \(\sigma_u^2\) is the genetic variance, e is the vector of residual effects with the distribution of N(0, \(I\sigma_e^2\)), where I is an identity matrix, and \(\sigma_e^2\) is the residual variance.

In the second model, which was applied in P₂ to P₅, the available genomic information was combined with pedigree data through the SSGBLUP. The model is similar to PBLUP, except that the distribution of random effects in u is equal to N(0, \(H\sigma_u^2\)), where H is an enhanced covariance matrix combining pedigree and genomic information, and constructed as follows (Legarra et al., 2009):

\[
H = A + A\Lambda.
\]

The \(A\Lambda\) in the equation above was calculated as:

\[
\begin{bmatrix}
A_{12} & 0 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
G_{\text{a}22} & A_{22} \\
A_{22} & A_{21}
\end{bmatrix}
\begin{bmatrix}
1 \\
0
\end{bmatrix}
\]

Where, \(G_{\text{a}22}\) was also equal to:

\[
G_{\text{a}22} = (aG_{22} + bA_{22}) + \omega.
\]

Subscripts 1 and 2 denote un-genotyped and genotyped individuals, respectively. Hence, \(G_{22}\) and \(A_{22}\) are genomic and
pedigree-based relationship matrices for genotyped birds. Having omitted SNPs with Minor Allele Frequencies (MAF) < 0.05, G was constructed using current allele frequencies following VanRaden (2008) in each generation. $a$ and $b$ weights in $G_{22}$ avoid potential singularity challenge and slightly improve predictions (VanRaden, 2008). The $\omega$ constant is the difference between the average values of $A_{22}$ and $G_{22}$ and is used to mitigate the bias of the predictions (Vitezica et al., 2011).

In each generation, top 2% males and top 20% females having the highest (G)EBVs were selected with truncation selection under the discrete generation assumption. Two mating scenarios comprising RM and MA were implemented for all subpopulations. The MA matrices were constructed with selected males in rows and selected females in columns whereby each element of matrices was equal to half of the relationship between pairs and extracted from respective relationship matrices in each subpopulation. Therefore, the best mates were chosen based on their predicted progeny inbreeding if mating between full and half-sibs was avoided. Additionally, the number of mating was similar across all birds (ten matings per male) in both RM and MA, where each male generates 100, and each female generates ten progeny. It should be noted that RM scenario in reality results in un-equal progeny as there is always mate preferences between individuals. Therefore, birds do not contribute equally to the next generation gene pool. It is known that un-equal genetic contribution eventually results in more restricted genetic variance and higher levels of $\Delta F$ in the offspring population. However, this research aimed at showing how MA is more capable in controlling the escalation of $\Delta F$ in breeding populations compared with that of the best possible RM, where all candidates contribute equally to the next generation gene pool.

Subsequently, comparisons were made based on reductions in $\Delta F$ and changes in the amounts of $\Delta G$ between two mating scenarios within each subpopulation, and two evaluation methods across all subpopulations. The $\Delta F$ levels were calculated from the changes in means of diagonal elements of pedigree-based relationship matrices while the rates of $\Delta G$ were calculated from the changes in means of TBVs. Besides, the correlation of TBVs and (G)EBVs was used as a criterion for evaluation accuracies. Eventually, the results were presented as the averages of ten replicates for all scenarios. The whole data were simulated and analyzed with self-developed R codes that can be accessed at https://github.com/VNF1981/MateAllocation_2019.

**RESULTS**

Figure 1 represents an overall design of the simulation of subpopulations. The $\Delta F$ levels at the last generation of selection and mating and changes in their values over time can be seen in Table 1 and Figure 2, correspondingly. MA outperformed RM in controlling the level of $\Delta F$ such that it resulted in 20 to 30% less $\Delta F$ in all subpopulations. Besides, it delayed the accumulation of inbreeding for at least one generation in all subpopulations compared with RM. Additionally, in $P_2$ to $P_5$, SSGBLUP realized 15, 30, 46, and 61% less $\Delta F$ in RM, and 11, 22, 44, and 55% less $\Delta F$ in MA compared to the PBLUP in $P_1$, respectively.

Furthermore, Table 2 and Figure 3 represent the average values of evaluation accuracies through the target generations on each subpopulation. In both mating scenarios, the SSGBLUP realized 6 to 29% higher accuracies in $P_2$ to $P_4$ compared to those of PBLUP in $P_1$. More precisely, the rates of improvements were within 34 to 52% for genotyped birds and 1 to 7% for un-genotyped birds. Moreover, early predictions in $P_5$ were 4% more accurate in MA, but 2% less accurate in RM compared to the PBLUP in $P_1$. The amounts of improvements for genotyped birds were 20% in RM and 27% in MA. Nevertheless, the early prediction was
almost 27% less accurate for un-genotyped birds in both mating scenarios. Also, the rates of ΔG and changes in their values over time in two mating scenarios reached by each subpopulation are presented in Table 3 and Figure 4, respectively.

**Table 1.** The average levels of ΔF at the last generation of selection and mating (g = 60) with Standard Errors (SE) in parentheses.

<table>
<thead>
<tr>
<th>Subpopulation</th>
<th>P₁</th>
<th>P₂</th>
<th>P₃</th>
<th>P₄</th>
<th>P₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inbreeding</td>
<td>RM</td>
<td>0.13 (0.02)</td>
<td>0.11 (0.02)</td>
<td>0.09 (0.01)</td>
<td>0.07 (0.02)</td>
</tr>
<tr>
<td></td>
<td>MA</td>
<td>0.09 (0.02)</td>
<td>0.08 (0.02)</td>
<td>0.07 (0.01)</td>
<td>0.05 (0.02)</td>
</tr>
</tbody>
</table>

*P₁ to P₅: Five subpopulations; RM and MA as defined under Figure 1.*

**Figure 1.** Simulation design. G: Generation; P: Five subpopulation (P₁ to P₅) with 0, 10, 20, 50, and 50% of genotyped individuals, respectively; RM: Random Mating; MA: Mate Allocation using the information of respective relationship matrices on each subpopulation; PBLUP: Pedigree-based BLUP; SSGBLUP: Single Step Genomic BLUP.

**Figure 2.** Changes in average levels of inbreeding (ΔF) in subpopulations (P₁ to P₅) through the generations 54 to 60. RM and MA as defined under Figure 1.
Table 2. The average accuracy of evaluations through the generations 54 to 60.

<table>
<thead>
<tr>
<th>Populations</th>
<th>Random Mating</th>
<th>Mate Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>Genotyped birds</td>
</tr>
<tr>
<td>P1</td>
<td>0.49 (0.04)</td>
<td>-</td>
</tr>
<tr>
<td>P2</td>
<td>0.52 (0.03)</td>
<td>0.66 (0.05)</td>
</tr>
<tr>
<td>P3</td>
<td>0.55 (0.03)</td>
<td>0.69 (0.03)</td>
</tr>
<tr>
<td>P4</td>
<td>0.62 (0.02)</td>
<td>0.73 (0.03)</td>
</tr>
<tr>
<td>P5</td>
<td>0.48 (0.05)</td>
<td>0.59 (0.06)</td>
</tr>
</tbody>
</table>

*P*1 to P5: Five subpopulations; RM and MA as defined under Figure 1. Accuracies are represented for all birds and two subsets of genotyped and un-genotyped birds calculated as the correlation between EBVs and TBVs with Standard Errors (SE) in parentheses. In each generation, all available data were used in P1 to P4 to conduct the evaluations while in P5, early predictions were applied to estimate breeding values without utilizing the phenotypic records of newly born birds.

In two mating scenarios, P2 to P4 realized 2 to 25% more ΔG than P1. The final rates of ΔG were also 3% higher through applying MA in P5. However, RM led to 2% lower ΔG in this subpopulation compared with P1. On a different perspective, the final rates of ΔG in two mating scenarios were close such that MA led to 1.5% more ΔG in P3 while RM resulted in 3 and 1% higher amounts of ΔG in P1 and P2, respectively. In P3 and P4, the final rates of ΔG were almost similar, and differences were less than 1% in both mating scenarios.

**DISCUSSION**

According to the results, SSGBLUP in P2 to P5 led to lower levels of inbreeding...
The rates of Genetic Gain ($\Delta G$) in subpopulations ($P_1$ to $P_5$) through the generations 54 to 60. RM and MA as defined under Figure 1.

**Table 3.** The average rates of Genetic Gain ($\Delta G$) at the last generation of selection and mating ($g=60$).\(^a\)

<table>
<thead>
<tr>
<th>Subpopulation</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
<th>$P_4$</th>
<th>$P_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic Gain</td>
<td>RM</td>
<td>72.55</td>
<td>73.65</td>
<td>78.81</td>
<td>88.34</td>
</tr>
<tr>
<td></td>
<td>MA</td>
<td>70.33</td>
<td>72.86</td>
<td>78.51</td>
<td>87.98</td>
</tr>
</tbody>
</table>

\(^a\) $P_1$ to $P_5$: Five subpopulations; RM and MA as defined under Figure 1. The average rates of Genetic Gain calculated as means of TBVs realized by subpopulations. In each generation, all available data was used in $P_1$ to $P_4$ to conduct the evaluations but in $P_5$, early predictions were applied to estimate breeding values without utilizing the phenotypic records of newly born birds.

compared with PBLUP in $P_1$. It has been very well known that the same relationship coefficients for close relatives in traditional covariance matrix leads to co-selection of entire families in PBLUP, and that results in higher levels of inbreeding. Furthermore, as the phenotypes cannot efficaciously mirror the genetic background of low heritable traits, there is a heavy emphasis to capture information from relatives (Daetwyler et al., 2007). Hence, even more, inbreeding would be expected to accumulate in low heritable traits in PBLUP.

In contrast, the stronger Bulmer effect by applying SSGBLUP leads to a reduction of the between-family variance and, therefore, the impact of selection becomes stronger within families. Coupled with above, exploiting part of the ancestral and recent Mendelian sampling information makes the relationships of closer relatives less correlated. Thus, birds can be more effectively ranked, the co-selection of relatives through truncation selection can be decreased and, accordingly, the levels of inbreeding can be effectively minimized. This can be seen in the results, where on average, SSGBLUP and RM led to 15 to 61\% less inbreeding in $P_2$ to $P_5$ compared with PBLUP and RM in $P_1$. The same outcome is also reported by Alemu et al. (2016), where the SSGBLUP resulted in 53\% less inbreeding than PBLUP. It is noticeable that the above process continues only for a few generations until those identical alleles due to the intense selection
restrain the regeneration of Mendelian sampling variance.

The results of the present study also support the hypothesis that the MA using the information of combined genomic-pedigree relationships can minimize the level of ΔF beyond what was described above while having a negligible impact on the final rate of ΔG. MA, in comparison with RM, made use of the available relationship information to pair the most proper candidates and, therefore, it brings about the lower level of ΔF in the progeny. Besides, MA delayed the accumulation of inbreeding for at least one generation in all subpopulations. As a result, the escalation of inbreeding was mitigated throughout the generations. From a different perspective, SSGBLUP and MA led to 38 to 69% less inbreeding in P<sub>2</sub> to P<sub>3</sub> compared with the commonly used PBLUP and RM in P<sub>1</sub>. This indicates that the information of H should not only be utilized to improve the accuracy of EBVs but also in mating designs as they can give rise to additional benefits for breeding programs by controlling the level of inbreeding. Such outcomes are valuable as the method can be easily applied using the same information that is used to predict GEBVs through a simple algorithm without logistical constraints and extra costs for the breeding program.

As expected, SSGBLUP improved the accuracy of evaluations and consequently realized higher rates of ΔG than that of PBLUP in subpopulations. Generally, improvement in accuracy due to utilizing genomic information is quite case-specific. It depends on a variety of factors, such as genotyping density, size of the training population, relationships between individuals in training and validation sets, amounts of available phenotypic records, and particulars of the traits of interest (Muir, 2007). These factors determine the potential benefits of utilizing genomic information in breeding programs as compared with PBLUP. For instance, Brinker et al. (2017) did not find any superiority for SSGBLUP on PBLUP, owing to the limited number of individuals in the training set. Nevertheless, Alemu et al. (2016), Chen et al. (2011), and Lourenco et al. (2015) reported about 33, 50, and 17% improvements on accuracies for SSGBLUP in their studies, respectively. Overall, the lower improvements on accuracies for SSGBLUP over PBLUP despite a higher percentage of genotyped individuals in this study compared with the above studies are mainly due to the higher accuracies produced by PBLUP, smaller population size, and moderate marker's density. Furthermore, since the selection intensity was equal and the selection criterion was based on the highest EBVs in truncation selection, equal rates of ΔG would be expected for MA and RM in the last generation of selection and mating. Therefore, the lower rates of ΔG in MA can be attributed to the nature of the random number generation and relatively a few numbers of replications. Moreover, results from P<sub>5</sub> showed that, under the above conditions, decreasing the generation interval is possible if the genomic information of at least 50% of birds have been considered in early predictions in order to achieve equal accuracies and the same rates of ΔG for SSGBLUP as compared with that of regular evaluations in PBLUP.

In addition, it should be emphasized that a problem with partially genotyped populations is the lack of any strategy for pre-selection of candidates for genotyping. Practically, the accuracy of evaluations can be maximized if all parents have been genotyped, which rarely happens because animals are often selected after genotyping and evaluations. We used different strategies in our simulations and found that the accuracy of evaluations can be maximized when at least one bird from each full-sib family has been randomly genotyped. This strategy results in more effective differentiation between families as it more efficiently traps part of both current and ancestral Mendelian sampling variance. On the other hand, using the genotyped individuals as parents of the progeny or genotyping animals that have the best phenotypic records would result in lower
rates of $\Delta G$ (results not shown). Overall, the best strategy would be quite case-specific as per the amounts of available information and breeding objectives.

As Preisinger (2012) and Wolc et al. (2016) have elucidated in their researches, utilizing the genomic technologies should be an appropriate choice to address the challenging issues in current chicken breeding programs (i.e., animal welfare considerations) despite the high-priced data recording. However, as long as maximization of the economic benefit is the priority of a breeding program, it is expected that the extra benefits due to implementing genomic information exceed, or at least counterbalances, the costs of genotyping.

To conclude, this research has shown that the information of $H$ relationship matrix should not only be utilized in improving the accuracy of EBVs but also in mating designs as they can precipitate additional benefits for breeding programs by controlling the level of inbreeding while having a negligible impact on final rates of genetic response. Besides, this study has also verified that SSGBLUP can result in lower levels of $\Delta F$ in addition to higher accuracies and more rates of $\Delta G$ compared to PBLUP.

REFERENCES


وضعیت همخونی و پیشرفت زنتیکی در آمیش‌های تصادفی و برنامه‌ریزی شده با استفاده از روابط خویشاوندی ترکیبی شجره‌ای-زنومی در طیور و نیکونزاد فرد و ح. مهراب‌نیا یگانه

چکیده

در این مقاله، همخونی و پیشرفت زنتیکی (MA) در آمیش‌های تصادفی با استفاده از اطلاعات ترکیبی شده‌زنومی (SSGBLUP) با نوشتاری تخصصی گرفته شده‌است. البته، این متغیرها در برنامه‌ریزی زروان‌ها و روش‌زنومی تک محلی (PBLUP) را پیشنهاد کرده‌اند. این اطلاعات در طیور و نیکونزاد فرد با استفاده از روابط خویشاوندی ترکیبی شجره‌ای-زنومی و نوشتاری تخصصی گرفته شده‌اند.

Inbreeding and Genetic Gain in Chickens

ΔG حداقل تأثیر بر مقادیر خواهد شد. همچنین، این پژوهش نشان داد که به کارگیری روش SSGBLUP می‌تواند سبب کنترل موثرتر در مقایسه با PBLUP در جمعیت‌های اصلاحی ΔF گردد.