

# IDI: an index to assess the usefulness of cryopreserved collections for the management of animal genetic diversity

A. Jacques<sup>1\*</sup>, D. Duclos<sup>2</sup>, C. Danchin-Burge<sup>2</sup>, M.J. Mercat<sup>3</sup>, M. Tixier-Boichard<sup>1</sup> and G. Restoux<sup>1</sup>

<sup>1</sup> Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France  
<sup>\*</sup> [alicia.jacques@inrae.fr](mailto:alicia.jacques@inrae.fr); <sup>2</sup> Institut de l'Élevage, 149 rue de Bercy, 75595 Paris CEDEX 12, France; <sup>4</sup> Institut du porc, La Motte au Vicomte, 35651 Le Rheu CEDEX, France.

## Abstract

Confronted with the erosion of genetic diversity in domestic animal populations, many countries such as France have created gene banks to limit this loss. However, it is crucial to evaluate the potential of these collections before using them for genetic diversity management. Thus, we propose a new Index of Diversity Impact, IDI, to evaluate and compare the diversity potential of collections. IDI is negative when the use of collections decreases the overall relatedness of the current population and positive when it increases it. IDI was used to evaluate the cryopreserved collections of the French National Cryobank of 14 breeds of three livestock species, pig, sheep and cattle. We obtained significantly negative IDIs for three breeds that differ in terms of collection strategy and population management. The IDI appears to be a relevant tool to help managers and breeders to quickly identify promising collections to be used in future breeding programmes.

## Introduction

Genetic diversity is essential for the sustainability of livestock populations by allowing for continuous response to selection and for adaptation to new conditions. However, selection has led to an erosion of genetic diversity over time, which can limit further genetic gain. In addition, for small populations, such as local breeds, the effect of genetic drift is even more important and leads to accelerated depletion of genetic diversity within the breed.

To address this risk, many countries have invested in conserving genetic diversity through the creation of gene banks. These *ex situ* conservation entities have been storing genetic material for several years, including reproductive material such as semen. However, these genetic resources are still underused, even though they could have a potential interest for managing breed diversity. In their study, Leroy et al. (2011) showed that the use of frozen semen from cryobanked bulls can be a valid choice for introducing genetic diversity, provided that attention is paid to gene diffusion, or when there is a change in breeding goals. Two other studies showed that the use of cryopreserved semen from ancient bulls could introduce genetic diversity into a local breed under selection (Eynard et al., 2018) but also in an international breed (Doekes et al., 2018).

France is a country with a rich livestock biodiversity, with for instance 47, 57 and 12 French breeds officially recognized in 2015 for cattle, sheep and pigs, respectively, but a majority of local ones are threatened (Verrier et al., 2015). French breeds greatly differ in terms of population size and management programme, with, for example, the Prim'Holstein, first French dairy breed under selection, having nearly 3,400,000 females born in the period 2016-2019 versus the local Tarentaise dairy breed, having only 17,266 females in the same period (Danchin-Burge, 2020). This raises the question of genetic erosion and motivated the creation of the French National Cryobank, FNC, in 1999, in order to preserve the reproductive material from French breeds of domestic animals. Currently, the FNC brings together reproductive resources of 21 species and nearly 8,000 donors with stored material over the last 20 years.

However, until today, the use of these available resources to manage the genetic diversity of current populations is very scarce. Thus, this study aims at characterizing the collections of the FNC and evaluating their potential use for the management of breed diversity in future breeding programmes.

## Materials & Methods

### *Data for individuals in the National Cryobank.*

Donor animal data were obtained from the FNC database ([www.cryobanque.org](http://www.cryobanque.org)). We focused on three species of livestock that are likely to use frozen semen in their management: cattle, sheep and pig. The national identifier (unique identification in a national pedigree database), the breed and the year of birth of donor animals were recorded.

### *Breeds, pedigree data and definition of cohorts.*

A total of 14 breeds were chosen to cover a large range of situations encountered in France, encompassing two pig breeds, the Pietrain (PIE) and the Cul Noir du Limousin (CNL), five sheep breeds, the Dairy Lacaune (LAC), the Manech Tête Noire (MTN), the Manech Tête Rousse (MTR), the Basco-Béarnaise (BAS) and the Mouton Vendéen (VEN), and seven cattle breeds, the Prim'Holstein (HOL), the Montbéliarde (MON), the Abondance (ABO), the Froment du Léon (FRO), the Tarentaise (TAR), the Saosnoise (SAO) and the Maraîchine (MAR). We defined contemporary male and female cohorts, M and F respectively, as individuals with at least one known parent, born between 2018 and 2020 for pigs, between 2017 and 2020 for sheep, and between 2015 and 2020 for cattle, according to the species' generation interval. These contemporary cohorts are assumed to be currently alive and available for breeding. A cohort composed of all the cryopreserved sires present in the FNC for each breed, C, was also defined. For the pig breed (PIE), 13 boars with semen cryopreserved by the breeding companies were added to our study. The pedigrees of all individuals in the three cohorts were extracted from the national database and limited to the 25 last generations. The pedigree quality of the contemporary male and female cohorts were evaluated with the overall equivalent generations as the weighted average of the generation equivalents (EqG) for each year using PEDIG (Boichard, 2002).

### *Computation of kinship and definition of an Index of Diversity Impact.*

The calculations of the average kinship between the three cohorts were carried out using the *colleau* module of PEDIG (Colleau, 2002). For each breed, we calculated the average kinship between all the sires present in the cryobank (cohort C) and the contemporary female population (cohort F) ( $\overline{\phi_{C-F}}$ ); and between the contemporary male and female cohorts ( $\overline{\phi_{M-F}}$ ). The average kinship within cohort C between cryobank sires ( $\overline{\phi_C}$ ) was also calculated. Then, we defined an Index of Diversity Impact (IDI) for cryobank individuals. This indicator is an estimation of the expected impact of the use of cryobank individuals in random mating with contemporary females on the level of diversity of the next generation under the assumption that only the males in the cryobank would be utilized. It is calculated following the Equation (1).

$$IDI = \frac{(\overline{\phi_{C-F}} - \overline{\phi_{M-F}})}{\overline{\phi_{M-F}}} \quad (1)$$

It is expressed as a percentage of loss (negative) or gain (positive) in mean kinship relative to the current kinship of the population. The confidence intervals (CI) at 95 % were defined from the individual IDI of each cryobank individual and applying Equation (2).

$$CI = \left[ \overline{IDI} \pm 1,96 \frac{sd(IDI)}{\sqrt{n}} \right] \quad (2)$$

With  $sd(IDI)$  the standard deviation observed within the cryobank population of  $n$  donors. All statistical analyses were conducted using R, version 4.0.2 (R Core Team, 2019).

## Results

### *Composition of the cryobank's collections*

For the three species, a total of 84 breeds are represented in the FNC, with 21 breeds for cattle, 15 for pigs and 48 for sheep. This corresponds to 977 bulls, 451 boars and 1,177 rams with frozen semen. The average number of donors per breed is 46.5 for cattle (min=1, max=210, se=13.4), 30.1 for pig (min=10, max=76, se=4.4) and 24.5 for sheep (min=1, max=302, se=6.4). Donor birth years ranged from 1959 to 2013, 1979 to 2018 and 1974 to 2019 for cattle, pig and sheep respectively. This shows that material from ancient sires was retrieved by the FNC after 1999. The range of birth years of donors varied according to species and breed (Table 1).

### *Genetic diversity and IDI for the 14 breeds analysed with pedigree data*

The number of donors in the cryobank for the 14 selected breeds ranged from 10 bulls in the SAO to 302 rams in the LAC. The size of the contemporary cohorts was very variable according to breed, with the number of females ranging from 534 cows for the FRO to 5,069,731 for the HOL.

**Table 1. Table of genetic diversity measures for the 14 breeds studied by pedigree data.**

Breed	Cryobank sires		Contemporary cohort size <sup>1</sup>		No. of EqG <sup>2</sup>		Average Kinship (in %)			Index of Diversity Impact (in %)	
	No.	Birth year range	M	F	M	F	C	C-F	M-F	IDI	[CI] <sup>3</sup>
SAO	10	1991;2011	3084	3089	3.2	3.3	00.8	03.1	01.9	66.5	[25.5;107.5]
FRO	12	1978;2006	536	534	3.1	3.1	10.1	12.8	11.3	13.5	[-08.2;35.2]
TAR	75	1994;2006	25781	26030	6.5	6.6	10.6	09.3	08.6	08.5	[03.1;14.0]
ABO	66	1995;2006	470	71698	6.7	6.8	15.1	12.9	12.0	07.2	[01.1;13.4]
MON	142	1994;2013	3658	1201388	9.4	9.5	12.1	10.8	10.8	00.0	[-03.0;03.0]
MAR	13	1989;2004	3195	3368	6.7	6.5	05.7	07.6	08.0	-04.7	[-28.5;19.2]
HOL	190	1993;2013	10776	5069731	10.4	10.0	10.8	08.9	10.1	-12.6	[-14.7;-10.6]
PIE	27	2016;2017	49073	48503	19.3	19.1	12.8	12.7	12.4	02.3	[00.4;04.1]
CNL	32	1976;2013	1243	1367	15.5	15.5	24.3	24.9	36.1	-31.0	[-35.4;-26.6]
MTN	26	2007;2016	1209	11977	5.9	4.5	11.3	06.0	04.1	44.1	[34.1;54.1]
MTR	48	2008;2016	6546	104184	9.7	8.2	06.8	05.5	05.1	07.2	[-01.3;15.7]
LAC	302	1980;2018	19206	708459	13.3	10.8	05.7	03.5	03.4	03.3	[00.3;06.2]
BAS	32	2009;2016	2418	32547	8.4	6.1	10.1	06.8	06.6	03.2	[-04.9;11.3]
VEN	69	1989;2012	306	33855	13.7	10.3	05.2	03.7	04.4	-16.3	[-20.8;-11.9]

<sup>1</sup> Number of individuals with at less one parent known in the contemporary cohorts

<sup>2</sup> Number of Equivalent generations in the pedigree computed for the contemporary cohorts

<sup>3</sup> Confidence intervals at 95 %, computed from the individual IDI of each cryobank individual

The quality of the pedigrees ranged from 3.1 to 19.1 EqG on the female pathway and from 3.1 to 19.3 EqG on the male pathway (Table 1). The mean within-cryobank kinship values ranged from 0.8 % to 24.3 %. The average kinship values between cryobank sires and contemporary females (cryo-F) ranged from 3.1% to 24.9%, and the average kinship values between contemporary males and females (M-F) ranged from 1.9% to 36.1%. The IDI values varied between -31% for the CNL and 66.5% for the SAO. Three breeds exhibiting different situations have a significantly negative IDI: a local pig breed with quite ancient sires in the cryobank (CNL), a moderately selected sheep breed with a large number of males in the cryobank as compared to contemporary males (VEN), and a highly selected cattle breed (HOL) for which the cryobank favored donors with extreme breeding values and includes old material, thus,

increasing the genetic distance with the current population. The 11 remaining breeds have a positive or null IDI.

## Discussion

With more than 8000 donor animals, the FNC needs to know the potential value of its collections for the management of genetic diversity, all the more that ancient sires are represented in the collections and that living populations can suffer from selection and drift. For sheep, the collections include donors born fairly continuously over time since 1985. For pigs, local breeds have ancient donors whereas breeds under selection have quite recent donors. For cattle, few donors are born after the 2010s following the switch to genomic selection, which considerably shortened the reproductive carrier of males and decreased the quantity of material available for long term storage.

For the breeds studied, the number of donors in the collection is not correlated with the average intra-breed relatedness, which rather depends on the quality of the pedigree of the individuals in collections and the breed management programmes. The IDI reflects the average impact of randomly mating sires in collections with contemporary females relative to expectations for the use of contemporary males only; and makes it possible to rank the usefulness of collections resulting from different sampling strategies. A positive or null IDI reveals a representative collection of the current population, either by the presence of the strongest contributors to the breed (12 AI bulls in collection out of the 13 used in the FRO), or by recent collections likely more related to current females (example of PIE or BAS). However, a positive index does not definitively mean that the use of cryobanked sires is not useful for genetic diversity management, but that on average there will be fewer individuals in the collection that are genetically distant from the current population. In these cases, large confidence intervals suggest that the use of some well-chosen cryobank sires could produce matings with a lower level of relatedness than with contemporary sires, even if the IDI is positive, as for FRO or MON.

In conclusion, the IDI summarizes the impact, on genetic diversity, of using cryopreserved individuals relative to the current level of genetic diversity of the population; it is thus a convenient tool for managers to rapidly screen collections. In our study, IDI values were obtained from pedigree data but could also be calculated from genomic data, especially to be applicable to local breeds without pedigrees. The next step is to analyse options for individual mating plans between cryobank sires and contemporary females.

## References

- Boichard D. (2002) Proc. of the 7<sup>th</sup> WCGALP, Montpellier, France.
- Colleau J.J. (2002) Genet. Sel. Evol. 34:409–421. <https://doi.org/10.1051/gse:2002015>
- Danchin-Burge C. Indicateurs de variabilité génétique – races bovines – Edition 2020. Available at: [https://idele.fr/fileadmin/medias/Documents/2020\\_varume\\_BL.pdf](https://idele.fr/fileadmin/medias/Documents/2020_varume_BL.pdf)
- Doekes H.P., Veerkamp R.F., Bijma P., Hiemstra S.J., and Windig J. (2018) J. Dairy Sci. 101(11):10022–10033. <https://doi.org/10.3168/jds.2018-15217>
- Eynard S.E., Windig J.J., Hulsegge I., Hiemstra S.J., and Calus M.P.L. (2018) J. Anim. Breed. Genet. 135(4):311–322. <https://doi.org/10.1111/jbg.12333>
- Leroy G., Danchin-Burge C., and Verrier E. (2011) Genet. Sel. Evol. 43:36. <https://doi.org/10.1186/1297-9686-43-36>
- R Core Team. R: A language and environment for statistical computing. Available at: <https://www.R-project.org/>
- Verrier E., Audiot A., Bertrand C., Chapuis H., Charvolin E. et al. (2015) Anim. Genet. Resources 57:105–118. <https://doi.org/10.1017/S2078633615000260>