Managing the genetic diversity of animal populations using cryobanks: optimizing the constitution of *ex situ* collection?

Alicia Jacques[1,2*], Michèle Tixier-Boichard[1] and Gwendal Restoux[1]

[1] Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 – Jouy-en-Josas, France

[2] Eliance, 149 rue de Bercy, 75012 – Paris, France

*Corresponding author

Correspondence: alicia.jacques@inrae.fr

Abstract

The genetic diversity of populations is important for their long-term conservation and adaptive potential. Nevertheless, there is an erosion of genetic variability in all populations, whether a small population under conservation or a large breed under selection. Many countries have conserved animal genetic resources *ex situ*, with the creation of cryopreserved collections. These cryopreserved genetic resources could open up new prospects for managing the genetic diversity of animal populations. Currently, there are as many strategies for conserving cryopreserved genetic resources as there are cryobanks. The aim of our study is to examine some parameters for optimizing the constitution of a gene bank.

Results

We performed simulations of animal populations (*i.e.* selection or conservation) using different types of breeding strategy and changing (i) the size and age of the cryopreserved collection or (ii) the management of cryopreserved collections (entry and exit of frozen semen).

We found that the age of genetic resources was the predominant factor for populations under selection, as this factor also quantifies the lag in breeding values. A more massive stocking of recent individuals seems to be the most relevant strategy for mobilizing *ex situ* genetic resources in a breeding scheme. For populations in conservation, older reproductive material becomes highly relevant. A cryobank with a large number of old individuals enables better management of matings and avoids too rapid an increase in inbreeding over the long term, while the use of a more moderate number of old individuals enables some genetic progress to be maintained, but offers fewer possibilities for managing genetic diversity. Continuous collection supply is a relevant option for population management. Whatever the breeding strategy, the use of OCS can help mobilize *ex situ* genetic resources in animal breeding programs.

Conclusions

Cryopreserved genetic resources are an invaluable help in the management of domestic populations. It is important to continue research in order to optimize the use of cryopreserved genetic resources, and to set up collection management recommendations that are relevant and adapted to breeding programs.

Keywords: gene bank, simulations, biodiversity, breeding scheme, genetic resources

Introduction

Populations must be able to evolve in various situations if they are to survive and not become extinct. Genetic diversity is an essential component in the evolution and adaptation of populations. Many studies have investigated the impact of a genetic variability loss in populations at different time scales (Frankham, 2005; Wiggans et al., 2017; Kardos et al., 2021). Indeed, in the long term, allele variability favors the adaptive potential of a population (Notter, 1999; Reed and Frankham, 2003; de Cara et al., 2013). The more different genotypes there are in a group of individuals, the higher the probability that some of these genotypes will be advantageous in the face of a disturbance (*e.g.* emergence of a new pathogen, new environmental constraints *etc...*). In the medium term, the erosion of genetic diveristy has an impact on the evolution of selection, and in particular on genetic progress for domestic populations under selection (Kinghorn et al., 2009; Meuwissen et al., 2013; Woolliams and Oldenbroek, 2017). Finally, on a short timescale, reduced variability leads to inbreeding depression, *i.e.* a reduction of selective value (O'Grady et al., 2006; Charlesworth and Willis, 2009; Leroy, 2014). Consequently, whatever the type of population, conserving genetic diversity is an extremely important issue.

For several decades, initiatives have been taken to conserve the genetic diversity of domestic animal breeds in different countries. Animal genetic resources, defined as reproductive material (*e.g*. semen, oocytes) (Vavilov, 1931; Convention on Biological Diversity (CBD), 1992), have been preserved through the development of gene banks. In this way, reproductive material for many animal species and populations has been *ex situ* cryopreserved since several generations. Some studies, based on datasets from real populations, have shown that the use of genetic resources in cryobanks can help manage genetic variability and genetic progress in the next generation (Doekes et al., 2018; Eynard et al., 2018). A simulation studie had shown that the use of old frozen semen can help to manage genetic diversity within a cattle breeding scheme at long term (Leroy et al., 2011). However, despite these few studies, there are currently very few concrete examples of the cryopreserved genetic resources use in breeding schemes. To our knowledge, the mobilization of *ex situ* genetic resources has been described for two local pig breeds in France (Mercat et al., 2008), the use of old cryopreserved bulls has locally reintroduced genetic diversity on the Y chromosome in Holstein population in the United States (Dechow et al., 2020) and the semen of an old bull has been used to reintroduce diversity into local French dairy cattle breeds (Jacques et al., 2023).

The Food and Agriculture Organization of the United Nations (FAO) has identified cryoconservation of genetic resources as an essential element in national conservation strategies (FAO, 2007; FAO, 2015). Over the last few years, germplasm collections have been developed in different country (Blackburn, 2018; Leroy et al., 2019). In 2016, the IMAGE project (Innovative Management of Animal Genetic Ressources) goal was to improve the use of genetic collections and the management of animal genebanks. In the same way, the CRB-Anim project (French Biological Animal Resource Center, https://doi.org/10.15454/1.5613785622827378E12) has enabled the enrichment of cryopreserved collections for various animal species and breeds. These various actions have contributed to the developpement of methodological and technical skills associated with the development and management of germplasm collections, or to providing financial support for the collections of somes samples. For instance, the French National Cryobank, created in 1999, contains currently a large variety of *ex situ* genetic resources, which can help to manage the genetic diversity of contemporary populations (Jacques et al., 2024a).

Although the FAO has proposed recommendations on the genetic material required to reconstitute a breed (FAO, 2012), there are no other recommendations for the creation of a cryobank. Presently, there are as many strategies for the sampling of *ex situ* genetic resources as there are germplasm collections. However, maintaining cryopreserved genetic resources generates continually different costs (*e.g.* economic, energy and spatial for the storage) (Pizzi et al., 2016; De Oliveira Silva et al., 2019). In the long term, these constraints may influence the cryopreserved material quantity and cryobank's operation. For exemple, presently there is no limit concerning the quantity of genetic resources that can be stored in a cryopreserved collection. Consequently, investigating the optimization of *ex situ* collections constitution to manage livestock populations over the long term could enable to anticipate solutions facing with emergence of these operating constraints. To our knowledge, no study has documented sampling strategies for animal genetic resources according to breeding programs. However, it is necessary to envisage strategies for collections management considering their future use or anticipated costs. In our

study, we propose to study the use of various cryobanks, which differ in terms of constitution, in order to provide some initial answers.

Methods

We performed stochastic simulations of different breeding scheme using the R package MoBPS, version 1.6.64 (Pook et al., 2020; Pook, 2021). Some simulation steps follow the same protocol as Jacques *et al*., 2024 (Jacques et al., 2024b) and will be described in the same way (*e.g.* pre-burn-in, burn-in and the size of breeding scheme).

R version 4.2.1 (R Core team, 2020) and ggplot2 R packages (Wickham, 2011) were used to perform all additional calculations and graphical representations. Statistical tests were performed using the lm function, post-hoc tests were conducted using the emmeans package (Lenth et al., 2021), and type II ANOVA were performed using the car package (Fox et al., 2019).

Definition of scenarios: type of selection, prolificity and uses of germplasm collections

We defined four types of scenarios that differed by the selection process, considering two traits of interest, thereafter named Trait 1 and Trait 2. The scenario pop_rm corresponds to a random choice of parents among the candidates; the scenario pop_max_BV corresponds to a choice of parents so as to maximize genetic progress on a synthetic index combining both traits; the scenario pop_max_DG corresponds to a choice of parents so as to minimize the loss of genetic diversity; the scenario pop_OCS corresponds to a choice of parents which maximizes genetic progress under the constraint of a maximal increase in kinship of 0.5% per generation, as recommended by the FAO (FAO, 1998). The pop_max_BV, pop_max_DG and pop_OCS scenarios were modeled using the optimal contribution strategy (OCS) (Meuwissen, 1997) implemented in MoBPS via the optiSel package, version 2.0.5 (Wellmann, 2019; Wellmann, 2021). The relationship matrix in the OCS was performed using the vanRaden method (VanRaden, 2008). For the three scenarios involving the use of OCS, we set a uniform constraint for the female kernel, allowing us to consider their contributions to be equal. The maximum contribution of each male candidate was set at 5% for the pop_max_BV scenario, or 1% for the pop_max_DG and pop_OCS scenarios, in order to obtain a realistic minimum number of fathers at each generation. For the pop_OCS scenario, an upper bound to restrict the average kinship in the progeny is calculated for each generation, based on an estimate of the average kinship within the male and female kernels proposed as candidates and prohibiting an increase beyond 0.5%.

The effect of prolificity was studied by simulating three levels: 1, 2 or 10 offsprings per mating. These values were chosen to cover the majority of breeding livestock species (ruminants, horse as well as pig).

Finaly, these scenarios were implemented either with the contemporary kernels, but also by adding different types of germplasm collection (defined below) as candidates for reproduction. In this last case, cryopreserved individuals were systematically re-evaluated at the same time as contemporary individuals.

Initialization of simulated populations and pre-burn-in

Genetic architecture and definition of a synthetic index

Firstly, we simulated a founder population with 5000 genetic markers (SNPs) distributed over 4 chromosomes for 500 individuals with a sex ratio of 0.5. SNPs were regularly distributed along the genome, *i.e.* an average distance in centimorgans between neighboring SNPs equal to 4.105 cM.

Each trait was simulated with 500 additive quantitative trait loci (QTLs), randomly distributed along the genome, the effects of which are drawn from a Gaussian distribution. The parameters for the two traits were initialised with a genetic mean of 0, a variance of 1, a negative correlation of -0.3 and a heritability of 0.4. The true breeding values (BV) were generated by the genotype simulation and used to standardise the EBV for each trait. Then, EBVs were combined in a synthetic index with a weight of 0.8 for Trait 1 and a weight of 0.2 for Trait 2. Thus, for each animal, the index is obtained as follows:

$$(1) \quad Index_i = weight_{Trait1} \frac{EBV_{Trait1_i} - \overline{EBV_{Trait1_{cand}}}}{sd(BV_{Trait1_{cand}})} + weight_{Trait2} \frac{EBV_{Trait2_i} - \overline{EBV_{Trait2_{cand}}}}{sd(BV_{Trait2_{cand}})}$$

With $Index_i$ the synthetic index value of animal *i*, EBV_{Trait1_i} and BV_{Trait1_i} respectively the estimated breeding value and the true breeding value of animal *i* for Trait1 (idem for Trait 2), $weight_{Trait1}$ and $weight_{Trait2}$ the weights of Trait1 and Trait2 in the selection objective, respectively (*i.e.* here 0.8 and 0.2).

The candidate sires were ranked according to this index. The estimated genetic values of individuals (EBV) were obtained using GBLUP evaluation models. The SNPs involved in the QTLs of the two traits were not taken into account in the evaluations, as is generally the case with the SNP chips used routinely in breeding schemes, for which the causal SNPs are not available. The reference population for the estimates consisted of all phenotyped individuals available at the time of evaluation.

Pre-burn-in and initial population creation

We performed a pre-burn-in process to generate linkage disequilibrium and genetic structure for 20 simulated populations (corresponding to each of the 20 replicates for the different scenarios that will follow). To do this, we simulated 20 founder kernels of 250 males and 250 females, then random breeding cycles for 10 generations. Each new cohort included 250 males and 250 females (candidates to selection) resulting from the mating of the 20 best males and the 250 best females of the previous cohort with prolificity set at 2, prohibiting half-sib matings, and with a selection of sires based on the previously defined synthetic index (Figure 1). At the end of these 10 generations of pre-burn-in, the aim was to obtain a population with an average kinship of 10% in generation 10.

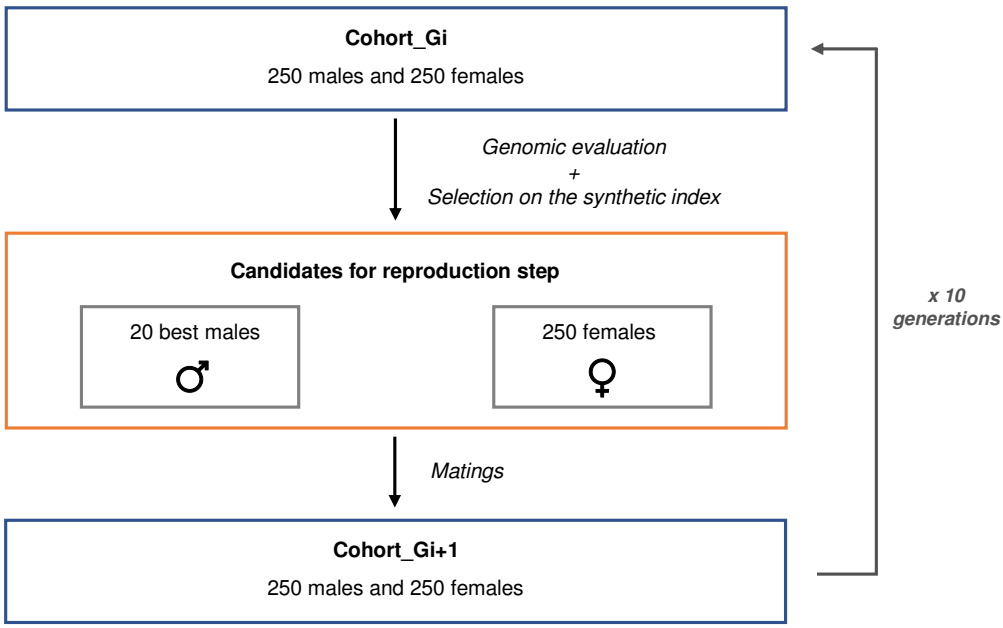


Figure 1 - Breeding steps for the pre-burn-in of simulations.

Burn-in phases, breeding scheme and definition of cryopreserved individuals

Burn-in and simulated population size

Secondly, following the pre-burn-in, we performed two types of burn-in: one to mimic a selected population and the other to mimic a conserved population.

We have simulated a breeding program livestock breeding in generations overlapping. Each new cohort was composed of 500 males and 500 females, obtained from matings between a kernel of 100 males and 800 females from the last two generations, with a prolificity always set at 2. Individuals retained for reproduction were selected (i) on the basis of the synthetic index defined above for selection burn-in or (ii) by random sampling in the case of conservation burn-in (*i.e.* only effect of genetic drift). These burn-in were performed over 10 generations to obtain an average kinship of (i) 14/15% in the case of the selection scheme or (ii) 10/11% in the case of the conservation scheme for the 20th generation.

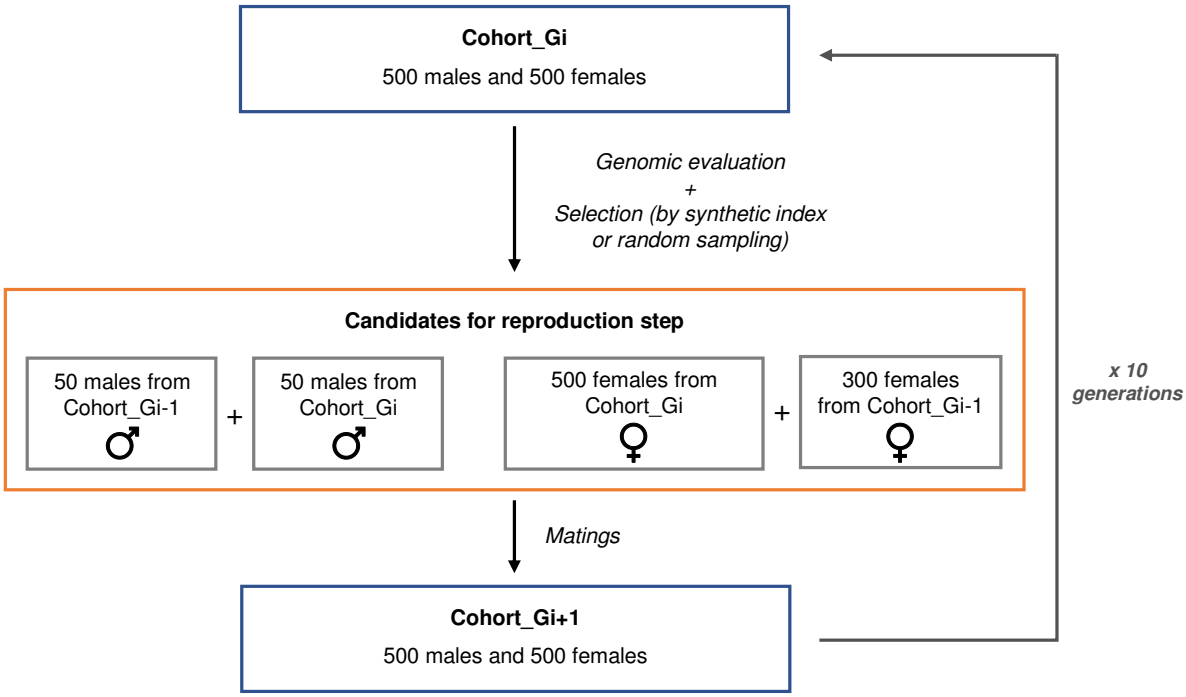


Figure 2 - Breeding steps for the selection or conservation burn-in.

Definition of germplasm collection types

Fixed cryobanks

Firstly, different types of cryobank were defined according to the age of the cryopreserved sires: (a) an old cryobank composed of individuals cryopreserved from generations 10, 11, 12 and 13, (b) a recent cryobank composed of individuals cryopreserved from generations 15, 16, 17 and 18 and (c) a mixed cryobank composed of individuals cryopreserved from generations 11, 13, 15 and 17. The aim was to study whether the age of the cryopreserved collection had an impact on the use of *ex situ* genetic resources.

Secondly, during the burn-in steps (for the selection or conservation population), some individuals were randomly selected from sires of each generation defined above to create the cryopreserved collections. We varied the number of individuals cryopreserved in each generation, storing 5, 10 or 20 individuals. In this way, three total cryobank sizes were obtained, with 20, 40 or 80 cryopreserved individuals respectively. The aim was to study whether the size of the cryopreserved collection had an impact on the use of *ex situ* genetic resources.

Mobile cryobanks

We defined mobile cryobanks, whose composition varied over time (*i.e.* with each generation). In this case, the collections had a fixed size but a variable composition over time, as they were regularly updated, with older individuals being eliminated in favor of more recent ones joining the collections. This mimics the dynamic management of fixed-budget or fixed-capacity collections.

Two types of mobile cryobank were created: a recent mobile cryobank and a continuous mobile cryobank, both of which contained a total of 40 individuals for a given generation. For each generation, noted *gen*, the recent mobile cryobank was composed of ten cryopreserved individuals from the last four generations that were no longer proposed as sires to create the *gen*+1 cohort : *gen*-2, *gen*-3, *gen*-4 and *gen*-5. Similarly, the continuous mobile cryobank contained the five cryopreserved individuals from the last eight generations which were no longer proposed as sires to create the *gen*+1 cohort : *gen*-2, *gen*-3, *gen*-4, *gen*-5, *gen*-6, *gen*-7, *gen*-8 and *gen*-9. An example of fixed cryobank and dynamic cryobank for the creation of two different generations is shown in Additional File 1 Figure S1.

Finally, a final type of mobile cryobank has been created, whose size varies from generation to generation: the cumulative mobile cryobank. For this collection, the ten individuals from *gen*-2 were added

to the collection at each generation. In this way, the cryobank was enriched every generation, while the other two types of mobile cryobank operated at constant size.

The different cryobanks were then used in different management programs with different objectives from generation 21 onwards. The results were compared to identify which types of collections best met the expectations of the different objectives.

Indicators of scenarios, genetic diversity and uses of different germplasm collections

Quantify the use of cryopreserved individuals

For each scenario involving the use of *ex situ* genetic resources, we quantified the effective use of cryopreserved individuals by counting their average number of direct descendants (*i.e.* sons or daughters), then indirect descendants (grandchildren, great-grandchildren...). Thus, for each rank (*i.e.* the hierarchical level separating a cryopreserved father and his descendants, *e.g.* rank 1 for sons/daughters, rank 2 for grandchildren, rank 3 for great-grandchildren, *etc*....) the number of descendants resulting from the use of cryopreserved individuals was monitored.

Measurements for performances

For each generation, the "true" genetic values (BV) of the two traits were calculated, together with the empirical genetic variances for these two traits (*i.e.* variance of the genetic values of the simulated individuals). The EBVs for traits 1 and 2, as well as the value of the sires' standardized synthetic index, were evaluated at each generation. We also calculated a value for the "true" selection index, *i.e.* based on EBVs and without standardization, for each generation by weighting their respective weights in the selection objective. In this way, we were able to compare the simultaneous evolution of the two traits.

Measurements for genetic diversity

The evolution of genetic diversity was studied by comparing its level at the start of the program (generation 20) with that obtained at the end of the program (generation 35). Average kinship was calculated for each generation using “kinship.emp.fast function” of the MoBPS package, which estimates the real kinship based on the recombination points of genotypes by sampling the population. This estimator is based on the notion of shared identical segments per progeny. From this, the average kinship increase (ΔK) was computed.

Finally, the genetic distance between generation 20 and generation 35 was estimated using Nei's method (Nei, 1972) for each scenario with MoBPS's "get.distance" function.

Results

Use of *ex situ* genetic resources in breeding schemes

For fixed *ex situ* collections, the age of the collections had no impact on their use in conservation programs aimed purely at optimizing genetic diversity (see Table 1 and Figure 3). Indeed, in these programs, all the proposed individuals are used, and the resulting number of offspring is significantly different according to the number of sires used (2-factor ANOVA test, F= 628.8, df=171, p-val<0.05 for pop_rm and F= 1188, df=171, p-val<0.05 for pop_max_DG). The larger the collection, the greater its contribution to reproduction.

Table 1 - Number of cryopreserved individuals used and production of direct offspring for selected population.

CRYOBANK TYPE	pop_rm		pop_max_DG		pop_OCS		pop_max_BV	
	No. cryopreserved sires used	No. offspring per cryopreserved sire	No. cryopreserved sires used	No. offspring per cryopreserved sire	No. cryopreserved sires used	No. offspring per cryopreserved sire	No. cryopreserved sires used	No. offspring per cryopreserved sire
20 cryopreserved sires								
Ancient collection	20 [sd=0]	126.6 [sd=3.65]	20 [sd=0]	217.7 [sd=12.25]	0 [sd=0]	NA	0 [sd=0]	NA
Mixed collection	20 [sd=0]	125.4 [sd=3.91]	20 [sd=0]	228.7 [sd=9.44]	0.05 [sd=0.22]	20 [sd=NA]	0 [sd=0]	NA
Recent collection	20 [sd=0]	127.1 [sd=4.74]	20 [sd=0]	232.2 [sd=9.93]	0.35 [sd=0.49]	32.86 [sd=25.61]	0.10 [sd=0.31]	114 [sd=36.8]
40 cryopreserved sires								
Ancient collection	40 [sd=0]	108 [sd=3.65]	40 [sd=0]	155 [sd=5.13]	0 [sd=0]	NA	0 [sd=0]	NA
Mixed collection	40 [sd=0]	107.8 [sd=2.41]	40 [sd=0]	160 [sd=5.19]	0.20 [sd=0.41]	20 [sd=7.12]	0 [sd=0]	NA
Recent collection	40 [sd=0]	108.9 [sd=3.56]	40 [sd=0]	168.5 [sd=5.03]	0.9 [sd=1.37]	25.12 [sd=11.29]	0.15 [sd=0.37]	107.3 [sd=12.86]
80 cryopreserved sires								
Ancient collection	80 [sd=0]	83.6 [sd=2.36]	80 [sd=0]	99.1 [sd=2.23]	0 [sd=0]	NA	0 [sd=0]	NA
Mixed collection	80 [sd=0]	83.6 [sd=2.25]	80 [sd=0]	103.2 [sd=2.47]	0.70 [sd=0.80]	23.80 [sd=15.40]	0 [sd=0]	NA
Recent collection	80 [sd=0]	83.5 [sd=2.38]	80 [sd=0]	108.3 [sd=2.36]	1.80 [sd=1.36]	24.51 [sd=9.26]	0.45 [sd=0.83]	98.3 [sd=8.82]
Mobile recent	171.70 [sd=2.05]	25.20 [sd=0.51]	178.60 [sd=1.14]	32.73 [sd=0.64]	11.10 [sd=2.81]	26.15 [sd=3.98]	0.40 [sd=0.60]	111.71 [sd=15.98]
Mobile continuous	87.65 [sd=1.09]	44.85 [sd=1.54]	89.10 [sd=0.91]	66.61 [sd=1.84]	3.80 [sd=1.61]	25.98 [sd=8.29]	0.05 [sd=0.22]	108 [sd=NA]
Mobile cumulative	199.80 [sd=3.04]	44.01 [sd=1.10]	215.05 [sd=1.61]	47.71 [sd=0.62]	8.40 [sd=2.95]	24.24 [sd=4.39]	0.20 [sd=0.41]	106 [sd=20.72]

Note: Scenario abbreviations correspond to those developed in the Method section.
No. cryopreserved sires used and No. offspring par cryopreserved sire are average values computed with the 20 replicates

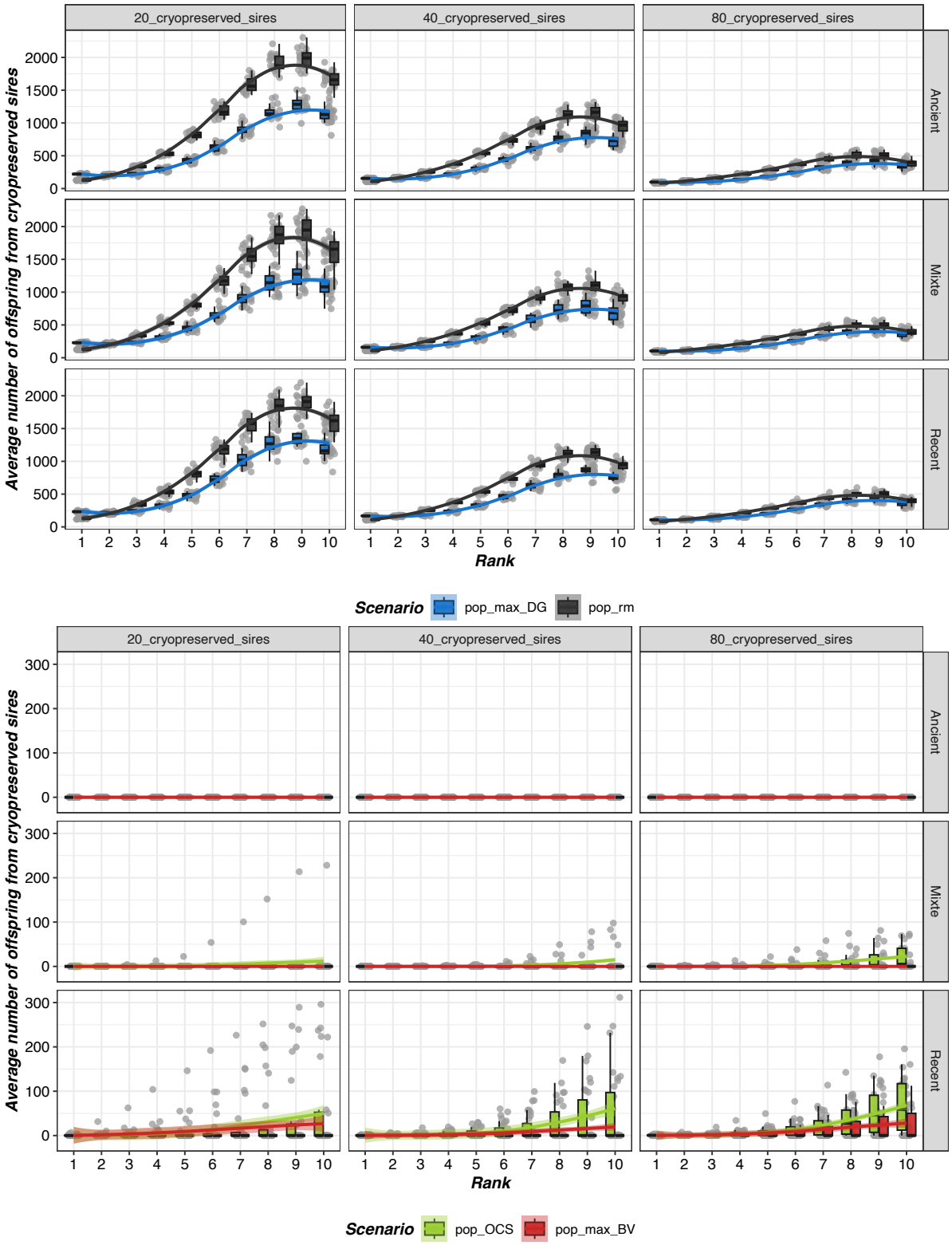


Figure 3 - Number of offspring from cryopreserved individuals in different management programs and scenarios over generations with fixed collections.

The x-axis corresponds to the genealogical rank of individuals in relation to their relatives from ex situ collections (i.e. rank 1 for children, rank 2 for grandchildren, rank 3 for great-grandchildren...).

On the other hand, in the case of the selection programs (*i.e.* pop_max_BV and pop_OCS), the most recent collections were used more, while the oldest collections were hardly used at all (see Table 1). *Ex situ* collections that were more extensive in time were used intermediately. Less markedly, the size of the cryopreserved collection had a positive impact on its use in breeding programs (3-factor ANOVA test, F=16.84, df=346, p-val<0.05). This trend is confirmed with the cumulative collection, which is the largest and most extended in time, and for which the cryopreserved individuals used are mainly the most recent (see Figure 4). The other two dynamic cryobanking strategies confirm this trend, with recent mobile collections being much more widely used than extended mobile collections (test ANOVA à 2 facteurs, F= 137.3, df=114, p-val<0.05), which include more older generations but with a lower number of individuals cryopreserved per generation. However, it should be noted that, on average, few individuals are used in breeding programs, particularly when these do not incorporate constraints on genetic diversity. Nevertheless, the use of dynamic collections drastically increases their utilization compared with fixed collections (Table 1).

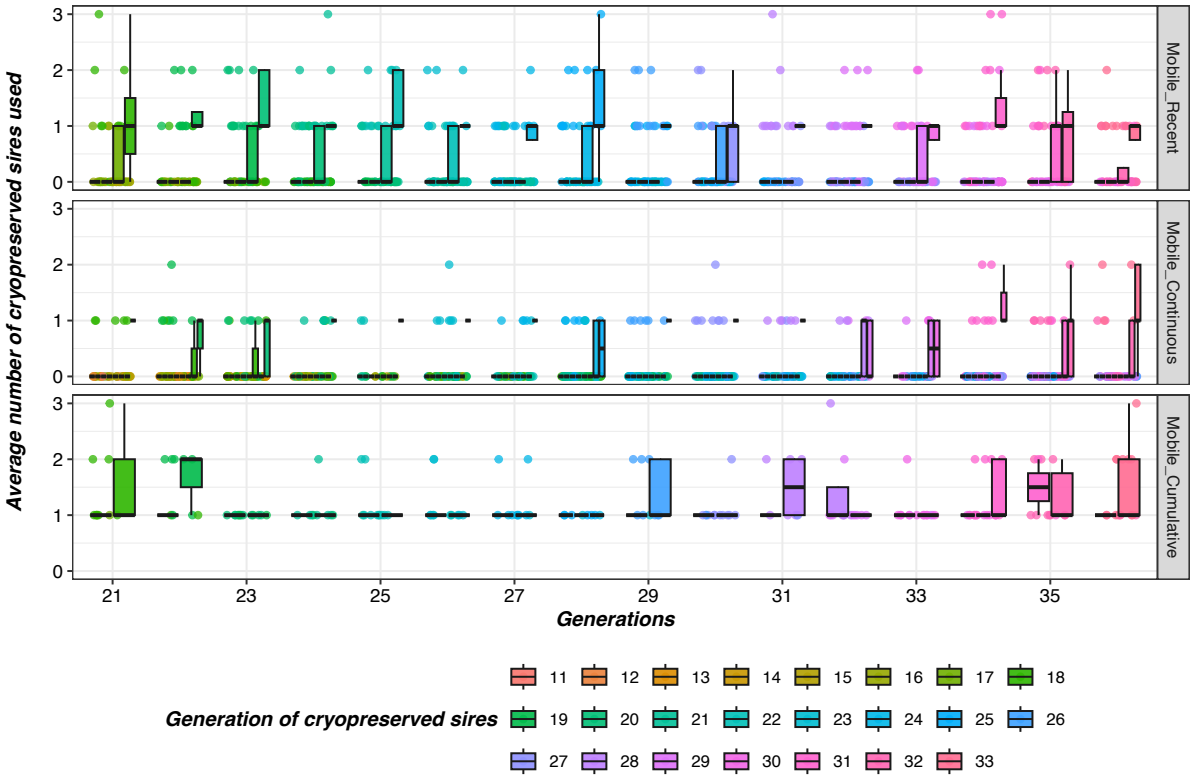
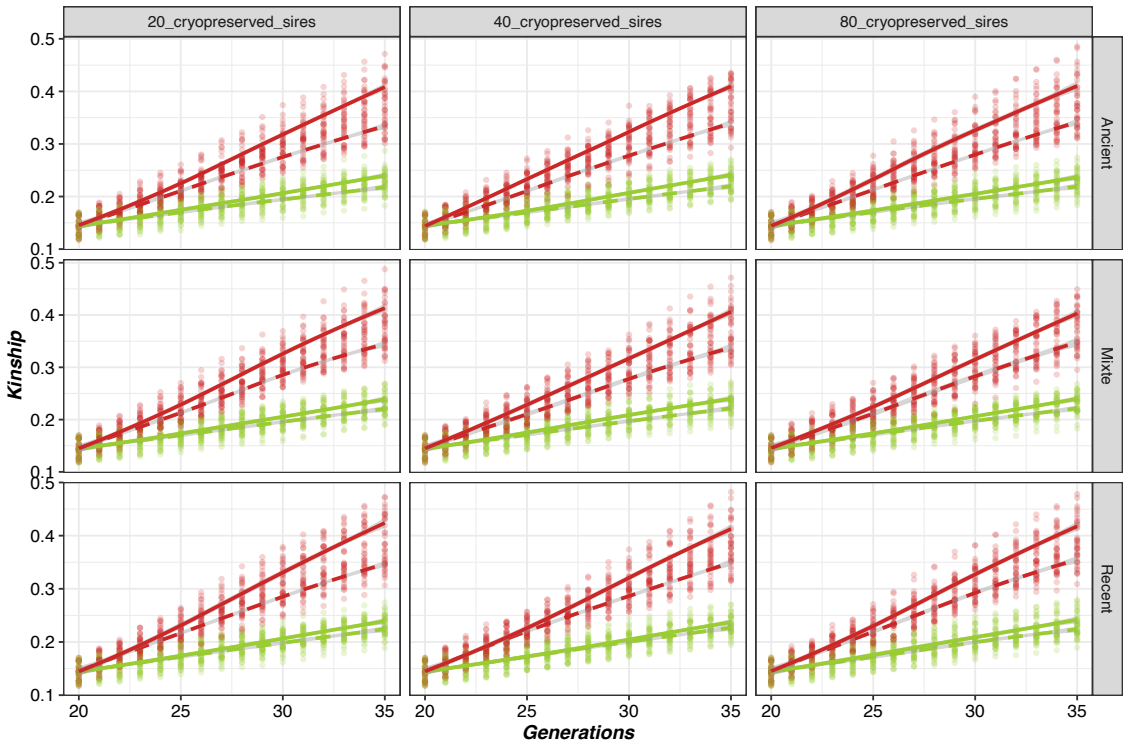


Figure 4 - Age of genetic resources used in a program maximizing progress under diversity constraints (pop_OCS) with mobile cryobanks.

Impact on genetic diversity

In conservation programs focused only with maintaining genetic diversity, there is a significant effect of the size of fixed collections (3-factor ANOVA test, F=659.7, df=346, p-val<0.05 for pop_rm and F=229.5, df=346, p-val<0.05 for pop_max_DG), while the age of collections seems to less affect the level of diversity (see Figure 5). However, when using dynamic collections, it is more advantageous to use older cryopreserved individuals (see Figure 6). The greatest impact on diversity was observed with the cumulative collection (3-factor ANOVA test, F=410.6, df=230, p-val<0.05), which was the most extended in time and had the highest numbers of cryopreserved individuals .



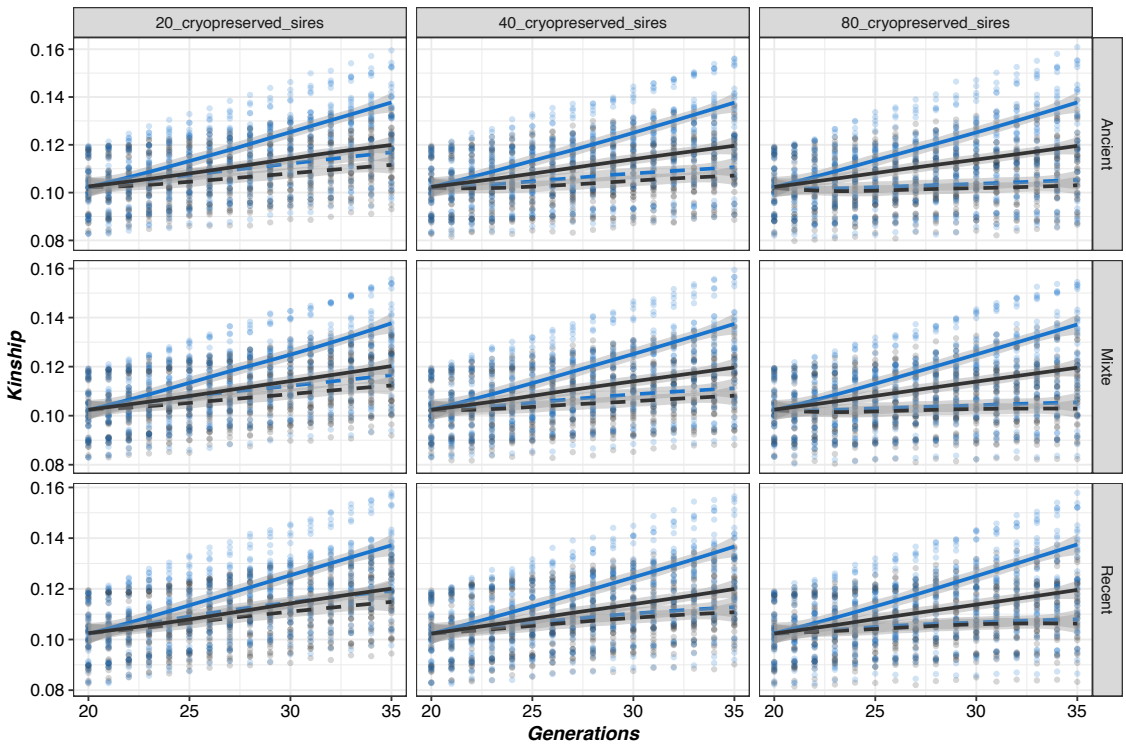


Figure 5 - Evolution of average kinship over the generations according to the different management programs and constitution strategies of fixed cryopreserved collections.

In black the pop_rm program, in blue the pop_max_DG program, in red the pop_max_BV program and in green the pop_OCS program. Solid and dotted lines correspond respectively to scenarios without and with the use of *ex situ* collections.

For selection programs (pop_max_BV and pop_OCS), the age or size of fixed collections doesn't appear to have any effect on the evolution of genetic diversity within populations. However, when using dynamic collections, there is an advantage of cumulative strategies and recent collections over those extended over time for the program aimed solely at improving performance, pop_max_BV (see Figure 6). In the latter case, collections are used very little, which may explain this lower effect. In breeding programs involving genetic diversity (*i.e.* pop_OCS), the impact of *ex situ* collections on genetic diversity is positive overall, although low, and doesn't differ between the different strategies (t-ratio=0.58, p-val=0.57).

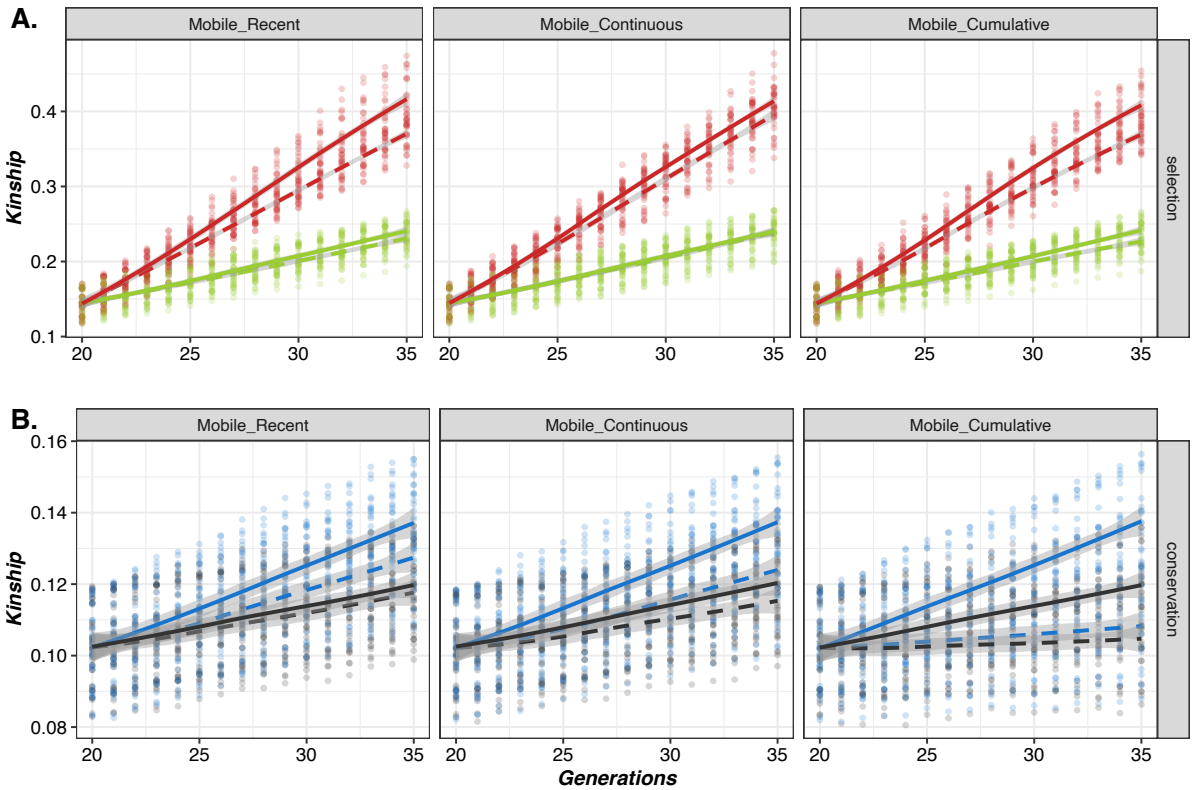


Figure 6 - Evolution of average kinship over generations according to the different management programs and constitution strategies of mobile cryopreserved collections.

In black the pop_rm program, in blue the pop_max_DG program, in red the pop_max_BV program and in green the pop_OCS program. Solid and dotted lines correspond respectively to scenarios without and with the use of *ex situ* collections.

Finally, for all the management programs involving a constraint on diversity (*i.e.* pop_rm, pop_max_DG, pop_OCS), genetic drift is globally more strongly reduced with the most complete collections (*i.e.* cumulative) than with recent collections (2-factor ANOVA test, F=245.7, df=114, p-val<0.05 for pop_rm, F=982.1, df=114, p-val<0.05 for pop_max_DG and F=22.84, df=114, p-val<0.05 for pop_OCS), with lower Nei genetic distances between generations 20 and 35 (see Table 2). On the other hand, for the program aimed solely at improving performance (*i.e.* pop_max_BV), there were no differences between the different strategies to build collections.

Table 2 - Genetic distances (Nei distance) between generations 20 and 35 according to the management program and the strategy used to constitute the cryobank.

Cryobank types	pop_rm	pop_max_DG	pop_OCS	pop_max_BV
Mobile recent	4.02E-03 [sd=2.50E-04]	5.52E-03 [sd=3.84E-04]	2.00E-02 [sd=1.87E-03]	5.86E-02 [sd=7.83E-03]
Mobile continuous	3.87E-03 [sd=2.54E-04]	4.95E-03 [sd=1.69E-04]	1.72E-02 [sd=1.46E-03]	5.23E-02 [sd=4.38E-03]
Mobile cumulative	2.23E-03 [sd=1.74E-04]	2.14E-03 [sd=1.29E-04]	1.56E-02 [sd=1.09E-03]	5.23E-02 [sd=4.76E-03]

Note: Scenario abbreviations correspond to those developed in the Method section.

Impact on genetic progress

The synthetic index is only slightly affected by the strategies used to build fixed cryopreserved collections, whatever their size or age (Figure 7) (3-factor ANOVA test, F=7.86, df=346, p-val<0.05 for pop_max_BV and F=1.26, df=346, p-val=0.24 for pop_OCS).

However, the impact on the two traits that make up this index differs (see Figure 8 and Figure 9). The use of cryopreserved individuals led to a slowdown in genetic progress for Trait 1 (3-factor ANOVA test, F=39.65, df=230, p-val<0.05) and improved Trait 2 (3-factor ANOVA test, F=41.63, df=230, p-val<0.05), minority trait in the index and negatively correlated with the majority trait. Thus, the larger the collection and the more ancient individuals it contains, the greater this impact will be.

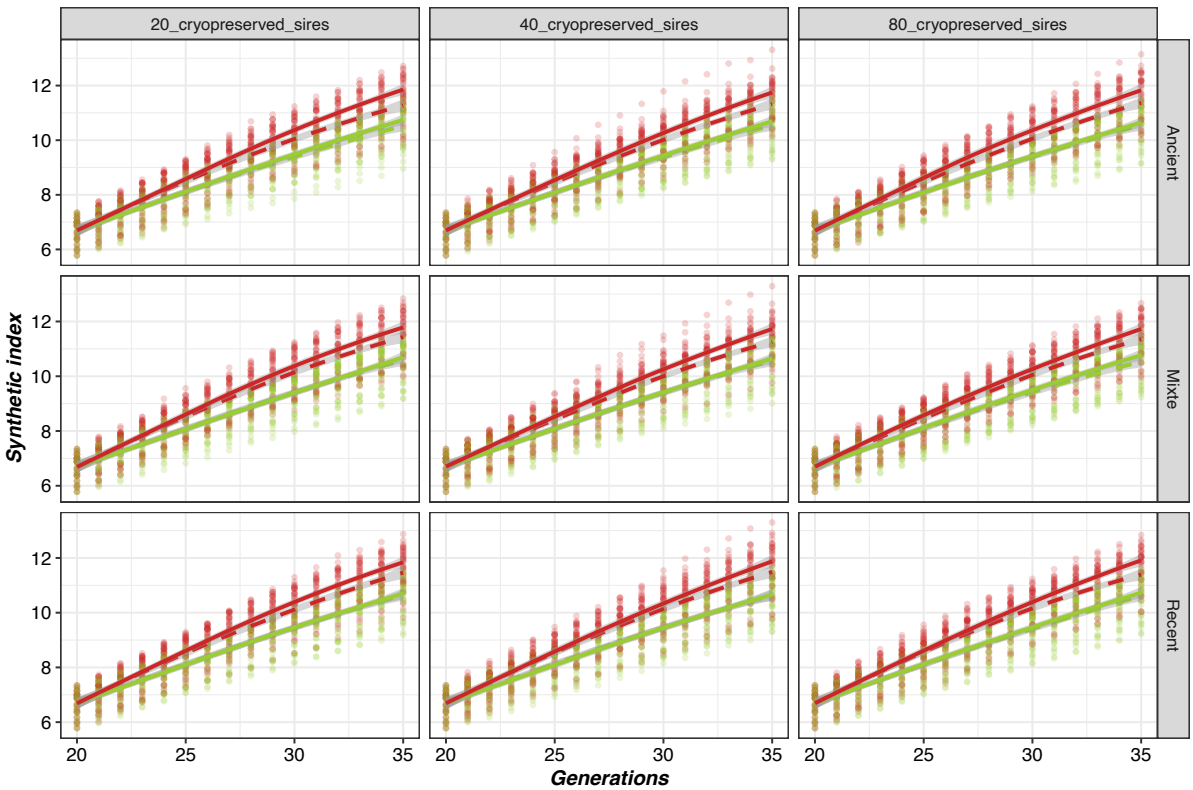


Figure 7 - Evolution of the synthetic index value over the generations for the pop_max_BV (in red) and pop_OCS (in green) programs according to the different strategies for building up fixed cryopreserved collections.

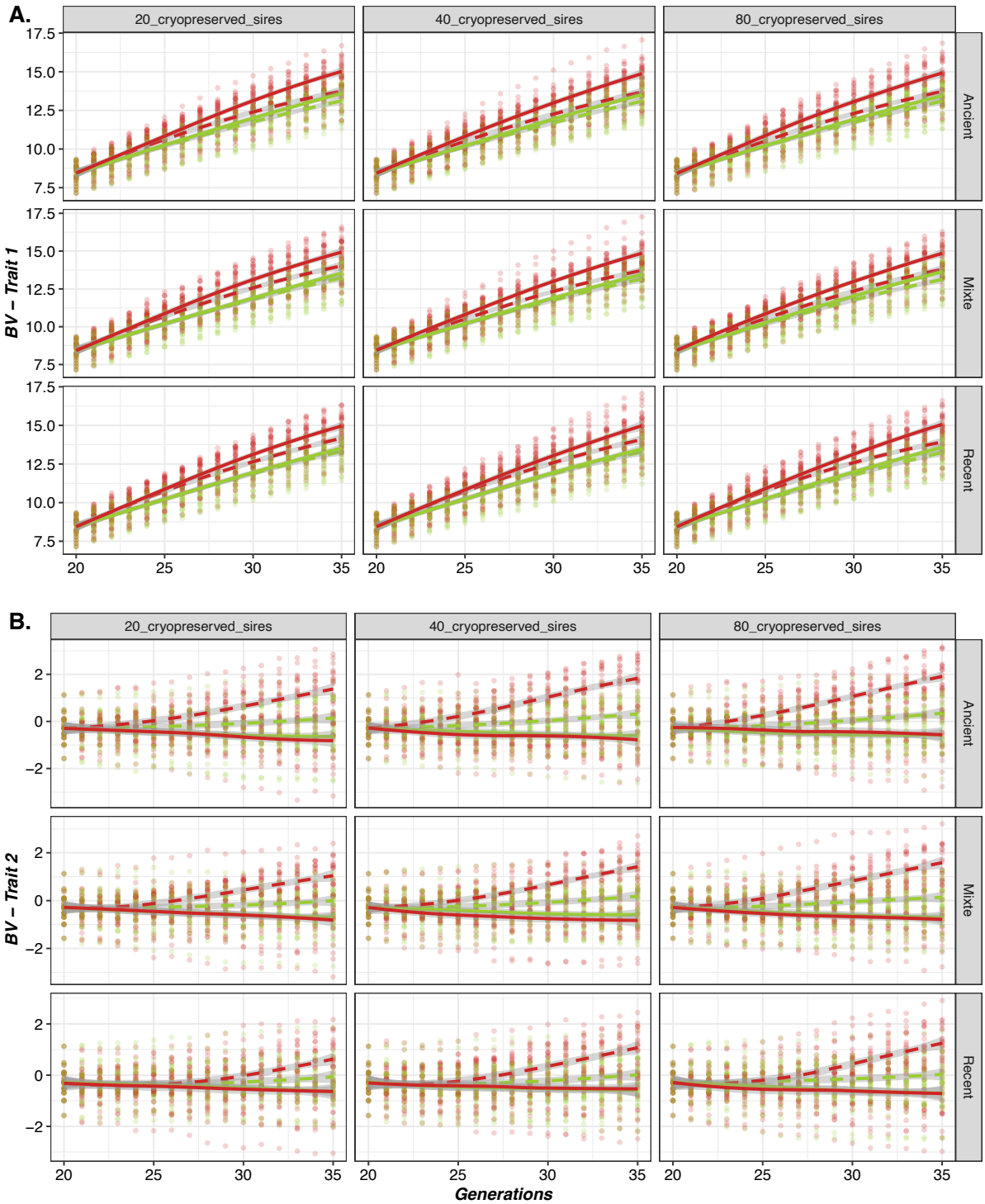


Figure 8 - Evolution of the genetic value of the two traits over the generations for the pop_max_BV (in red) and pop_OCS (in green) programs, according to different strategies for building fixed cryopreserved collections.

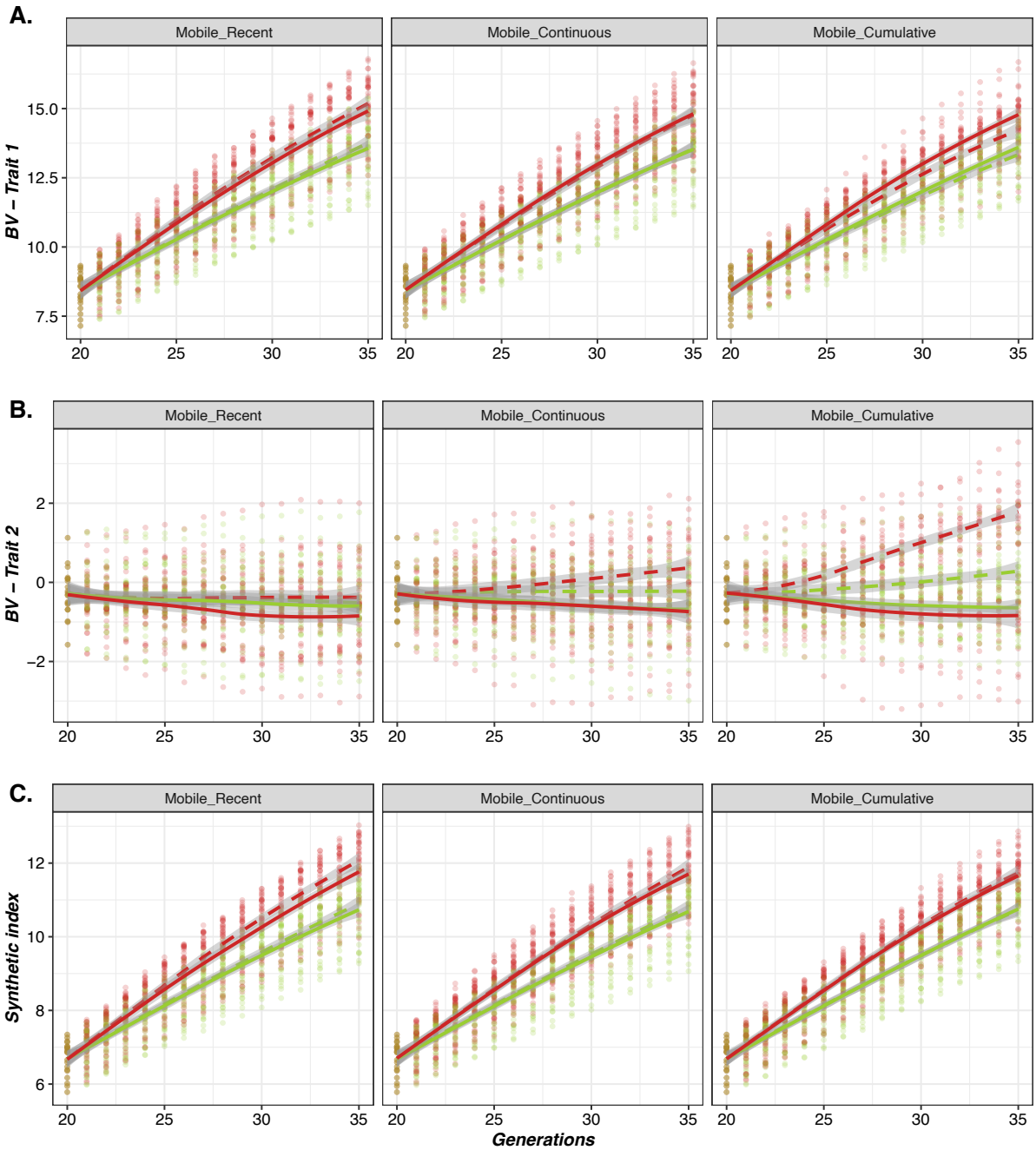


Figure 9 - Evolution of the genetic value of the two traits over the generations for the pop_max_BV (in red) and pop_OCS (in green) programs, according to the different strategies for building dynamic cryopreserved collections.

Impact on genetic variance

For fixed cryobanks, there were no significant differences in the delta of genetic variances of Trait 1 and Trait 2 between generation 20 and 35 in scenarios using OCS (3-factor ANOVA test, F=1.09, df=346, p-val=0.37 for Trait 1 and F=1.19, df=346, p-val=0.19 for Trait 2). Details of the values for the various metrics are given in Additional file 2.

For mobile collections, the additive genetic variances of Trait1 and Trait2 were not significantly different according to the use of cryopreserved individuals for the three types of germplasm mobile collections in the pop_OCS scenarios (2-factor ANOVA test, F=0.42, df=114, p-val=0.83 for Trait 1 and F=1.29, df=114, p-val=0.28 for Trait2) and in the pop_max_BV scenarios (2-factor ANOVA test, F=1.98, df=114, p-val=0.09 for Trait 1 and F=0.18, df=114, p-val=0.97 for Trait2) (see Figure 10).

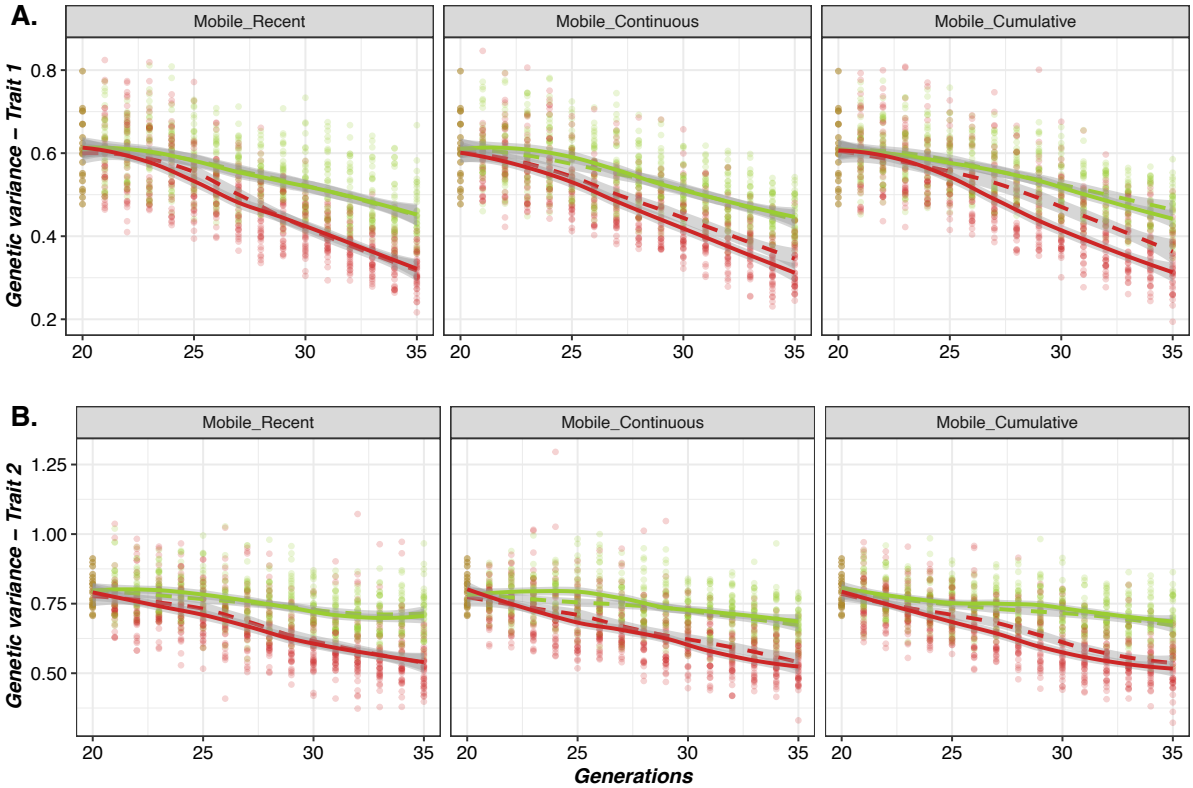


Figure 10 - Evolution of genetic variances for the Trait 1 (A) and Trait 2 (B) according to the type of mobile cryobank for selected population.

In green, scenario maximizing genetic gain with a constraint on kinship increase (pop_OCS). In red, scenario maximizing genetic gain (pop_max_BV). The dotted lines represent scenarios with the use of cryopreserved genetic resources, and the solid lines represent scenarios without the use of cryopreserved individuals.

Discussion

Constitution of collections for a breed under selection

For populations under selection, the age of cryopreserved individuals seems to be more important than the size of *ex situ* collections. Indeed, the genetic values of cryobanks containing old individuals become too far removed from those of contemporary candidates and greatly penalize genetic progress, whereas for *ex situ* collections containing recent individuals, this performance gap is more moderate. In addition, when improving genetic progress is the main objective, a recent cryobank with many individuals becomes more advantageous, as it maximizes the chances of having cryopreserved individuals with interesting genetic values. In this situation, cryopreserved genetic resources are not relevant for immediate use, but remain very interesting for other purposes (*e.g.* reorientation of the breeding scheme discussed in the first part of this study). Moreover, in certain cases, there may be a desire to reintroduce certain traits which have been little taken into account, or totally ignored, in the selection objective. This can be an important motivation for the use of cryopreserved animal genetic resources. One example is the reintroduction of threatened or already extinct families, even if the genetic values are not as high as in the contemporary population. In these situations, the use of older individuals is often forced, as was the case for a local French breed (Jacques et al., 2023). Consequently, for selected populations, older genetic resources are not necessarily relevant, as Leroy et al. (2011) point out, unless there is an additional objective associated with the history and structure of the population.

Constitution of collections for a breed in conservation

The use of cryopreserved resources is a real advantage for conservation programs, whatever the composition of the cryobank. The use of old cryopreserved individuals not only allows diversity to be reintroduced, but also ensures that not too much genetic progress is lost. Indeed, as mentioned above, the main objective for populations in conservation is to maintain a correct level of genetic variability, and genetic gain is often not a priority or even ignored. This differs greatly from selected populations, and leads to a very interesting added value of old *ex situ* genetic resources, since the problem of lagging performance no longer arises. The regular use of cryopreserved genetic resources slows down the increase in inbreeding and the effect of genetic drift, all the more so when cryopreserved collections contain a large number of individuals, as this allows better mixing of matings at each generation, drastically increasing the number of available breeding stock. The use of a cryobank containing few individuals generates a more moderate slowdown in the increase of inbreeding. The use of ex-situ genetic resources is therefore relevant to the conservation of endangered breeds, but the optimal collection-building strategy depends mainly on the resources available for conservation. A cryobank with a large number of aged individuals enables better management of matings and avoids too rapid an increase in inbreeding over the long term, while the use of a more moderate number of aged individuals offers fewer possibilities for managing genetic diversity.

The importance of economic and technical factors in the constitution of cryopreserved collections

For all scenarios, a few cryopreserved individuals were used in each generation, demonstrating that effective long-term management is only possible with the rational use of cryopreserved genetic resources over several generations. However, these strategies for mobilizing *ex situ* collections are costly and also require a high level of operational organization for population managers and breeders, particularly for certain local species and breeds where techniques for using frozen semen are often less widespread than for cattle breeds. Consequently, a compromise might be to use one or two individuals from the cryobank with contemporary breeders, then change donors with each generation. This system would be similar to the rotational mating strategy (Windig and Kaal, 2008; Theodorou and Couvet, 2010). This possibility could be simulated in order to study its impact, in particular with constraints in terms of the quantity of material (*e.g.* AI doses) to be used for each breeder, so as not to deplete stocks too quickly. Indeed, collections of certain breeds may have been difficult to build up. In the case of the French populations, some breeds have benefited from financial support from the CRB-Anim project to build up or enrich their collections, which shows that stocking strategies must also take into account the heterogeneous financial resources of different species and breeds. In this context, it may be appropriate to rationalize collections and consider their dynamic management.

The dynamic collections

The cumulative cryobank, which consists in continuously enriching collections, seems to be the best strategy as it combines large size (*i.e.* genetic variance), originality via older individuals (*e.g.* beneficial for new or minority traits) and good connection with current populations via recent individuals (*i.e.* limited performance gap). It is suitable for all types of population management programs, whether aimed at conservation or selection, with or without changes in objectives. This is currently the case for the French National Cryobank, which is therefore of undoubted interest for the future of animal populations, particularly in the context of agro-ecological transition. However, this option is the most costly to set up and is unlikely to be sustainable. It is therefore necessary to consider more parsimonious scenarios, while maximizing their usefulness in the management of animal populations.

For selected populations, where the main objective is genetic progress, cryobanks are rarely used. It is mainly when selection objectives change that cryopreserved collections become most useful. In this case, preserving reproductive material from the last few generations prior to the current generation would be the most appropriate storage strategy (*i.e.* little gap in genetic values). This could avoid too rapid an increase in inbreeding over time, while retaining potential for new traits introduced for the new purpose. This management strategy is already known for private collections in cattle breeding companies, where old stocks are regularly replaced by contemporary bull semen. This "updating" of stocks has, moreover, enabled the enrichment of the French National Cryobank, which then recovered these old semen doses from breeding companies. This shows that certain storage strategies could be complementary between public and private players. We could envisage extensive dynamic collections, but with different numbers

of individuals depending on their age. In this way, a dynamic collection that samples a wide range of generations upstream, but with more recent individuals and fewer older ones, could be an optimal solution and should be tested at a later date.

Finally, it would be interesting to use simulations to explore more heterogeneous situations in terms of the type of individuals cryopreserved. Indeed, the *ex situ* genetic resources in our simulations were a sample of breeding males from some generations, meaning that there was a good chance that they were connected to the contemporary population. This situation corresponds to type III as defined by the French National Cryobank (Danchin Burge et al., 2006), where the aim is to have a representation of the breed's diversity over a given period. Type II, created for entries of original individuals (*e.g.* family origins, particular genotypes, breeding values etc.) has not been studied, but could be of interest from the point of view of reorienting selection objectives. For example, it would be possible to create a sampling of individuals to be included in the collection by selecting a few individuals whose genetic values show a large deviation from the generation average, or by taking into account new original traits.

Conclusions

In conclusion, strategies for the collections management of *ex situ* animal genetic resources should be in harmony with the potential mobilization according to the breeding program. For populations in selection, based only on the criterion of managing genetic progress and genetic diversity, the sampling of very old semens does not appear to be relevant, given that there is very little chance of these resources being used. For populations in conservation, the number of donors in collections greatly impcat the potential for reintroducing genetic diversity. The regular conservation of certain sires at each generation would enable a wider range of uses, whatever the type of breeding scheme, but would raise logistical issues hat we would need to be able to address to guarantee the long-term supply of the germplasm collections.

Acknowledgements

We are grateful to the INRAE MIGALE bioinformatics facility (MIGALE, INRAE, 2020. Migale bioinformatics Facility, doi: 10.15454/1.5572390655343293E12) for providing help and computing resources. The authors would like to acknowledge Torsten Pook for his assistance with the MoBPS software.

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Funding

This study was partially funded by INRAE institute and the project GenResBridge. GenResBridge has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 817580. The PhD of AJ was funded by IDELE (Institut de l’Élevage), IFIP (Institut français du porc), SCC (Société Centrale Canine) and GenResBridge.

Authors' contribution

GR et MTB conceived the project. AJ performed the analysis and drafted the manuscript. GR and MTB supervised the analyses. AJ, MTB and GR interpreted the results. AJ, GR, and MTB contributed to the writing of the manuscript. All authors read and approved the final manuscript.

References

Blackburn, H. D. 2018. Biobanking Genetic Material for Agricultural Animal Species. Annual Review of Animal Biosciences. 6:69–82. doi:10.1146/annurev-animal-030117-014603.

de Cara, M. Á. R., B. Villanueva, M. Á. Toro, and J. Fernández. 2013. Using genomic tools to maintain diversity and fitness in conservation programmes. Mol Ecol. 22:6091–6099. doi:10.1111/mec.12560.

Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. Nat Rev Genet. 10:783–796. doi:10.1038/nrg2664.

Convention on Biological Diversity (CBD). 1992. Convention on Biological Diversity, Rio de Janeiro.

Danchin Burge, C., E. Verrier, S. Moureaux, M. Tixier-Boichard, and B. Bibe. 2006. Sampling strategies and overview of the French national cryobank collections. In: 8th World Congress on Genetics Applied to Livestock Production. Belo Horizonte, MG, Brasil.

De Oliveira Silva, R., B. V. Ahmadi, S. J. Hiemstra, and D. Moran. 2019. Optimizing ex situ genetic resource collections for European livestock conservation. J Anim Breed Genet. 136:63–73. doi:10.1111/jbg.12368.

Dechow, C. D., W. S. Liu, L. W. Specht, and H. Blackburn. 2020. Reconstitution and modernization of lost Holstein male lineages using samples from a gene bank. J Dairy Sci. 103:4510–4516. doi:10.3168/jds.2019-17753.

Doekes, H. P., R. F. Veerkamp, P. Bijma, S. J. Hiemstra, and J. Windig. 2018. Value of the Dutch Holstein Friesian germplasm collection to increase genetic variability and improve genetic merit. J Dairy Sci. 101:10022–10033. doi:10.3168/jds.2018-15217.

Eynard, S. E., J. J. Windig, I. Hulsegge, S. J. Hiemstra, and M. P. L. Calus. 2018. The impact of using old germplasm on genetic merit and diversity - a cattle breed case study. J Anim Breed Genet. 135:311–22.

FAO. 1998. Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans - Management of small populations at risk. FAO, Rome, Italy. Available from: https://www.fao.org/documents/card/fr/c/d074fc39-aca0-58ab-8024-beaf80ba6014/

FAO. 2007. Global plan of action for animal genetic resources and the Interlaken declaration. FAO, Rome.

FAO. 2012. Cryoconservation of animal genetic resources. Food and Agriculture Organization of the United Nations, Rome.

FAO. 2015. The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture. FAO Commission on Genetic Ressources for Food and Agriculture Assesments. edited by B. D. Scherf and D. Piling, Rome, Italy.

Fox, J., S. Weisberg, B. Price, D. Adler, D. Bates, G. Baud-Bovy, B. Bolker, S. Ellison, D. Firth, M. Friendly, G. Gorjanc, S. Graves, R. Heiberger, P. Krivitsky, R. Laboissiere, M. Maechler, G. Monette, D. Murdoch, H. Nilsson, D. Ogle, B. Ripley, T. Short, W. Venables, S. Walker, D. Winsemius, A. Zeileis, and R-Core. 2019. car: Companion to Applied Regression. https://cran.r-project.org/web/packages/car/index.html. Accessed 24 Oct 2022.

Frankham, R. 2005. Genetics and extinction. Biological Conservation. 126:131–140. doi:10.1016/j.biocon.2005.05.002.

Jacques, A., D. Duclos, C. Danchin-Burge, M.-J. Mercat, M. Tixier-Boichard, and G. Restoux. 2024a. Assessing the potential of germplasm collections for the management of genetic diversity: the case of the French National Cryobank. Peer Community Journal. 4:e13. doi:10.24072/pcjournal.369.

Jacques, A., G. Leroy, X. Rognon, E. Verrier, M. Tixier-Boichard, and G. Restoux. 2023. Reintroducing genetic diversity in populations from cryopreserved material: the case of Abondance, a French local dairy cattle breed. Genet Sel Evol. 55:28. doi:10.1186/s12711-023-00801-6.

Jacques, A., M. Tixier-Boichard, and G. Restoux. 2024b. Optimizing the use of cryopreserved genetic resources for the selection and conservation of animal populations. doi:10.1101/2024.09.17.613440. Available from: http://biorxiv.org/lookup/doi/10.1101/2024.09.17.613440

Kardos, M., E. E. Armstrong, S. W. Fitzpatrick, S. Hauser, P. W. Hedrick, J. M. Miller, D. A. Tallmon, and W. C. Funk. 2021. The crucial role of genome-wide genetic variation in conservation. Proceedings of the National Academy of Sciences. 118:e2104642118. doi:10.1073/pnas.2104642118.

Kinghorn, B. P., R. Banks, C. Gondro, V. D. Kremer, S. A. Meszaros, S. Newman, R. K. Shepherd, R. D. Vagg, and J. H. J. van der Werf. 2009. Strategies to Exploit Genetic Variation While Maintaining Diversity. In: J. van der Werf, H.-U. Graser, R. Frankham, and C. Gondro, editors. Adaptation and Fitness in Animal Populations: Evolutionary and Breeding Perspectives on Genetic Resource Management. Springer Netherlands, Dordrecht. p. 191–200. Available from: https://doi.org/10.1007/978-1-4020-9005-9_13

Lenth, R. V., P. Buerkner, M. Herve, M. Jung, J. Love, F. Miguez, H. Riebl, and H. Singmann. 2021. emmeans: Estimated Marginal Means, aka least-squares means. https://CRAN.R-project.org/package=emmeans/. Accessed 24 Oct 2022.

Leroy, G. 2014. Inbreeding depression in livestock species: review and meta-analysis. Animal Genetics. 45:618–628. doi:10.1111/age.12178.

Leroy, G., P. Boettcher, B. Besbes, C. Danchin-Burge, R. Baumung, and S. J. Hiemstra. 2019. Cryoconservation of Animal Genetic Resources in Europe and Two African Countries: A Gap Analysis. Diversity. 11:240. doi:10.3390/d11120240.

Leroy, G., C. Danchin-Burge, and E. Verrier. 2011. Impact of the use of cryobank samples in a selected cattle breed: a simulation study. Genet Sel Evol. 43:36. doi:10.1186/1297-9686-43-36.

Mercat, M.-J., H. Lenoir, S. Ferchaud, and P. Guillouet. 2008. Première expérience d'utilisation de semence de la Cryobanque Nationale comme outil de gestion de la variabilité génétique en race locale porcine. Journées Recherche Porcine. 93–98.

Meuwissen, T. H. 1997. Maximizing the response of selection with a predefined rate of inbreeding. J Anim Sci. 75:934–40. doi:10.2527/1997.754934x.

Meuwissen, T., B. Hayes, and M. Goddard. 2013. Accelerating improvement of livestock with genomic selection. Annu Rev Anim Biosci. 1:221–237. doi:10.1146/annurev-animal-031412-103705.

Nei, M. 1972. Genetic Distance between Populations. The American Naturalist. 106:283–292. doi:10.1086/282771.

Notter, D. R. 1999. The importance of genetic diversity in livestock populations of the future. J Anim Sci. 77:61–69. doi:10.2527/1999.77161x.

O'Grady, J. J., B. W. Brook, D. H. Reed, J. D. Ballou, D. W. Tonkyn, and R. Frankham. 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. Biological Conservation. 133:42–51. doi:10.1016/j.biocon.2006.05.016.

Pizzi, F., F. Turri, T. M. Gliozzi, and G. Gandini. 2016. Implementation and cost analysis of a regional farm animal cryobank: an Italian case study. Italian Journal of Animal Science. 15:207–210. doi:10.1080/1828051X.2016.1153406.

Pook, T. 2021. MoBPS: Modular Breeding Program Simulator. Available from: https://cran.r-project.org/web/packages/MoBPS/index.html

Pook, T., M. Schlather, and H. Simianer. 2020. MoBPS - Modular Breeding Program Simulator. G3: Genes, Genomes, Genetics. 10:1915–1918. doi:10.1534/g3.120.401193.

R Core team. 2020. R: a language and environment for statistical computing.

Reed, D. H., and R. Frankham. 2003. Correlation between Fitness and Genetic Diversity. Conservation Biology. 17:230–237.

Theodorou, K., and D. Couvet. 2010. Genetic management of captive populations: the advantages of circular mating. Conserv Genet. 11:2289–2297. doi:10.1007/s10592-010-0114-8.

VanRaden, P. M. 2008. Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science. 91:4414–4423. doi:10.3168/jds.2007-0980.

Vavilov, N. 1931. The Problem of the Origin of the World's Agriculture in the Light of the Latest Investigations. In: by the delegates of the U.S.S.R, Frank Cass and Co., London. Available from: https://www.marxists.org/subject/science/essays/vavilov.htm

Wellmann, R. 2019. Optimum contribution selection for animal breeding and conservation: the R package optiSel. BMC Bioinformatics. 20:25. doi:10.1186/s12859-018-2450-5.

Wellmann, R. 2021. optiSel: Optimum Contribution Selection and Population Genetics. Available from: https://CRAN.R-project.org/package=optiSel

Wickham, H. 2011. ggplot2. Wiley Interdiscip Rev Comput Stat. 3:180–185. doi:10.1002/wics.147.

Wiggans, G. R., J. B. Cole, S. M. Hubbard, and T. S. Sonstegard. 2017. Genomic Selection in Dairy Cattle: The USDA Experience. Annu. Rev. Anim. Biosci. 5:309–327. doi:10.1146/annurev-animal-021815-111422.

Windig, J. J., and L. Kaal. 2008. An effective rotational mating scheme for inbreeding reduction in captive populations illustrated by the rare sheep breed Kempisch Heideschaap. Animal. 2:1733–1741. doi:10.1017/S1751731108003029.

Woolliams, J. A., and J. K. Oldenbroek. 2017. Genetic diversity issues in animal populations in the genomic era. In: Genomic management of animal genetic diversity. Wageningen Academic Publishers, The Netherlands. p. 13–47.