Long-Term Impact of Conventional and Optimal Contribution Conservation Methods on Genetic Diversity and Genetic Gain in Chinese Indigenous Pig Breeds

Qingbo Zhao
Shanghai Jiao Tong University
Huiming Liu
Aarhus Universitet
Qamar Raza Qadri
Shanghai Jiao Tong University
Qishan Wang
Zhejiang University
Yuchun Pan (panyuchun1963@aliyun.com)
Shanghai Jiao Tong University https://orcid.org/0000-0002-1163-5963
Guosheng Su
Aarhus University: Aarhus Universitet

Research Article

Keywords: conventional conservation, indigenous pig resources, genetic diversity, genetic gain, inbreeding, optimal contribution

Posted Date: May 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-526041/v1

Long-Term Impact of Conventional and Optimal Contribution Conservation Methods on Genetic Diversity and Genetic Gain in Chinese Indigenous Pig Breeds

Qingbo Zhao[12#], Huiming Liu[2#], Qamar Raza Qadri[1], Qishan Wang[3], Yuchun Pan[3]*, Guosheng Su[2]*

[1]School of Agriculture and Biology, Department of Animal Science, Shanghai Jiao Tong University, Shanghai 200240, PR China
[2]Center for Quantitative Genetics and Genomics, Faculty of Science and Technology, Aarhus University, Tjele 8830, Denmark
[3]Department of Animal Breeding and Reproduction, College of Animal Science, Zhejiang University, Hangzhou 310030, PR China

Qingbo Zhao, E-mail: zhaoqingbo@sjtu.edu.cn
Huiming Liu, E-mail: huiming.liu@qgg.au.dk
Qamar Raza Qadri, E-mail: qamrq25@sjtu.edu.cn
Qishan Wang, E-mail: wangqishan@zju.edu.cn

*Corresponding author:
Yu-chun Pan, E-mail: panyuchun1963@aliyun.com
Guo-sheng Su, E-mail: guosheng.su@qgg.au.dk

#: These authors have contributed equally to this work

Abstract

Background

China has rich and vast genetic resources of indigenous pig breeds. Currently, great attention is paid to either crossbreeding or conservation of these indigenous pig breeds, and insufficient attention is paid to the combination of conservation and breeding along with their long-term effects on genetic diversity. The genetic diversity of livestock is essential to increase productivity and respond to future challenges such as climate change. The genetic stability and product consistency of these indigenous pig breeds should be focused on and further improved. Therefore, the objective of this study is to compare the long-term effects of using conventional conservation and optimal contribution selection methods on genetic gain and genetic diversity.

Results

A total of 11 different methods including conventional conservation and optimal contribution selection methods were investigated using stochastic simulations with a population size of 600 animals in each generation. Each scenario was run for 20 generations and 100 replicates. The long-term effects of using these methods were evaluated in terms of rate of genetic gain, rate of true inbreeding based on genome-wide identity-by-descient (IBD) markers and various genetic diversity metrices such as expected heterozygosity (He). The results indicated that the rates of true inbreeding in these conventional conservation methods were maintained at around 0.01. The optimal contribution selection methods based either on the pedigree (POCS) or genome (GOCS) information showed more genetic gain than conventional methods, and POCS achieved the largest gentic gain. Furthermore, the effect of using GOCS methods on most of the genetic diversity metrics was slightly better than the conventional conservation methods when the the rate of true inbreeding was the same, but this also required more sires used in OCS methods. According to the rate of true inbreeding, there was no significant difference among these conventional methods.

Conclusion

In conclusion, there is no significant difference in different ways of selecting sows on inbreeding when we use different conventional conservation methods. Compared with conventional methods, POCS method could achieve the most genetic gain. However, GOCS methods can not only achieve higher genetic gain, but also maintain a relatively high level of genetic diversity. Therefore, GOCS is a better choice if we want to combine conservation and breeding in actual production in the Chinese national-level conservation farms.

Keywords conventional conservation; indigenous pig resources; genetic diversity; genetic gain; inbreeding; optimal contribution;

Background

The genetic diversity of livestock is essential to improve productivity and respond to challenges including food security and climate change mitigation in the future [1]. However, due to agricultural innovation since the beginning of the 19th century and subsequent intensification of production, many local varieties can not adapt to the resulting changes. Pigs are one of the most common domestic animals, and more than one-third of indigenous pig breeds in the world are in China. These indigenous pig breeds generally have high fertility, good meat quality and high tolerance to harsh environmental conditions [2]. However, pig inductries currently pay more attention to crossbreeding for indigenous pig breeds, and insufficient attention to the combination of conservation and genetic improvement. To improve the production and economic value of breeding stocks, indigenous pig breeds usually cross with the foreign breeds which have high production performance. Thus, gene flow usually only occurs from breeds with superior economic characteristics to indigenous pig breeds [3]. Furthermore, due to the epidemic of diseases such as African swine fever, a lot of precious indigenous breeds are on the verge of extinction. The herds of indigenous pig breeds have reduced greatly during this period. To protect these indigenous pig breeds in China, the Chinese government has established national-level breeding farms for most indigenous pig breeds, and a large number of breeding funds are used to protect these unique indigenous pig breeds every year [4].

In these national-level conservation farms, our goal is to maintain each breed's genetic materials and control the rate of inbreeding as much as possible. Inbreeding is an important reason for the loss of genetic variation, and the rate of inbreeding mainly depends on the effective population size [5]. In order to reduce the impact of inbreeding on the loss of population genetic variation, the increase in the inbreeding coefficient of each generation in the conservation population is recommended to be controlled at 1-4%, and the conservation population needs to have at least 12-25 sires and 100-250 dams [6]. These guides are commonly used in conservation farms. Simultaneously, to keep the genetic diversity and control rate of inbreeding, conservation farms attempts to keep the same number of offspring for each family. However, current conservation methods do not combine conservation of genetic resources and genetic improvement of production performance. In conservation populations, it is also very important to properly select the dominant traits of each local pig breed, which will help to further consolidate the advantages of the breed and maintain the uniqueness of each breed. From a long-term sustainable perspective, how to combine conservation and selective breeding in the conservation field is a crucial issue. Great attention should be paid to genetic improvement of important economic traits while maintaining the overall genetic diversity of locsal breeds to meet pig industry's sustainable development and even other livestock industries. Therefore, we need to re-examine our current conservation strategy. The new conservation strategy should take into account at least two principles at the

same time. One of the principles is to keep genetic diversity as high as possible, and the other is to obtain genetic progress of some essential economic traits as large as possible. Optimal contribution selection (OCS) is an effective selection method that balances inbreeding and genetic gain [7, 8]. It maximises rates of predicted genetic gain while controlling inbreeding at given rates by optimizing the genetic contribution of each selection candidate to the next generation [9-11]. Optimal contribution selection can be based either on pedigree information (POCS) or genome information (GOCS). Several previous studies have investigated the impact of OCS on rate of inbreeding and long-term genetic gain based on simulated data [12, 13] and real data [9, 14]. Gourdine et al. [13] claimed that the genetic gain with optimal contribution selection could be similar to truncation selection, but the inbreeding was lower. Sánchez-Molano et al. [15] showed that genome-based optimal contribution strategies could effectively control inbreeding even when selected traits, adaptive traits and production traits, are negatively correlated using simulated data. Henryon et al. [16] reported that the optimal contribution selection based on pedigree information for controlling inbreeding could achieve more genetic gain than that based on genome information due to less restriction on the change of QTL allele frequencies. Nowadays, there are many ways to measure changes in genetic variation and its diversity. We can calculate the inbreeding coefficient if the pedigree information is known. However, in actual situations, the registration of pedigree is often incomplete and inaccurate, limiting the usage of this method. With the use of molecular markers, more and more genetic diversity indicators are used to assess the degree of diversity of a particular population [17-20].

The objective of this study is to compare the long-term effects on genetic diversity and genetic gain of using conventional conservation and OCS methods. We acheive this by using a stochastic simulation study approach, where 11 different conservation methods for a small pig population were compared. The results of this simulation study are expected to provide guidance to breeders and government departments on formulating better conservation programs.

Methods

Experimental design

We used stochastic simulation to estimate the long-term genetic gain and genetic diversity using different conservation methods. The conservation methods included conventional conservation methods and OCS methods. The sires selected were the males with the highest estimated breeding value (EBV) within each sire (half-sib) family for conventional conservation methods. The dams were selected using one of the six methods:

1. Selecting females with the highest EBV within each full-sib family (Sirehalf - Damfull scenario)

2. Selecting females with the highest EBV within each half-sib family (Sirehalf-Damhalf scenario)
3. Selecting females with the highest EBV without considering the families (Sirehalf-Damtrunc scenario)
4. Randomly sampling females from each full-sib family (Sirehalf-DamfullRandom scenario)
5. Randomly sampling females from each half-sib family (Sirehalf-DamhalfRandom scenario)
6. Randomly sampling females without considering the families (Sirehalf-DamRandom scenario)

For OCS methods, either pedigree (POCS) or two genomic (GOCS) relationship matrices were used to constrain the rate of inbreeding. We also simulated truncation selection and random selection as reference methods. In total, there were 11 selection methods studied. Each selection method was run for 20 discrete generations, and the animals were selected based on a single trait controlled by 360 quantitative trait loci (QTL). The heritability of the trait was set to 0.2. Furthermore, 36,000 markers were simulated to carry out GOCS. For the methods other than OCS, 12 sires were selected, and each sire was mated to 10 dams in each generation. For OCS methods, the males were allocated 0, 1, 2 … or 120 matings by the program, and females were allocated a single mating in each generation. Each dam produced five offspring with an equal sex ratio. The animals were phenotyped before selection.

OCS methods

POCS allocated matings of selection candidates in generations $t = 1 \ldots 20$ according to EBV and pedigree relationships between all the involved animals. It was done by maximizing, U_t, with respect to $\boldsymbol{c}$ [21] **:**

$U_t(\boldsymbol{c}) = \boldsymbol{c}'\widehat{\boldsymbol{\alpha}} + \omega \boldsymbol{c}'\boldsymbol{A}\boldsymbol{c}$**,**

Where $\boldsymbol{c}$ is an n dimentional vector of genetic contributions, where n is the number of selected candidates, $\widehat{\boldsymbol{\alpha}}$ is n vector of EBV, and $\boldsymbol{A}$ is a $n \times n$ matrix for selected candidates which is a submatrix from the full additive genetic relationship matrix for all animals in the pegigree. In this study, pedigree of the selected candidates was traced back to the base population [22]. Elements of $\boldsymbol{c}$ were constrained to $0 \leq c_i \leq 0.5$ $(i = 1 \ldots n)$ and the sum of contributions were 0.5 for each sex.The component $\boldsymbol{c}'\widehat{\boldsymbol{\alpha}}$ is the expected breeding value, and the component $\boldsymbol{c}'\boldsymbol{A}\boldsymbol{c}$ is the expected average relationship of the proposed offspring. The penalty, ω, is applied to the expected average relationship of the next generation, which was constant across generations. GOCS was performed by replacing $\boldsymbol{A}$ with a $n \times n$ genomic relationship matrix ($\boldsymbol{G}$) which was calculated with genotypes for all markers of all the selected candidates using the method described by Yang et al.[23]. A range of ω (1, 5, 10 …100) was applied to examine the pattern of genetic gain with different inbreeding rates.

We used also an additional method to build a $\boldsymbol{G}$ matrix for GOCS, where the markers with a distance from a random QTL less than 1cM were excluded from the $\boldsymbol{G}$ (we call the corresponding GOCS the GOCS-1cM). To differentiate this GOCS method from the conventional GOCS method, we called the classical GOCS method the GOCS-0cM. We used EVA [24] to perform POCS and GOCS.

Simulations

Simulations of each conservation method were carried out in three stages: 1) a single founder population was generated as the basis for the subsequent stages, 2) a unique base population was sampled from the last generation of founder population, and 3) a selected population was generated based on the base population. Stage 2 and 3 were run for 100 times to produce 100 replicates. To simplify the simulation, instead of direct calculation of EBV, the EBV was approximated by the breeding values of a genetically correlated pseudo-trait [25]. The genetic correlation was set to 0.6, mimicking a genomic selection with an accuracy of 0.6 [26].

Founder population and genetic architecture: Generations -2000 to -1

The linkage disequilibrium of QTLs and the markers was generated by simulating a founder population with QMSim [27] using a Fisher-Wright inheritance model. The population had an effective population size of 200 animals (100 males and 100 females) and 2,000 discrete generations. The simulated genome consisted of eighteen 1 Morgan long chromosomes, on which 10,000 loci were equidistantly distributed, resulting in 180,000 loci in total across the genome. The recurrent mutation was allowed at a rate of 2.5×10^{-5} and recombination per chromosome was sampled from a Poisson distribution with a mean of 1.
At the last generation of the founder population (generation -1), among all segregating loci, every second locus with a minor allele frequency (MAF)>0.05 were used as potential markers. In total, we selected 36,000 markers from these potential marker loci. In total, 360 QTLs were selected from the remaining segregating loci with MAF > 0.01. The QTL allelic effects were assumed to follow a gamma distribution with a shape parameter of 1.48, which was derived from distributing QTL effects in pig breeds [28].

Base population: Generation 0

In generation 0, 200 animals were generated by random mating of 100 males and 100 females in generation -1. From these 200 animals, 12 males and 120 females were randomly selected as base animals to produce 600 offspring with an equal sex ratio.

Selected population: Generations 1 to 20

In each of generations 1 to 20, 120 matings were allocated to sires and dams, and each dam was allocated a single mating to produce five offspring with an equal sex ratio. The offspring in each generation inherited alleles of markers and QTLs from their parents, following Mendel's laws of heredity allowing for recombinations following a Poisson diostribution with a mean of 1. The phenotype of the trait for the i^{th} animal, y_i, was calculated as $y_i = \alpha_i + e_i$, where α_i is the animal's true additive genetic value and e_i is the residual environmental value. The true additive genetic value was calculated as the sum of all QTL effects. Those QTL effects were scaled in generation 0 to achieve an initial additive genetic variance equal to the heritability of 0.2. The additive QTL variance explained all additive genetic variance ($\sigma_\alpha^2 = \sigma_{qtl}^2$). Thus the true breeding value (TBV) for an individual was equal to the the sum of QTL effects of the individual. The environmental values were sampled from the distribution $N(0, \sigma_e^2 = 1 - h^2)$. The environmental variance σ_e^2 was constant through the generations of the simulation, such that genetic variance and heritability decreased throughout the generations of selection due to random drift, fixation, and Bulmer-effect [29].

Tracing identity-by-descent

To compute the rate of true inbreeding, 2,000 identical-by-descent (IBD) loci were equidistantly placed on each chromosome of animals in the base populations. Unique alleles were assigned to these IBD loci in the base population to trace each base animal's contribution to their descendants [30]. A descendant was IBD at an IBD locus when it inherited two copies of a unique allele. These IBD loci were not used for prediction or selection.

Statistical analyses

Rate of genetic gain and inbreeding

The rates of genetic gain and the rates of true inbreeding are presented as means ($\pm SD$) of the 100 replicates. The rate of genetic gain in each replicate was calculated as a linear regression of G_t on t, where G_t is the average TBV of animals born in generation, $t = 1 \ldots 20$ in each replicate. The rate of inbreeding was calculated as $1 - exp(\beta)$, where β is a linear regression coefficient of $\ln(1 - F_t)$ on t, F_t was the average true inbreeding coefficient of all the individuals born in generation t ($t = 1 \ldots 20$), and the inbreeding coefficient of each individual was calculated as the proportion of IBD loci being IBD to total IBD loci in the genome [30].

Genetic diversity metrics

We calculated the following genetic diversity metrics.

Expected heterozygosity (H_e) is the probability that an individual will be heterozygous at a given locus in one population. It is calculated by Nei's [31] method as follows:

189 $$H_e = \frac{2n}{2n-1}\frac{1}{N}\sum_{k=1}^{N}\left(1-\sum P_{k_i}^2\right)$$

190 Observed heterozygosity (H_o) refers to the ratio of the observed heterozygous
191 individuals in the population to the total number of individuals. The calculation formula
192 is as follows:

193 $$H_o = \frac{1}{N}\sum_{K=1}^{N}\frac{H_k}{n}$$

194 where n is the number of individuals in the population, N is the total number of loci,
195 H_k is the number of individuals with locus K is heterozygous and P_{k_i} is the
196 probability of allele i at locus K.
197 The number of the polymorphic gene loci (M01 and M05) is defined as the minor allele
198 frequency of a gene locus is larger than or equal to 0.01 or 0.05 [32].
199 Effective allele number (A_e) in one population is calculated by $A_e = 1/{p_i}^2$, where p_i
200 is the frequency of the *i*-th allele of the gene locus [33].

201 # Results

202 ## The rate of genetic gain and rate of true inbreeding

203 We presented long-term response frontiers by plotting the rate of genetic gain
204 against the rate of inbreeding with all possible solutions by applying different penalties
205 for POCS and GOCS. As shown in Fig.1, the rate of true inbreeding for all the
206 conventional methods was around 0.01, except for the truncation selection scenario,
207 consistent with basic conservation theory. Most importantly, inbreeding increment of
208 each generation in the random scenario was also around 0.01, which indicates that the
209 rate of true inbreeding could also be controlled around 0.01 as long as we maintain an
210 approperiate population size (such as12 males and 120 females in this study) and
211 guarantee complete random mating. Among these conventional conservation methods,
212 when the inbreeding increment was 0.01, the scenarios of Sirehalf-Damtrunc, Sirehalf-
213 Damhalf, and Sirehalf-Damfull obtained higher genetic gains. However, the genetic
214 gain was much smaller when dam was randomly selected in three various form (i.e.,
215 Sirehalf-DamfullRandom, Sirehalf-DamhalfRandom and Sirehalf-DamRandom). As
216 expected, no genetic gain was obtained when both sire and dam were randomly selected.
217 Compared with the six convervational methods, truncation selection on both sire and
218 dam increased genetic gain by 7.5% (vs. Sirehalf-Damtrunc) to 67.5% (vs. Sirehalf-
219 DamfullRandam), but it tripled the rate of inbreeding.
220 All the OCS methods based on the genome and pedigree information realized more
221 genetic gain than the conventional conservation methods when the inbreeding rate was
222 almost around 0.01 (Fig. 1 and Table 1, *p*=10). Interestinly, there was no significant
223 difference in both the rate of true inbreeding and genetic gain between these two GOCS

methods (GOCS-0cM and GOCS-1cM). However, POCS could achieve more genetic gain than GOCS when the inbreeding rate was the same. POCS with penalty p=10 obtained the rate of genetic gain as high as the Truncation scenario, but the rate of inbreeding was only one third. The rate of inbreeding of two GOCS methods and POCS were similar to that of the Sirehalf-Damtrunc scenario when the penalty p =7 ,which used the same number of sires (Fig. 6 and Additional file 3), but they would be similar to that of the four methods (Sirehalf-Damfull, Sirehalf-DamfullRandom, Sirehalf-Damhalf and Sirehalf-DamhalfRandom) when the penalty p=10.

Changes in genetic diversity of different methods

In terms of He and Ho, as shown in Fig. 2a and 2b, 3a and 3b, GOCS methods were better than the conventional conservation method when the penalty p was increased to 10. As the number of generations increases, the declined slope of the He and Ho was smaller in GOCS methods compared to the conventional conservation methods, which indicated that GOCS had better effect than that of the conventional conservation methods. However, POCS was not superior to the conventional conservation methods and was only better than the truncation scenario, in terms of He and Ho. Furthermore, there was no significant difference in He and Ho among the conventional conservation methods.

Fig. 2c, 3c and Table 2 showed that there were more effective alleles in two GOCS scenarios when the weight p was 10. Regardless of p=7 or p=10, the POCS led to low Ae (see Fig. 3c, Table 2 and Additional file 2), only higher than the truncation scenario. For M01 and M05 (Fig. 2d, 3d and Additional file 1), several conventional conservation methods, such as Sirehalf-Damfull, Sirehalf-DamfullRandom, Sirehalf-Damhalf, and Sirehalf-DamhalfRandom, were better than those of OCS methods.

Changes of additive genetic variances across generations are presented in Fig. 4. From Fig. 4, the additive genetic variance in optimal contribution selection methods had the fastest decline, compared with other scenarios except for the method of truncation selection on both sire and dam. The variance in the POCS method was lower than that in GOCS methods. The four conventional scenarios with different types of random selection had the highest additive genetic variance, and the order was Random, Sirehalf-DamfullRandom, Sirehalf-DamhalfRandom, and Sirehalf-DamRandom. The trends of additive variance and inbreeding were generally inversely consistent.

Number of ancestors and sires used

As for the number of ancestors for different methods (Fig. 5), the pattern of the number of ancestors in the Sirehalf-Damfull and Sirehalf-DamfullRandom methods were different from the other methods. These two methods remained the same number of ancestors in the first generations. The trends began to decline after the fifth generation, indicating that some ancestors failed to make contributions as selection

proceeds due to selection and genetic drift. For other methods, the number of ancestors declined rapidly in the previous generations, and then gruadually fallened out. Therefore, keeping the same number of offspring from each sire and dam family will have the best effect in the first few generations. In addition, Sirehalf-Damfull and Sirehalf-DamfullRandom methods also retained the largest number of ancestors in the last few generations. The second largest number of ancestors was observed in the OCS scenarios, including GOCS-0cM, GOCS-1cM, and POCS.

The number of sires was around 12,which is the same as in conservational methods, when the weight *p* was 7 in two GOCS scenarios (Fig. 6 and Additional file 3). However, when the same number of sires were included, the weight *p* could be about 6 in the POCS scenario. When the weight *p* was 7 in the POCS scenario, the number of sires was around 13. This may indicate POCS method needs to use a little more sires to maintain the same level of genetic diversity or the rate of inbreeding compared to the GOCS method, although POCS could achieve more genetic gain. All OCS methods used more sires when the rate of true inbreeding was the same as in conventional methods. The number of sires used was about 18 in POCS and about 16 in two GOCS methods.

Discussion

There are huge indigenous pig breed resources in China, and these indigenous pig breeds have formed relatively unique characteristics under long-term environmental and artificial pressure. Nowadays, the focus of conventional conservation methods is only to control rate of inbreeding, and not much attention is paid to the selection of favorable traits for each breed in the existing conservation field. It will be helpful to further improve the advantage and uniqueness of each breed if we can combine the maintenance of genetic diversity and the selection of favorable traits. In this study, we studied conventional conservation and optimal contribution scenarios to conserve indigenous pig breeds with small population sizes using simulation studies. We explored the genetic diversity changes and genetic gain of these conservation scenarios during 20 generations. The founding is helpful in guiding the current conservation programs.

To utilise indigenous pig breeds for pig production, genetic improvements for some important economic traits in conservative pig populations is necessary. In the current study the genetic gain obtained by the optimal contribution selection methods show a trend of increasing with increasing weight *p* when *p* was small, and then decreases as the weight *p* increases. This may be because the increase in selection intensity with small *p* accelerates the reduction of genetic variation within the population, thereby reducing the further improvement of genetic gain. This implies that selection without a restraint on inbreeding will lead to the selection limit [34]. Long-term high-intensity selection will reduce the population's genetic variation, and the reduction of genetic variation will counteract the increase in genetic gain. In addition, from the changes in

genetic diversity of different scenarios, we could see that the penalty on genetic relationship should at least 10 if we want to apply GOCS methods for maintaining a higher heterozygosity than the conventional conservation scenarios.

As expected, the truncation method caused largest rate of inbreeding (Fig. 1, Table 1), and the trend became very significant as the number of generations increased. It indicates that we should not use this method to conserve indigenous pig breeds when the population size is small among conservation farms, which is different from selecting and breeding in the breeds for commercial production. Using conventional conservation methods, the trends of rate of inbreeding were amost the same in the senarios with different methods of selection on dam, except for trucation selection. This indicates that selection of dam within full-sib or half-sib family or random selection of dam from the whole population could all be used in actual conservation operation. The most important thing is that the boars should come from each sire family. The relaxion on restriction on dam could significantly reduce the farmers' workload and benefit for genetic improvement. Therefore, this result could guide the actual conservation operation.

Different indicators have been used to measure genetic diversity. Each indicator has its advantages and disadvantages [35]. Ayala et al. [36] summarized the study on genetic diversity of main domestic animals and indicated that the ratio of polymorphic loci and the average expected heterozygosity were the primary parameters to measure genetic diversity. Qian et al. [37] reported that the degree of expected heterozygosity was more effivtive than the ratio of polymorphic loci in accuracy of measuring genetic diversity. The variation of the number of polymorphic loci is relatively small, and the sensitivity to genetic diversity is relatively low [37]. In addition, the number of effective alleles could more effectively measure the change of genetic diversity in one population [38]. Therefore, in this study, we used multiple indicators to measure the impact of different conservation methods on genetic diversity changes after a number of generations to make our results more comprehensive and objective to a certain extent. These indicators could complement each other. For each indicator of genetic diversity, the results of GOCS-0cM and GOCS-1cM were similar. As the results shown in Fig. 4, the additive genetic variance in OCS methods had the fastest decline, which indicates the OCS methods results in larger increases in the frequencies of favorable alleles at QTL, compared with the other methods. Moreover, POCS is larger than GOCS, which is also consistent with the previous study [39]. The results of the number of polymorphic loci such as M01 and M05 (Fig. 3d, Additional file 1) also illustrate this point. When the penalty is 7, the optimal contribution selection methods (POCS and GOCS) were not better than several conventional conservation methods such as Sirehalf-Damfull, Sirehalf-DamfullRandom, Sirehalf-Damhalf, and Sirehalf-DamhalfRandom. This may be also due to the selection of QTLs that affect the traits, and the directional selection decreases polymorphic loci ratio.

From OCS methods, we can see that the genetic gain obtained by the optimal contribution selection method based on the pedigree relationship is higher than that

obtained based on genomic relationship when the rate of inbreeding in each generation is controlled at about 0.01. OCS automatically determines the number of male animals required to control inbreeding when we use OCS methods. Through comparison (Fig. 6 and Additional file 3), it is found that POCS method requires more sires than GOCS when the rate of inbreeding is controlled in the same level. In actual pig production, POCS method is often easier to put into practice. POCS method is based on pedigree information, and it only requires that the pedigree of each animal is registed in the conservation farms. Unlike POCS, GOCS method is based on genomic information, which requires individuals' genotype data. Thus the cost is relatively high for conservation farms. It is impossible to genotype all individuals to obtain genotype data for general conservation farms. Therefore, if keeing a little more males is accepatable by conservation farms, POCS method is a better way if we want to obtain more genetic gain. However, if we want to achieve the balance of conservation and selection, GOCS is a better choice, which allows to both control inbreeding and improve economic traits, compared to the other conservation methods.

Many factors could influence the objectives of a conservation or breeding program, such as economical value, historic bottlenecks, and the maintenance of genetic diversity level [40]. Nowadays, we pay more attention to increase economic merit for most livestock breeds,. Thus the most critical breeding objective is to maximize genetic gain. However, for example, historic bottlenecks are commonly suffered in companion animals because of an overuse of elite males. Therefore, the priority is to minimize inbreeding in these animal populations. In addition, the focus would be changed to increase conservation values in endangered breeds that get allowance for better conservation. This could be realized by increasing their genetic distance or recovering the native genetic background among these breeds. These goals conflict with each other to a certain extent. In order to maximize genetic gain, people would prefer to choose the animals with the highest breeding values for economic traits, which will increase rate of inbreeding, and may lead to inbreeding depression and new bottlenecks. Generally, commercial breeds often have the highest breeding values for economic traits, which would further lead to the loss of the genetic diversity of native breeds.

It is important to protect and conserve the indigenious pig breeds, especially when their population size has dropped sharply. However, it is not sensilbe and conducive if we only focus on protecting but not improving favorable traits. The current study shows that the optimal contribution selection method based on genomic information can maintain a high genetic diversity while improving the traits we want to improve, which is in line with our current needs.

Conclusion

In conclusion, our study showed conventional conservation scenarios resulted in the rate of inbreeding for each generation was at around 0.01. Different methods to select sow has small impact on inbreeding when we use conventional conservation methods.

Compared with conventional methods, POCS method could achieve the most genetic gain. However, two GOCS methods (GOCS-0cM and GOCS-1cM) can not only achieve higher genetic gain, but also maintain a relatively high level of genetic diversity, and the results of these two GOCS methods are similar. In particular, the advantages of GOCS that enable genetic diversity to be maintained at a higher level becomes more and more obvious as the number of generations increases. Therefore, GOCS is a better choice if we want to combine conservation and breeding in actual production in the Chinese national-level conservation farms. We can also choose whether to obtain higher genetic gain or maintain a higher level of genetic diversity according to our needs, and then appropriately adjust the conservation strategy according to our different concerns and goals.

List of abbreviations

OCS: Optimal contribution selection

POCS: Optimal contribution selection based on pedigree information

GOCS: Optimal contribution selection based on genome information

EBV: Estimated breeding value

QTL: Quantitative trait locus

IBD: Identical by descent

TBV: True breeding value

F: Inbreeding coefficient

He: Expected heterozygosity

Ho: Overserved heterozygosity

Ae: Population effective alleles

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (grant nos. 31772552, 31941007, 31872321).

Authors' Contributions

YP, GS, QZ, and HL conceived the study. YP and GS supervised the study, while QZ and HL ran the simulation, analyzed the data, and wrote the main manuscript. QZ, GS, and HL interpreted the results. All authors gave necessary suggestions, revised and approved the final manuscript.

Acknowledgments

The authors are grateful to the National Natural Science Foundation of China. The first author acknowledges the scholarship provided by the China Scholarship Council (CSC).

References

1. Groeneveld LF, Lenstra JA, Eding H, Toro MA, Scherf B, Pilling D, Negrini R, Finlay EK, Jianlin H, Groeneveld E *et al*: **Genetic diversity in farm animals--a review**. *Anim Genet* 2010, **41 Suppl 1**:6-31.
2. Zhao QB, Oyelami FO, Qadri QR, Sun H, Xu Z, Wang QS, Pan YC: **Identifying the unique characteristics of the Chinese indigenous pig breeds in the Yangtze River Delta region for precise conservation**. *BMC Genomics* 2021, **22**(1):151.
3. Wellmann R, Hartwig S, Bennewitz J: **Optimum contribution selection for conserved populations with historic migration**. *Genet Sel Evol* 2012, **44**:34.
4. Zhao QB, Lopez-Cortegano E, Oyelami FO, Zhang Z, Ma PP, Wang QS, Pan YC: **Conservation Priorities Analysis of Chinese Indigenous Pig Breeds in the Taihu Lake Region**. *Frontiers in Genetics* 2021, **12**.
5. Pekkala N, Knott KE, Kotiaho JS, Nissinen K, Puurtinen M: **The effect of inbreeding rate on fitness, inbreeding depression and heterosis over a range of inbreeding coefficients**. *Evol Appl* 2014, **7**(9):1107-1119.
6. Smith C: **Genetic-Aspects of Conservation in Farm Livestock**. *Livest Prod Sci* 1984, **11**(1):37-48.
7. Dagnachew BS, Meuwissen TH: **A fast Newton-Raphson based iterative algorithm for large scale optimal contribution selection**. *Genet Sel Evol* 2016, **48**(1):70.
8. Gomez-Romano F, Villanueva B, Fernandez J, Woolliams JA, Pong-Wong R: **The use of genomic coancestry matrices in the optimisation of contributions to maintain genetic diversity at specific regions of the genome**. *Genetics Selection Evolution* 2016, **48**.
9. Hayes BJ, Lewin HA, Goddard ME: **The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation**. *Trends Genet* 2013, **29**(4):206-214.
10. Wang Y, Bennewitz J, Wellmann R: **Novel optimum contribution selection methods accounting for conflicting objectives in breeding programs for livestock breeds with historical migration**. *Genet Sel Evol* 2017, **49**(1):45.
11. Grundy B, Villanueva B, Woolliams JA: **Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development**. *Genet Res* 1998, **72**(2):159-168.
12. Woolliams JA, Berg P, Dagnachew BS, Meuwissen THE: **Genetic contributions and their optimization**. *J Anim Breed Genet* 2015, **132**(2):89-99.
13. Gourdine JL, Sorensen AC, Rydhmer L: **There is room for selection in a small local pig breed when using optimum contribution selection: A simulation**

study. *J Anim Sci* 2012, **90**(1):76-84.
14. Sanchez-Molano E, Pong-Wong R, Banos G: **Genomic-Based Optimum Contribution in Conservation and Genetic Improvement Programs with Antagonistic Fitness and Productivity Traits**. *Frontiers in Genetics* 2016, **7**.
15. Sanchez-Molano E, Pong-Wong R, Banos G: **Genomic-Based Optimum Contribution in Conservation and Genetic Improvement Programs with Antagonistic Fitness and Productivity Traits**. *Front Genet* 2016, **7**:25.
16. Henryon M, Liu H, Berg P, Su G, Nielsen HM, Gebregiwergis GT, Sorensen AC: **Pedigree relationships to control inbreeding in optimum-contribution selection realise more genetic gain than genomic relationships**. *Genet Sel Evol* 2019, **51**(1):39.
17. Kaljund K, Jaaska V: **No loss of genetic diversity in small and isolated populations of Medicago sativa subsp falcata**. *Biochem Syst Ecol* 2010, **38**(4):510-520.
18. Ebrahimi A, Fatahi R, Zamani Z: **Analysis of genetic diversity among some Persian walnut genotypes (Juglans regia L.) using morphological traits and SSRs markers**. *Sci Hortic-Amsterdam* 2011, **130**(1):146-151.
19. Chen J, Peng J, Xiao Q, Pan Y, Zhang X, Lo LJ, Xu N: **The genetic diversity and population structures of indigenous pig breeds in Zhejiang Province revealed by GGRS sequencing**. *Anim Genet* 2018, **49**(1):36-42.
20. Zhao QB, Sun H, Zhang Z, Xu Z, Olasege BS, Ma PP, Zhang XZ, Wang QS, Pan YC: **Exploring the Structure of Haplotype Blocks and Genetic Diversity in Chinese Indigenous Pig Populations for Conservation Purpose**. *Evol Bioinform Online* 2019, **15**:1176934318825082.
21. Meuwissen TH: **Maximizing the response of selection with a predefined rate of inbreeding**. *J Anim Sci* 1997, **75**(4):934-940.
22. Henderson CR: **Simple Method for Computing Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values**. *Biometrics* 1976, **32**(1):69-83.
23. Yang JA, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW *et al*: **Common SNPs explain a large proportion of the heritability for human height**. *Nat Genet* 2010, **42**(7):565-U131.
24. Berg P, Nielsen J, Sørensen MK: **EVA: Realized and predicted optimal genetic contributions**. In: *Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006: 2006*: Instituto Prociência; 2006: 27-09.
25. Dekkers JCM: **Prediction of response to marker-assisted and genomic selection using selection index theory**. *J Anim Breed Genet* 2007, **124**(6):331-341.
26. Thomasen JR, Willam A, Egger-Danner C, Sorensen AC: **Reproductive technologies combine well with genomic selection in dairy breeding**

programs. *J Dairy Sci* 2016, **99**(2):1331-1340.
27. Sargolzaei M, Schenkel FS: **QMSim: a large-scale genome simulator for livestock**. *Bioinformatics* 2009, **25**(5):680-681.
28. Hayes B, Goddard ME: **The distribution of the effects of genes affecting quantitative traits in livestock**. *Genetics Selection Evolution* 2001, **33**(3):209-229.
29. Bulmer MG: **Effect of Selection on Genetic Variability**. *Am Nat* 1971, **105**(943):201-+.
30. Liu H, Henryon M, Sorensen AC: **Mating strategies with genomic information reduce rates of inbreeding in animal breeding schemes without compromising genetic gain**. *Animal* 2017, **11**(4):547-555.
31. Nei M: **Estimation of Average Heterozygosity and Genetic Distance from a Small Number of Individuals**. *Genetics* 1978, **89**(3):583-590.
32. Ayala FJ, Powell JR, Tracey ML, Mourao CA, Perez-Salas S: **Enzyme variability in the Drosophila willistoni group. IV. Genic variation in natural populations of Drosophila willistoni**. *Genetics* 1972, **70**(1):113-139.
33. Nei M: **Analysis of gene diversity in subdivided populations**. *Proc Natl Acad Sci U S A* 1973, **70**(12):3321-3323.
34. Zeng ZB, Hill WG: **The Selection Limit Due to the Conflict between Truncation and Stabilizing Selection with Mutation**. *Genetics* 1986, **114**(4):1313-1328.
35. Allendorf FW: **Genetic Drift and the Loss of Alleles Versus Heterozygosity**. *Zoo Biol* 1986, **5**(2):181-190.
36. Ayala FJ, Kiger JA: **Modern genetics**, 2nd edn. Menlo Park, Calif.: Benjamin/Cummings Pub. Co.; 1984.
37. Qian W, Ge S, Hong DY: **Genetic variation within and among populations of a wild rice Oryza granulata from China detected by RAPD and ISSR markers**. *Theor Appl Genet* 2001, **102**(2-3):440-449.
38. Nei M: **Analysis of Gene Diversity in Subdivided Populations**. *P Natl Acad Sci USA* 1973, **70**(12):3321-3323.
39. Henryon M, Liu HM, Berg P, Su GS, Nielsen HM, Gebregiwergis GT, Sorensen AC: **Pedigree relationships to control inbreeding in optimum-contribution selection realise more genetic gain than genomic relationships**. *Genetics Selection Evolution* 2019, **51**.
40. Wellmann R: **Optimum contribution selection for animal breeding and conservation: the R package optiSel**. *Bmc Bioinformatics* 2019, **20**.

Tables:

Table 1 The rate of genetic gain (ΔG) and inbreeding (ΔF) for different methods

Scenario	Selection Method*	ΔG		ΔF	
		mean	**SD**	**mean**	**SD**
1	Random	0.0013	0.0136	0.0107	0.0006
2	Truncation	0.3232	0.0162	0.0358	0.0078
3	Sirehalf-Damfull	0.2418	0.0128	0.0085	0.0006
4	Sirehalf-Damhalf	0.2741	0.0121	0.0089	0.0006
5	Sirehalf-Damtrunc	0.3007	0.0096	0.0105	0.0003
6	Sirehalf-DamfullRandom	0.1929	0.0118	0.0081	0.0006
7	Sirehalf-DamhalfRandom	0.1933	0.0114	0.0081	0.0006
8	Sirehalf-DamRandom	0.1984	0.0132	0.0086	0.0005
9	GOCS-0cM-p7	0.2957	0.0113	0.0119	0.0007
10	GOCS-0cM-p10	0.2782	0.0107	0.0085	0.0004
11	GOCS-1cM-p7	0.3027	0.0132	0.0121	0.0007
12	GOCS-1cM-p10	0.2820	0.0098	0.0085	0.0004
13	POCS-p7	0.3322	0.0122	0.0151	0.0013
14	POCS-p10	0.3230	0.0112	0.0110	0.0007

*: p7 means the penalty is 7 and p10 means the penalty is 10 in the OCS methods.

Table 2 The mean value of all genetic diversity metrics in all methods in the 20th generation

Scen-ario	Selection Method	H_e	H_o	A_e	M01	M05	IBD	varAdd	Nance-stor
1	Random	0.2549	0.2549	1.4308	0.7852	0.7080	0.1868	0.1745	109.7
2	Truncation	0.1758	0.1480	1.2945	0.5804	0.4941	0.5066	0.0543	70.4
3	Sirehalf-Damfull	0.2630	0.2619	1.4421	0.8217	0.7407	0.1522	0.1443	135.2
4	Sirehalf-Damhalf	0.2603	0.2591	1.4376	0.8145	0.7331	0.1587	0.1272	113.1
5	Sirehalf-Damtrunc	0.2502	0.2491	1.4209	0.7840	0.7032	0.1869	0.1103	108.1
6	Sirehalf-DamfullRandom	0.2666	0.2657	1.4484	0.8285	0.7493	0.1462	0.1697	134.8
7	Sirehalf-DamhalfRandom	0.2663	0.2663	1.4480	0.8278	0.7491	0.1472	0.1643	118.5
8	Sirehalf-DamRandom	0.2637	0.2624	1.4440	0.8178	0.7400	0.1556	0.1569	116.7

9	GOCS-0cM-p7	0.2490	0.2503	1.4225	0.7593	0.6845	0.2090	0.1151	107.5
10	GOCS-0cM-p10	0.2681	0.2693	1.4527	0.8251	0.7459	0.1546	0.1367	128.5
11	GOCS-1cM-p7	0.2484	0.2499	1.4215	0.7568	0.6831	0.2117	0.1079	105.8
12	GOCS-1cM-p10	0.2682	0.2695	1.4529	0.8258	0.7469	0.1541	0.1305	128.9
13	POCS-p7	0.2249	0.2239	1.3788	0.7107	0.6273	0.2583	0.0826	103.6
14	POCS-p10	0.2459	0.2452	1.4134	0.7812	0.6894	0.1935	0.0975	128.1

Note: Ae, Population effective alleles; varAdd; additive genetic variances; Nancestor, the number of ancestors;

Figures:

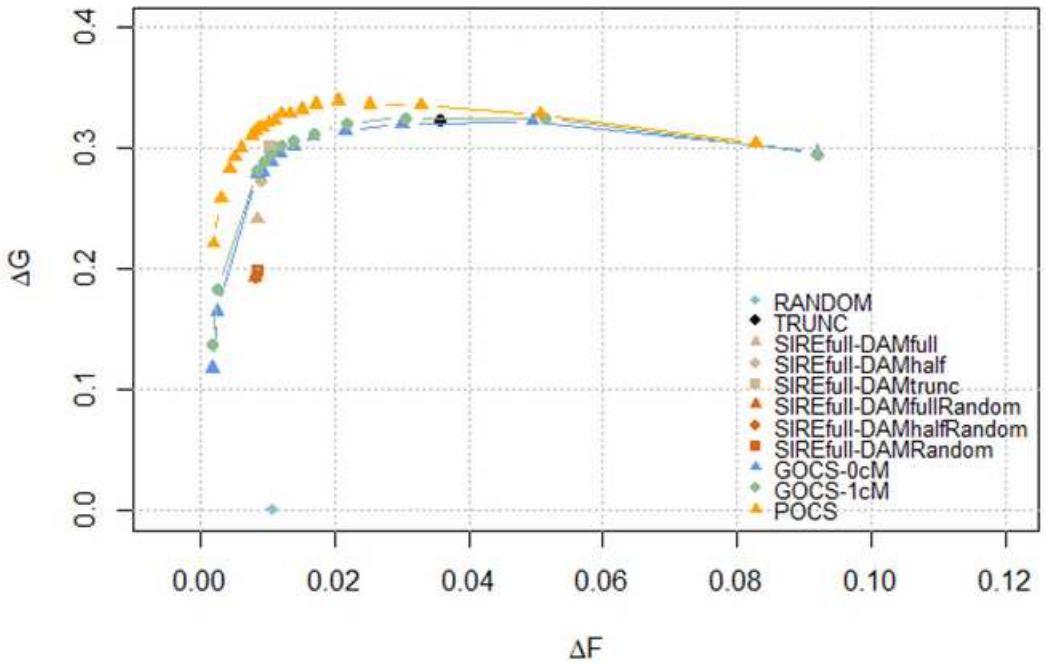


Fig. 1 The genetic gain and rate of inbreeding for different methods

Note: In OCS scenarios, different points represent different penalties. For GOCS, the penalties represented by each point from right to left are 1-10, 25, 50, 100. For POCS, the penalties represented by each point from right to left are 1-15, 25, 50, 100.

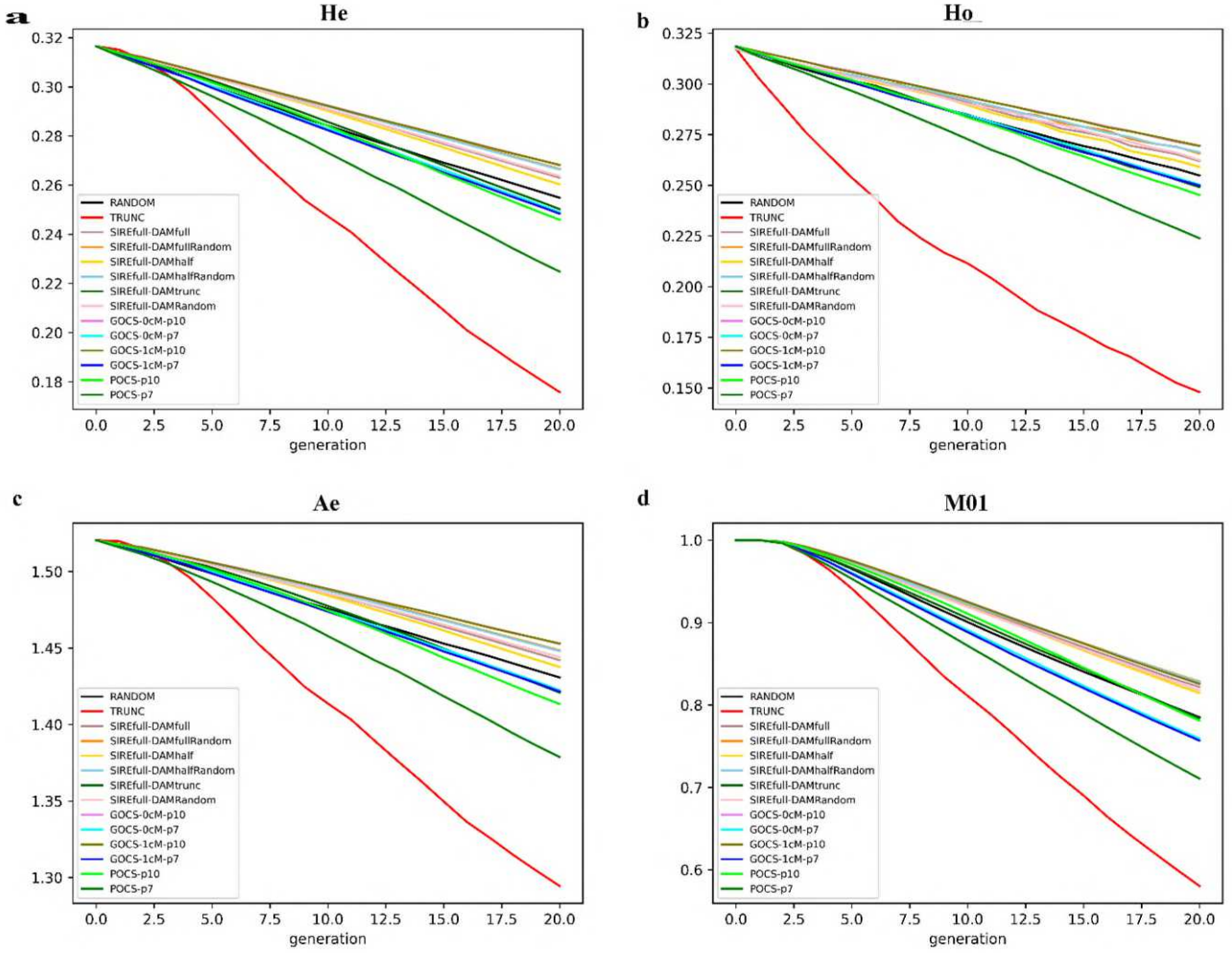


Fig. 2 The trends of genetic diversity metrics for different methods across 20 generations

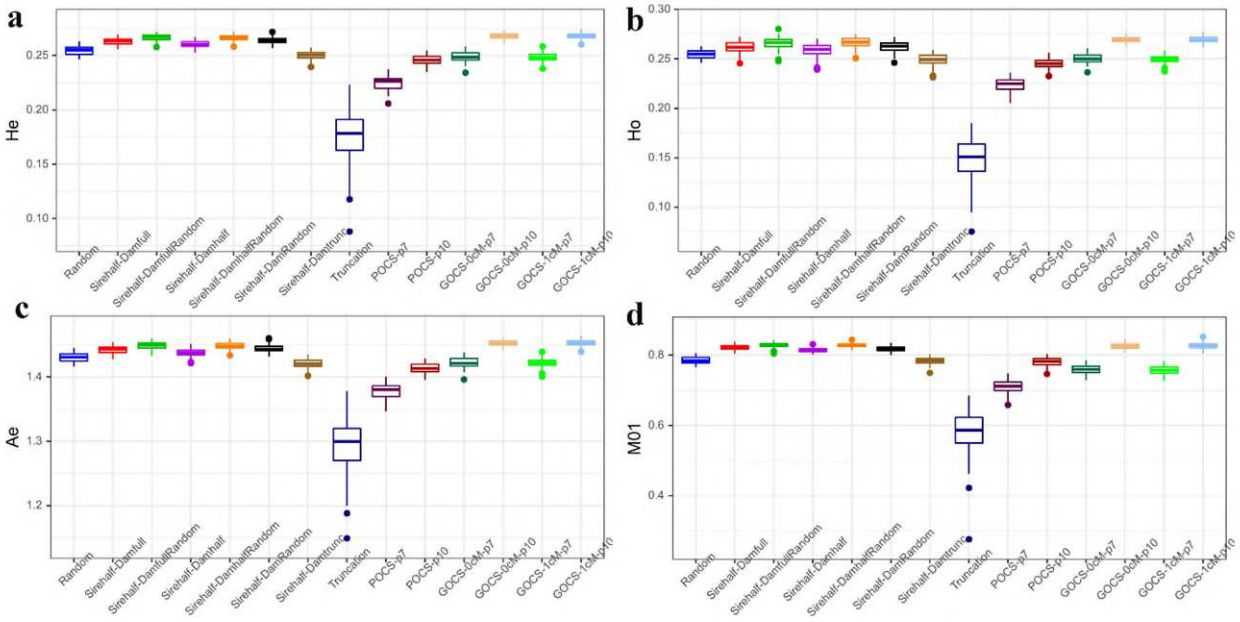


Fig. 3 The boxplots of genetic diversity metrics for different methods in 20th

generations

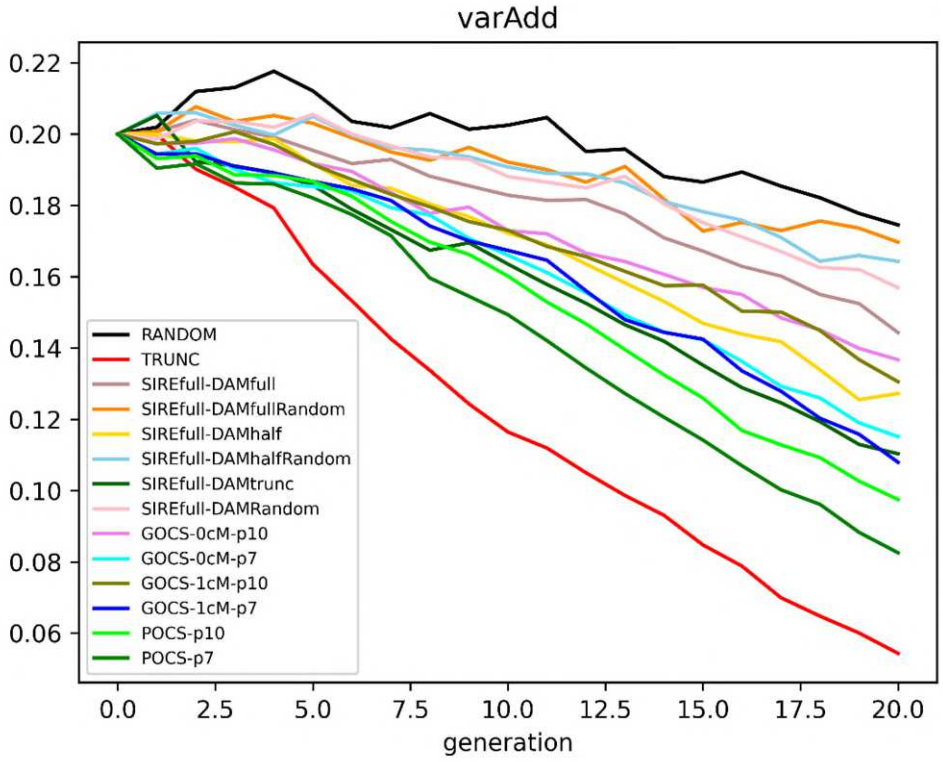
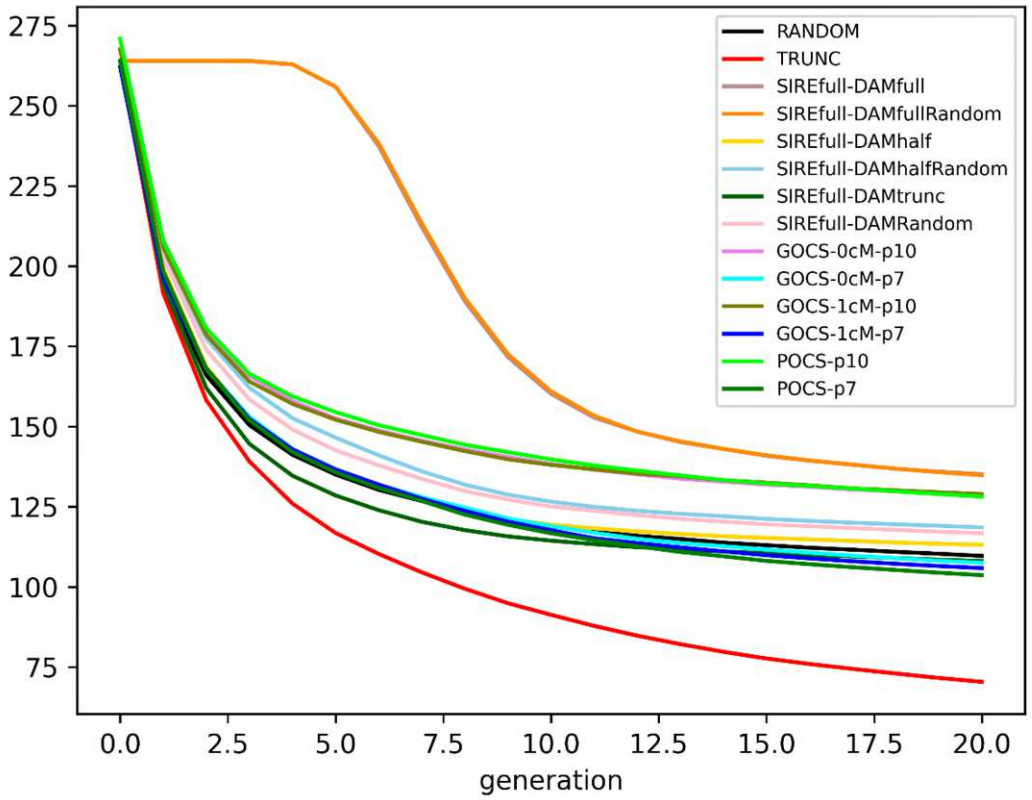


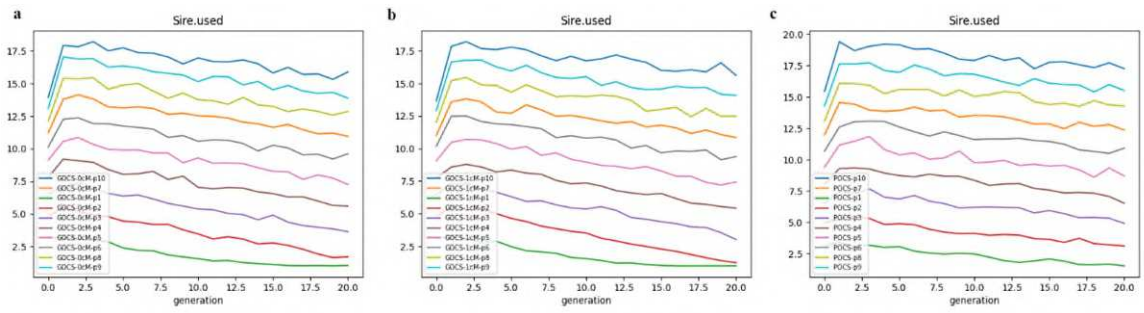
Fig. 4 The trends of additive variance (varAdd) for different methods across 20 generations



578

579 **Fig.5** The number of ancestor trends for different methods across 20 generations

580



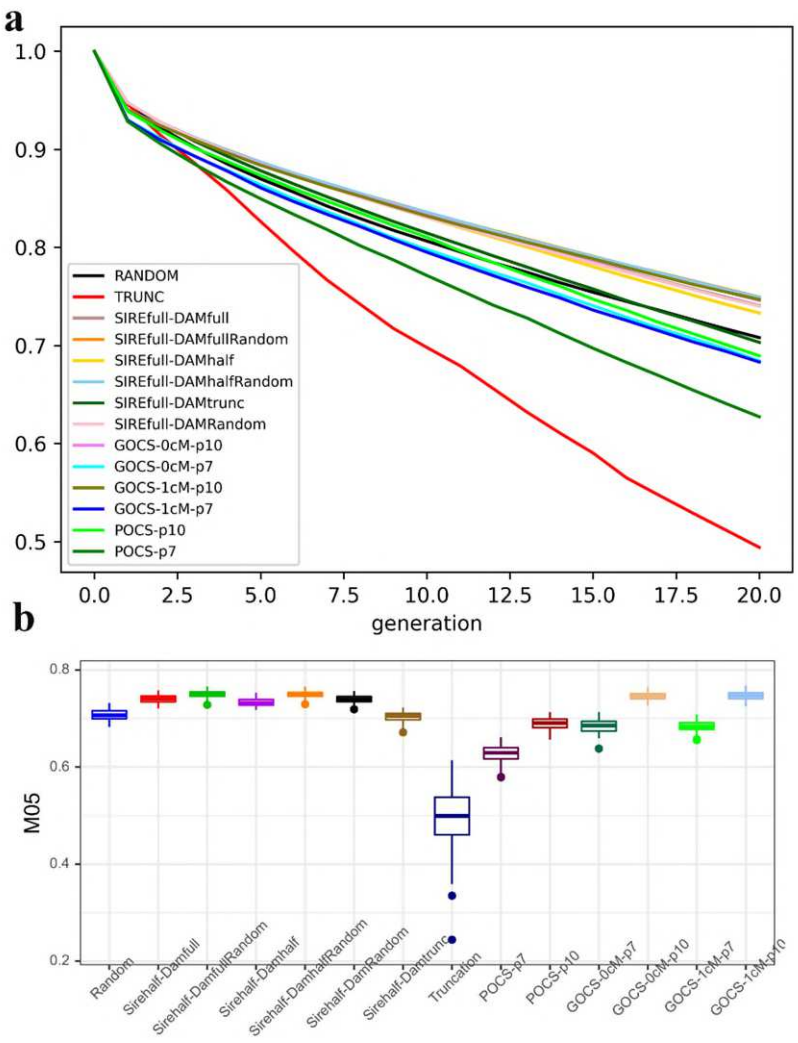
581

582 **Fig. 6** The number of sires used in OCS methods with penalty (1 to10) for 20
583 generations (**a**, GOCS-0Cm; **b**, GOCS-1cM; **c**, POCS)

584

585 # Additional files

586 **Additional file 1:**



Supplementary Fig. 1 a, M05 trend of different methods for across generations; **b**, M05 of different methods in 20th generation

Additional file 2: The mean value of all genetic diversity metrics in all methods across 20 generations

Additional file 3: The number of sires used in three OCS methods with all penalties across 20 generations

Figures

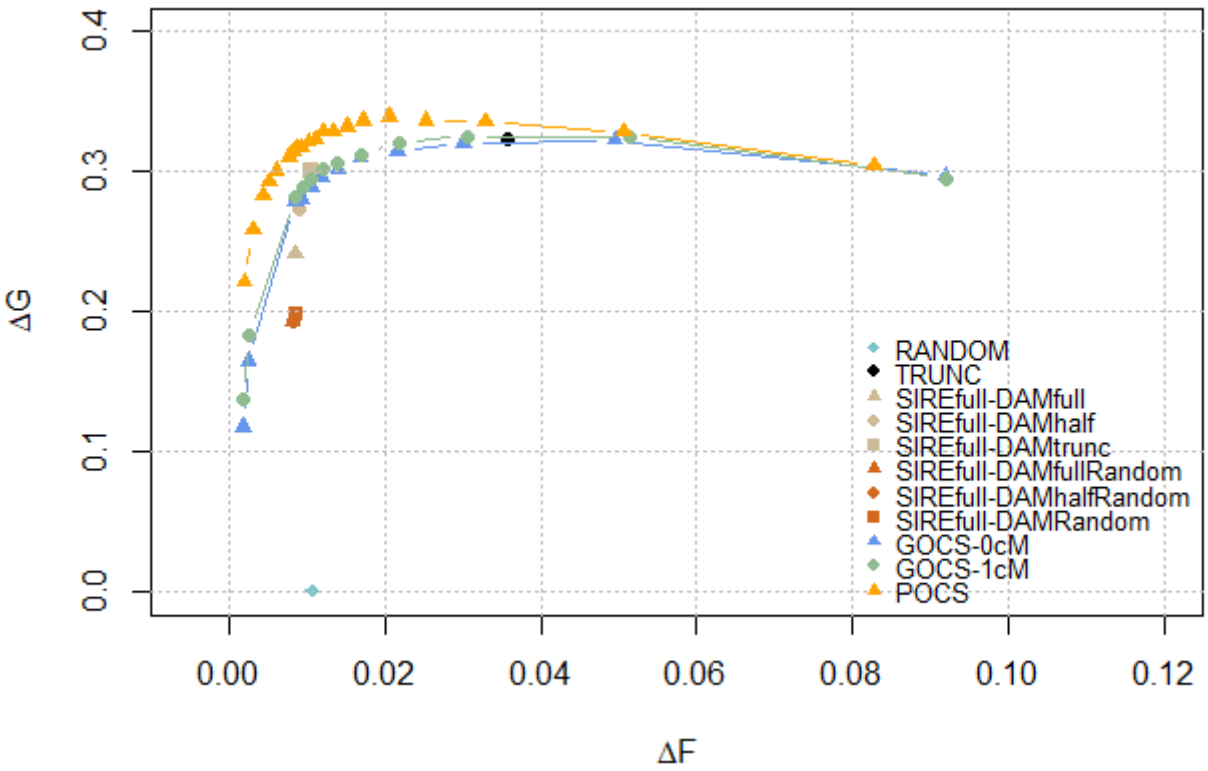


Figure 1

The genetic gain and rate of inbreeding for different methods Note: In OCS scenarios, different points represent different penalties. For GOCS, the penalties represented by each point from right to left are 1-10, 25, 50, 100. For POCS, the penalties represented by each point from right to left are 1-15, 25, 50, 100.

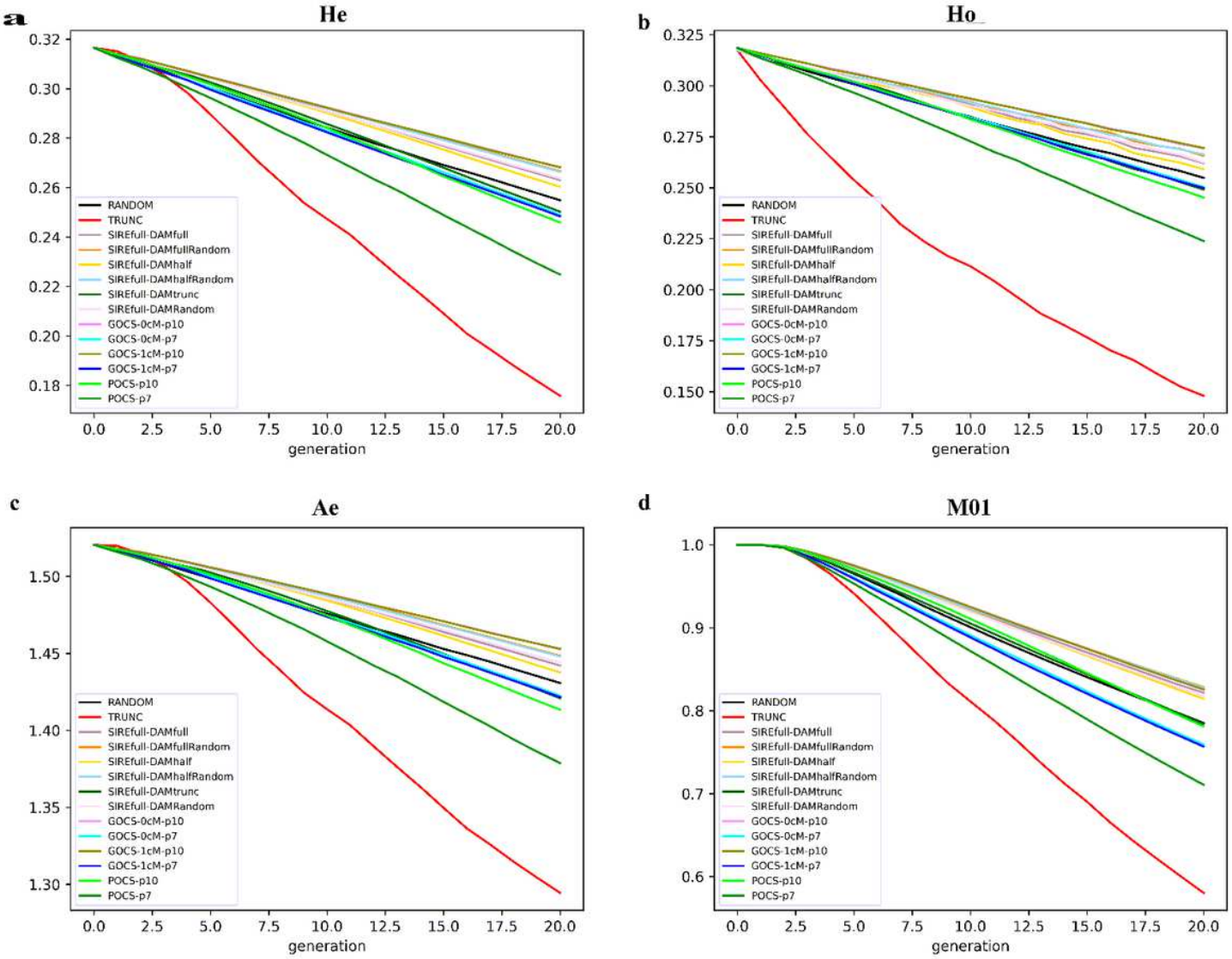


Figure 2

The trends of genetic diversity metrics for different methods across 20 generations

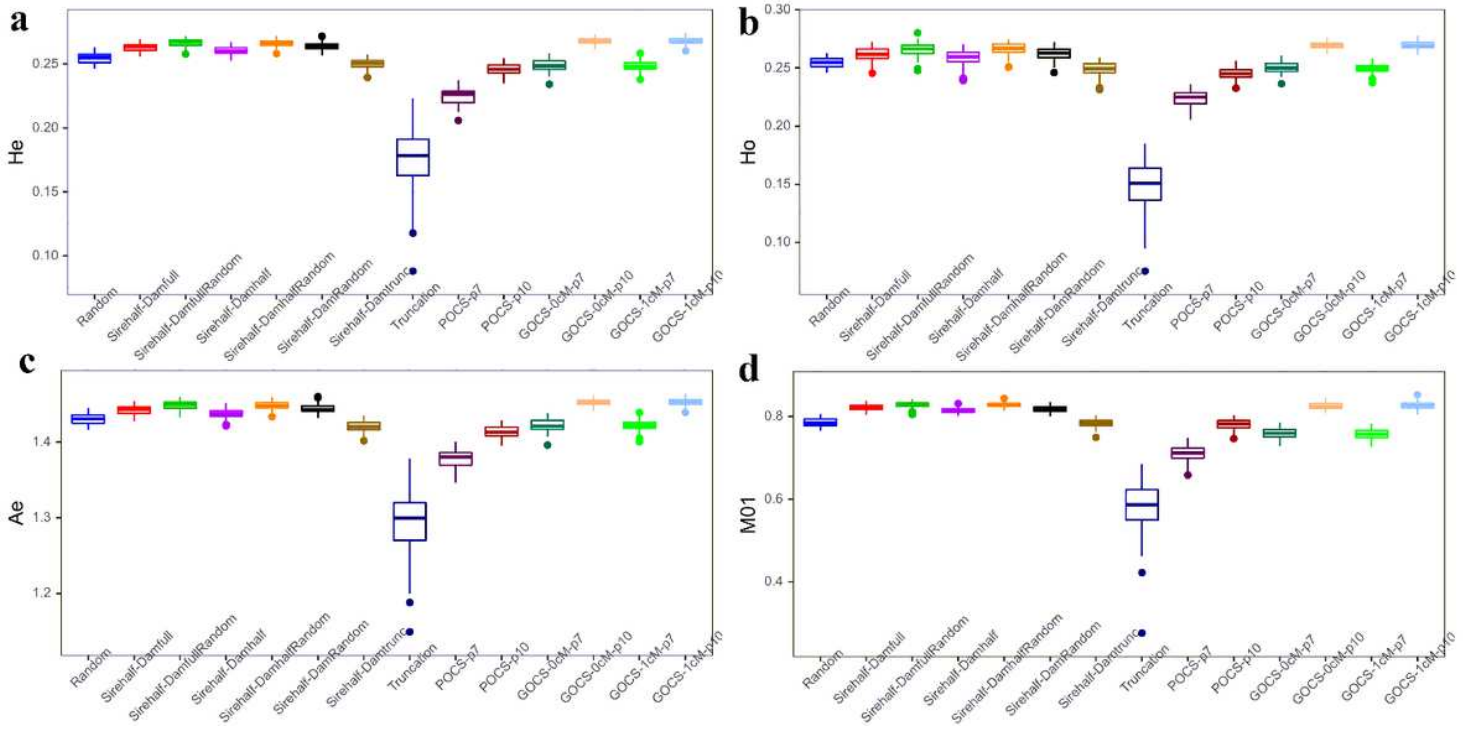


Figure 3

The boxplots of genetic diversity metrics for different methods in 20th generations

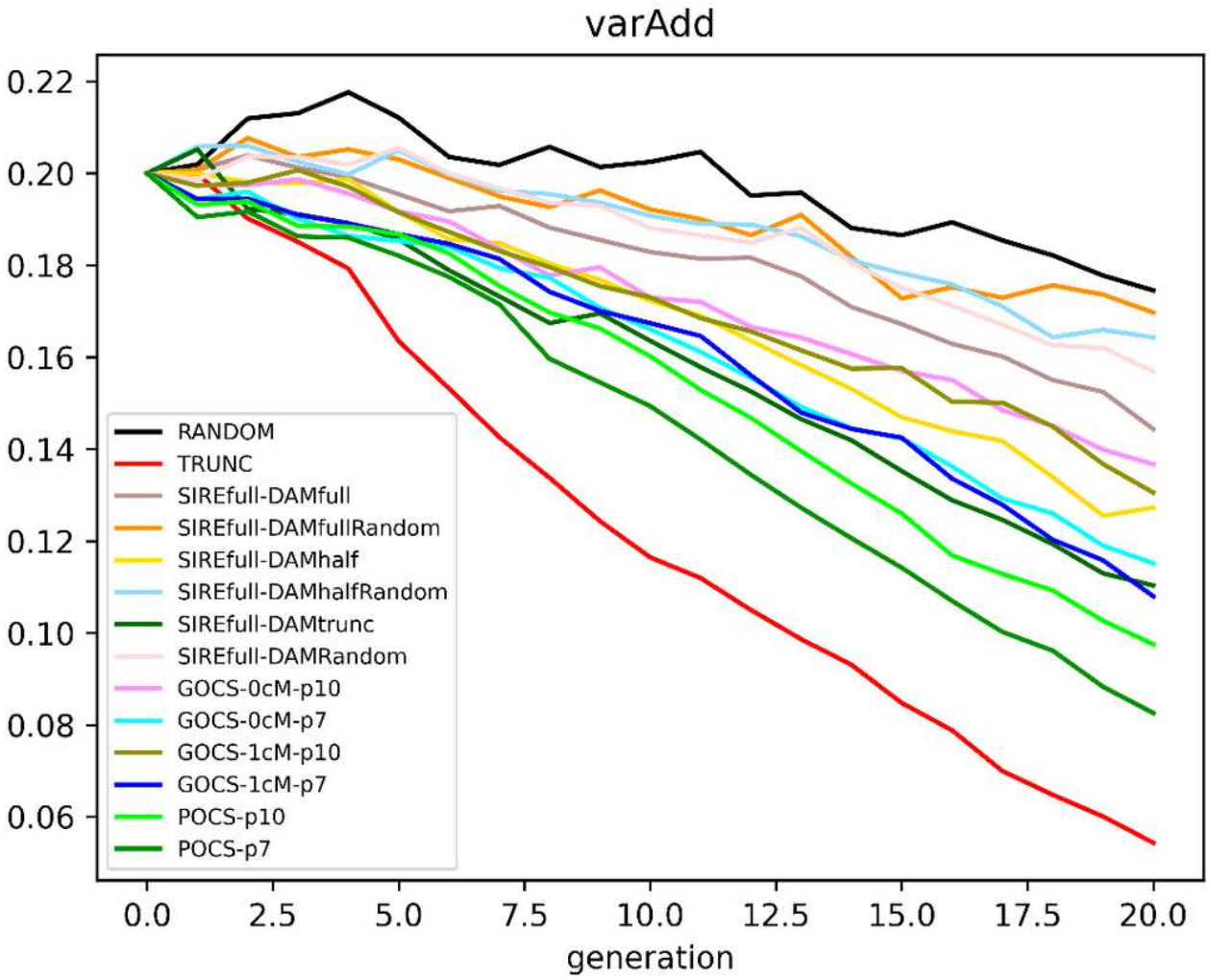


Figure 4

The trends of additive variance (varAdd) for different methods across 20 generations

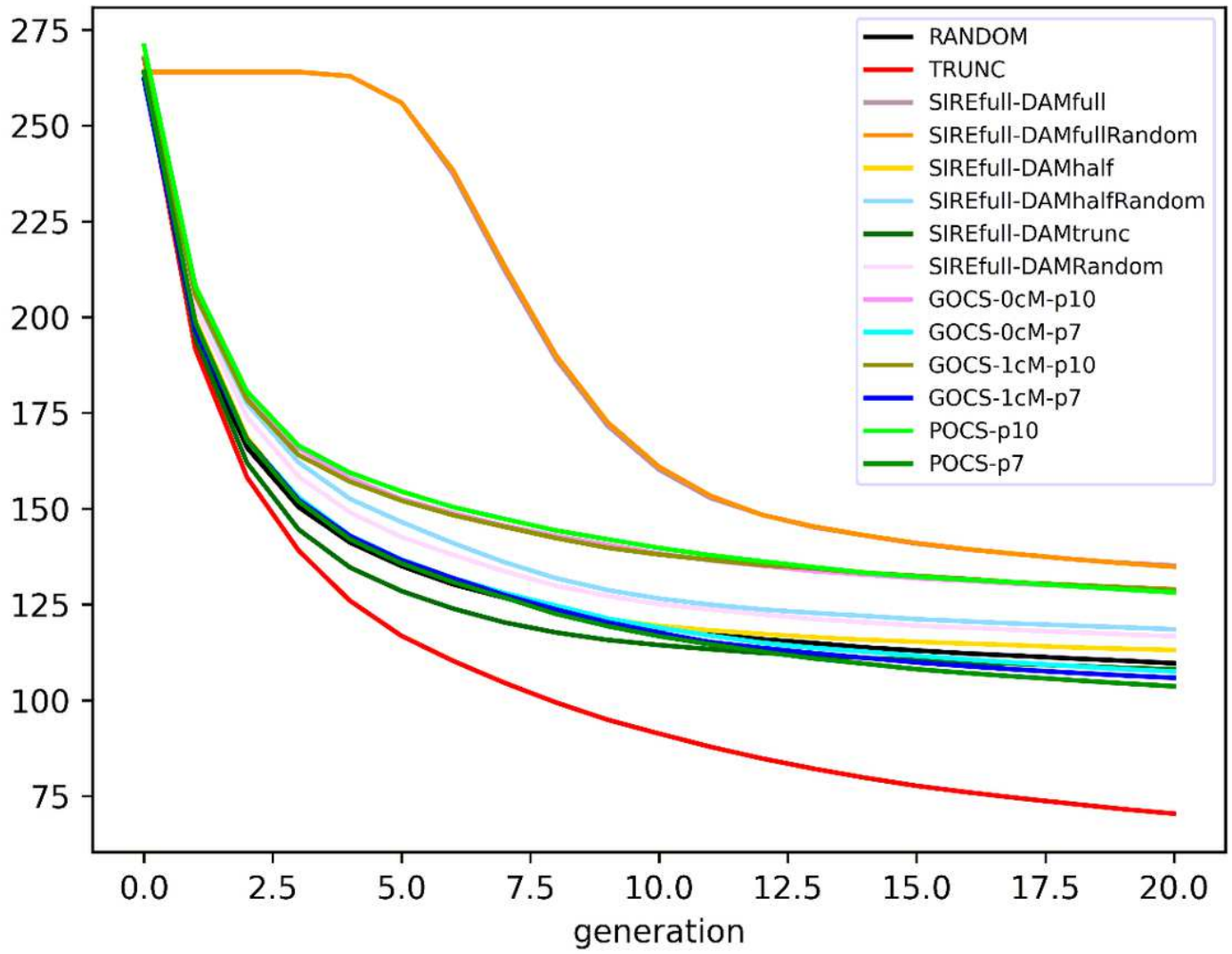


Figure 5

The number of ancestor trends for different methods across 20 generations

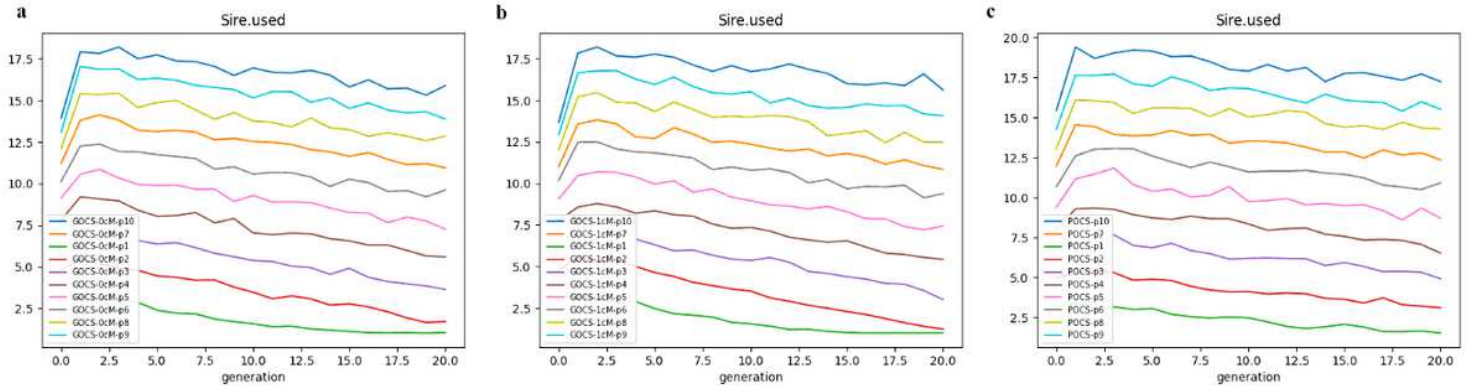


Figure 6

The number of sires used in OCS methods with penalty (1 to10) for 20 generations (a, GOCS-0Cm; b, GOCS-1cM; c, POCS)