

Conservation genetics

Drift, effective population size and sampling in genebank and conservation network

Identifier	SAMEN-PROV0001-P14
Author	Thomas Schüller, Syntagma Forschung Wien · ORCID 0009-0003-9799-6747
Contact	forschung@tschueller.com
Version and date	2.3 · 2026-09-28
Rights	Text all rights reserved; tools and data layers MIT
Disclaimer	Calculates with idealised models; real effective population sizes are usually well below the number of plants.

Abstract

An accession is not a single plant but a population. What makes it valuable is the diversity within this population, and this diversity is partly lost with every multiplication, even when germination capacity and quantity are in order. This paper sets out the calculations on which the plant numbers of the Collection Manager and the Genetics Calculator are based: drift, effective population size and sampling. The central finding: not storage but regeneration is the genetic bottleneck, and it is decided by the smallest plant number in the series of multiplications, not the average.

1. Why diversity within an accession matters

A landrace is usually a mixture of lines or, in outcrossers, a population with many alleles per locus. Rare alleles often carry precisely the traits for which a variety is later sought, such as a resistance or an adaptation to a site. The passport data (paper P11) record where an accession comes from; whether it retains its composition is decided during multiplication. Paper P05 puts it this way: whoever multiplies too few plants preserves the variety in name and loses it in composition. The following sections substantiate this statement by calculation.

2. Drift and loss of heterozygosity

2.1 The formula

In a small population, chance helps decide which alleles reach the next generation (genetic drift). This reduces heterozygosity, that is the proportion of heterozygous loci, in an ideal population according to the relation

$$H_t = H_0 \cdot (1 - 1/(2 \cdot N_e))^t$$

H_0 is the initial heterozygosity, t the number of generations, N_e the effective population size. Each generation, the proportion $1/(2 \cdot N_e)$ is lost. Harmon and Braude (2010) give the equivalent form of the increase in inbreeding F ; the proportion retained is $1 - F$.

2.2 What the formula means

At $N_e = 50$, 1 per cent of heterozygosity is lost per generation. After 50 generations 60.5 per cent is retained, after 100 generations 36.6 per cent. Heterozygosity is halved after about 69 generations. Figure 13 'Loss of heterozygosity through drift' shows the curves for N_e 10, 50, 100 and 500 over 100 generations: at $N_e = 10$ less than 1 per cent remains after 100 generations, at $N_e = 500$ about 90 per cent. A lost allele does not return from within the population itself.

3. Effective and census population size

3.1 The difference

The effective population size N_e is the size of an ideal population that loses as much diversity through drift as the actual one. It is almost always below the counted number N of plants, because plants leave unequal numbers of offspring and flowering times do not overlap completely. Frankham, Bradshaw and Brook (2014) mostly give values of 0.1 to 0.2 for the ratio N_e/N .

3.2 Unequal sex ratio

In dioecious species such as spinach or hemp, the rarer sex limits the effective size:

$$N_e = 4 \cdot N_m \cdot N_f / (N_m + N_f)$$

With 10 male and 40 female plants, $N_e = 4 \cdot 10 \cdot 40 / 50 = 32$. Of 50 counted plants, only 32 thus act genetically.

3.3 Fluctuating size across generations

Over several generations, N_e is the harmonic mean of the individual sizes:

$$N_e = t / \sum(1/N_i)$$

The harmonic mean is dominated by small values. A series of five multiplications with 100, 100, 12, 100 and 100 plants has an arithmetic mean of 82.4 but an effective size of only 40.5. A single bottleneck of twelve plants thus pushes the effective size down from 100 to about 40, even though four multiplications went properly. One poor growing year or an emergency multiplication from a remnant weighs more heavily than any good one.

4. Breeding system

Gametes per plant. An outcrossing plant contributes about two independent sets of gametes to the sample via egg cell and pollen, a selfer only about one, because both come from the same plant. Selfers therefore need about twice as many plants for the same certainty (section 5).

Isolation. Without isolation, outcrossers cross with other origins. Paper P05 calculates cautiously with about 500 metres for insect-pollinated and about 1000 metres for wind-pollinated species; practical guides give considerably smaller distances for individual species. Isolation protects the identity of the accession, the plant number its diversity.

5. The sample

The probability of capturing an allele of frequency p at least once in k independent gametes is $P = 1 - (1 - p)^k$. Solved for the number of plants n , with $g = 2$ gametes per plant for outcrossers and $g = 1$ for selfers:

$$n = \text{ceil}(\ln(1 - P) / \ln(1 - p) / g)$$

For $p = 0.05$ and $P = 0.95$, 58.4 gametes are needed. This gives 30 plants for outcrossers and 59 for selfers; the data layer and the Collection Manager round up to 60 for selfers. The values correspond to FAO Standard 4.1.5 (2014, already in the 2011 draft), which sets them for collecting; this paper transfers them to regeneration. Conversely, 30 plants of a selfer capture such an allele with a probability of only 78.5 per cent, ten plants with 40.1 per cent.

The rule applies to one allele at one locus. Whoever wishes to secure many loci at once needs more: Crossa et al. (1993) give 160 to 210 plants for alleles with a frequency of at least 0.05 at 150 loci at a probability of 90 to 95 per cent; Crossa (1989) gives about 100 individuals for outcrossers for alleles at 5 per cent and 300 to 400 for alleles at 1 per cent.

In apparent contradiction, the data layer on community conservation lists 10 plants for simple selfers such as tomato or bean, 40 for insect-pollinated outcrossers and 100 for wind-pollinated ones, expressly as empirical values. The contradiction resolves through the subject: a uniform line variety of a selfer is largely homozygous, its diversity lies between varieties, not within the variety; ten plants maintain it. A landrace or mixed stock of a selfer, by contrast, is a mixture of lines, and the sampling rule of about 60 plants applies. Since version 2.1 this note appears in the data layer and in the Conservation Network Planner.

6. Regeneration as bottleneck

6.1 Every regeneration is a generation

No drift takes place in storage. Every regeneration, however, is a generation in the sense of the formula, and it is there that the plant number takes effect. This is why the testing and regeneration intervals from paper P02 have a genetic side: the earlier regeneration is needed, the more generations fall within the same period. Under the FAO standards, regeneration takes place when germination falls below 85 per cent of the initial value or the remaining quantity no longer suffices for three sowings (Standard 4.4.1). Whether this is the case is shown by the germination test; without a prediction, it is repeated at -18°C and hermetic sealing every ten years for long-lived and at most every five years for short-lived species (Standard 4.3.3). A much-requested accession ages genetically faster than one that rests. Good storage saves generations.

6.2 Ten regenerations with 30 plants

Assuming regeneration is needed every ten years, ten regenerations span a century. With 30 plants and $N_e = N$, 84.5 per cent of heterozygosity is retained. At $N_e/N = 0.2$, the upper value of the range given by Frankham et al., N_e is only 6, and 41.9 per cent remains. The difference lies solely in how evenly the plants contribute to the next generation.

6.3 The worked examples at a glance

Calculation	Input	Result
Drift	$N_e = 50$, 50 generations	60.5 % of heterozygosity retained

Calculation	Input	Result
Drift	Ne = 50, 100 generations	36.6 % retained
Sex ratio	10 male, 40 female plants	Ne = 32 instead of 50
Fluctuating size	100, 100, 12, 100, 100 plants	Ne = 40.5; arithmetic mean 82.4
Sample, outcrosser	p = 0.05, P = 95 %	30 plants
Sample, selfer	p = 0.05, P = 95 %	59 plants, rounded 60
Regeneration	10 times 30 plants, Ne = N	84.5 % retained
Regeneration	10 times 30 plants, Ne/N = 0.2	41.9 % retained

All values can be reproduced in the Genetics Calculator.

7. Guide values: 50/500 and 100/1000

Franklin (1980) formulated the rule that an effective size of at least 50 protects against inbreeding depression in the short term and one of at least 500 maintains adaptive potential in the long term. Frankham, Bradshaw and Brook (2014) proposed raising the values to 100 and 1000: Ne = 100 would limit the loss of fitness over five generations to at most 10 per cent, Ne = 1000 would maintain evolutionary potential. The revision is contested; Franklin, Allendorf and Jamieson (2014) hold to the original rule.

A regeneration with 30 or 60 plants, at an Ne/N as in wild populations (0.1 to 0.2), does not even reach the lower threshold of the original rule. With an even harvest per plant, Ne is higher, but even then at most 30 or 60 respectively. The guide values come from the protection of wild populations; in ex situ conservation they are an argument for keeping the number of regenerations small. The FAO standards also recommend keeping, where possible, at least 50 seeds of the original sample as a reference (Standard 4.4.3).

8. Core collections and genetic erosion

Brown (1989) proposed forming, from a large collection, a core collection that captures about 70 per cent of the alleles with about 10 per cent of the accessions, calculated in a model of neutral alleles; the figure is listed as partially verified in the data layer. The core collection serves use, not conservation: the remaining 90 per cent carry precisely the rare alleles.

On the question of whether crop diversity as a whole has declined, van de Wouw et al. (2010), in a meta-analysis of 44 studies, found no long-term decrease in diversity within the bred cultivars of the 20th century, but a decline in the 1960s. Not covered is the replacement of landraces; on this loss see papers P17 and P18.

9. Conclusions for conservation networks

- 1. The smallest multiplication counts.** The harmonic mean forgives no bottleneck. A network should record for each accession the plant number of every multiplication, not only the result in grams. The Collection Manager records the plant number per multiplication and flags every multiplication below the sampling rule as a bottleneck.
- 2. Plant number by breeding system.** To capture an allele at 5 per cent frequency with 95 per cent certainty, 30 plants for outcrossers and 60 for selfers are regarded as the lower limit. If space is lacking, an accession can be distributed among several custodians and the harvest mixed in equal proportions.

3. **Harvest evenly.** Equal quantities per mother plant bring N_e closer to N .
4. **Multiply less often.** Every regeneration avoided saves one generation of drift (P02).
5. **Keep the original.** The reference sample allows comparison and a fresh start.
6. **On-farm is a different object.** There, selection and adaptation act in addition to drift (paper P05); the formulas of this paper do not represent this.

Figures

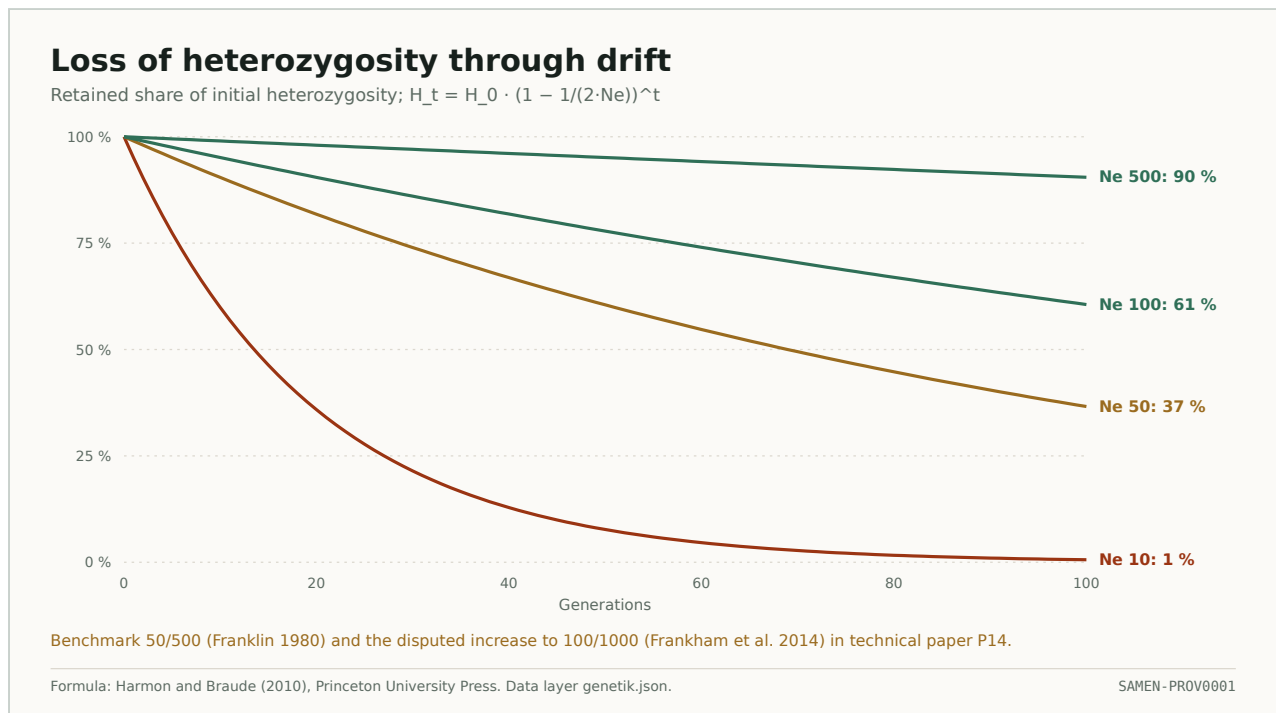


Figure 1. Loss of heterozygosity through drift. Source: data layer of the strand, see figure footer.

Limits of this paper

All calculations assume an ideal population: no selection, no mutation, no immigration of foreign alleles, random mating and non-overlapping generations. In regeneration, by contrast, unintended selection acts, for instance for the multiplication site, and in outcrossers introgression despite isolation. The ratio N_e/N is usually unknown for individual accessions; the range of 0.1 to 0.2 is a guide value, not a measurement. Two or one gamete per plant respectively is a simplification that does not represent partial outcrossing in selfers. The plant numbers of the data layer on community conservation are empirical values; the figure on core collections is only partially verified.

References

1. Harmon, L. J.; Braude, S. (2010): Conservation of small populations: effective population sizes, inbreeding, and the 50/500 rule. Chapter 12. Princeton University Press. https://assets.press.princeton.edu/chapters/s12_9242.pdf
2. Franklin, I. R. (1980): Evolutionary change in small populations. In: Soulé, M. E.; Wilcox, B. A. (eds.): *Conservation Biology. An Evolutionary-Ecological Perspective*. Sinauer, 135–149. <https://oalib.com/references/8896108>
3. Frankham, R.; Bradshaw, C. J. A.; Brook, B. W. (2014): Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation* 170, 56–63. <https://doi.org/10.1016/j.biocon.2013.12.036>

4. Franklin, I. R.; Allendorf, F. W.; Jamieson, I. G. (2014): The 50/500 rule is still valid. Reply to Frankham et al. *Biological Conservation* 176, 284–285. <https://doi.org/10.1016/j.biocon.2014.05.004>
5. Crossa, J. (1989): Methodologies for estimating the sample size required for genetic conservation of outbreeding crops. *Theoretical and Applied Genetics* 77, 153–161. <https://doi.org/10.1007/BF00266180>
6. Crossa, J. et al. (1993): Statistical genetic considerations for maintaining germ plasm collections. *Theoretical and Applied Genetics* 86, 673–678. <https://doi.org/10.1007/BF00222655>
7. Brown, A. H. D. (1989): Core collections: a practical approach to genetic resources management. *Genome* 31, 818–824. <https://doi.org/10.1139/g89-144>
8. van de Wouw, M. et al. (2010): Genetic diversity trends in twentieth century crop cultivars: a meta analysis. *Theoretical and Applied Genetics* 120, 1241–1252. <https://doi.org/10.1007/s00122-009-1252-6>
9. FAO (2011): *Draft Genebank Standards for Plant Genetic Resources for Food and Agriculture*. Draft. Rome: FAO. <https://www.fao.org/fileadmin/templates/agphome/documents/PGR/genebank/Genebankstandards-Nov2011english-2b.pdf>
10. FAO (2014): *Genebank Standards for Plant Genetic Resources for Food and Agriculture*. Revised edition. Rome: FAO. <https://www.fao.org/4/i3704e/i3704e.pdf>
11. Papers P02, P05 and P11 of this strand.