

Short Communication

Reliability of germination testing of *ex situ* conserved seeds: a genebank case study on outsourced analyses

Theo van Hintum* and Rob van Treuren

Centre for Genetic Resources, the Netherlands (CGN), Wageningen University and Research Centre, PO Box 16, NL-6700 AA Wageningen, The Netherlands

Received 9 March 2012; Accepted 13 April 2012 – First published online 11 May 2012

Abstract

Germination tests are performed on a routine basis to determine the viability of genebank accessions. The results determine which accessions have to be rejuvenated. The reliability of the germination test results used by the Centre for Genetic Resources, the Netherlands was determined by the retesting of 641 random samples anonymously, in the same year and by the same testing agency as the original tests. Results showed alarmingly low reliabilities, with error levels much higher than expected based on sampling effects. The result of a germination test of a random sample with a germination of 80% was shown to have a 95% confidence interval from 63 to 97%. The errors differed strongly over crops and testing years, and were larger for crop wild relatives than for crop species.

Keywords: error estimation; genebanks; genetic resources; germination; seed viability

Experimental

As decisions about the rejuvenation of seed lots conserved in *ex situ* collections are primarily based on germination test results, knowledge about the reliability of these results is of crucial importance. Rejuvenation is a costly activity that bears the dangers of genetic drift, shift and contamination, all compromising authenticity (van Hintum *et al.*, 2007; van de Wouw *et al.*, 2011). Therefore, unnecessary rejuvenation based on unreliable germination test results should be avoided, while also samples that do need rejuvenation should be reliably identified.

In publications that analyse seed storage behaviour in genebanks (e.g. Walters *et al.*, 2005; Nagel *et al.*, 2009), the reliability of the germination test results has not been discussed.

Germination tests are used as a proxy of seed viability. The 'Genebank Standards' (FAO/IPGRI, 1994) recommend seed viability testing soon after receipt of a seed lot, followed by retesting at 5 or 10 years' time intervals, depending on the expected storage life or the initial germination. Due to high testing costs and the substantial amount of seeds needed for testing, genebanks generally test at lower frequency. For example, the Centre for Genetic Resources, the Netherlands (CGN) tests at approximately half to one-third of the recommended frequency.

CGN outsources its germination tests to ISTA-certified testing agencies in the Netherlands. They apply ISTA protocols for the genebank material, using the possibilities of these protocols to break dormancy and to extend the observation period. Only in specific cases other protocols are used, such as for potato where the temperature and dormancy-breaking methods are different. A consistent deviation is that usually only 200 instead of 400 seeds of genebank accessions are tested in order to avoid rapid seed depletion. At CGN, since 2001, 5–10% of the annually

*Corresponding author. E-mail: theo.vanhintum@wur.nl

tested seed lots are retested in order to obtain reliability estimates. The 641 accessions in the randomly chosen and anonymized subsets involved all 25 CGN crops, ranging from a single sample of lamb's lettuce to 145 lettuce samples.

Pairwise plotting of the test results revealed large discrepancies (Fig. 1). The probability that pairwise differences were due to sampling effects was tested with the χ^2 distribution, using a two-tailed test. This allowed us to focus on the pairs that differed more than could be explained based on chance alone with a probability of 5%. Since in the margins of the distribution range, small pairwise differences may already be significant, further analysis was limited to those significantly different pairs where at least one value was smaller than 90%. These pairs, designated as 'suspicious', included 116 (18.1%) of the 641 seed lots, whereas based on the confidence level used statistically maximally 5% would be expected if it was only due to chance.

To compare among different crops, population types (cultivated, wild) and test years, a correction was made to avoid confounding underlying factors. The chance P_{CPY} of a crop C with population type P tested in year Y generating a suspicious result was calculated as:

$$P_{CPY} = P_0 + P_C + P_P + P_Y + \varepsilon,$$

where P_0 represents the overall probability on a suspicious result (18.1%), the factors P_C , P_P and P_Y represent

the additional probabilities based on the crop, population type and year, respectively, and ε is a normally distributed error. The factors were estimated by minimizing the variance of ε , i.e. the sum of the squared differences between the expected and observed number of suspicious results at the different estimated factor levels. Subsequently, the corrected chance of suspicious results per crop ($P_0 + P_C$), population type ($P_0 + P_P$) or year ($P_0 + P_Y$) could be calculated (Table 1).

The size of the pairwise differences obviously also depended on the average germination level of the tested pair. In case of very high values, e.g. 97%, the difference cannot be larger than 6% (100 and 94% for the first and second test, respectively). For this reason, the pairwise differences were classified in ten groups of approximately similar size according to average germination result of the pairs, and the standard deviation of the pairwise differences was calculated for each group. The resulting values allowed estimation of the standard deviation of the errors in the germination results by dividing by $\sqrt{2}$ (Supplementary Fig. S1, available online only at <http://journals.cambridge.org>). For a group with average germination between 77.5 and 82.5%, including 49 pairs of results, the standard deviation of the errors in the germination results was 8.0%. Based on the binomial distribution, a sample size of 200 seeds and a germination level of 80%, a value of 2.8% was to be expected if the error would be due to sampling effects only.

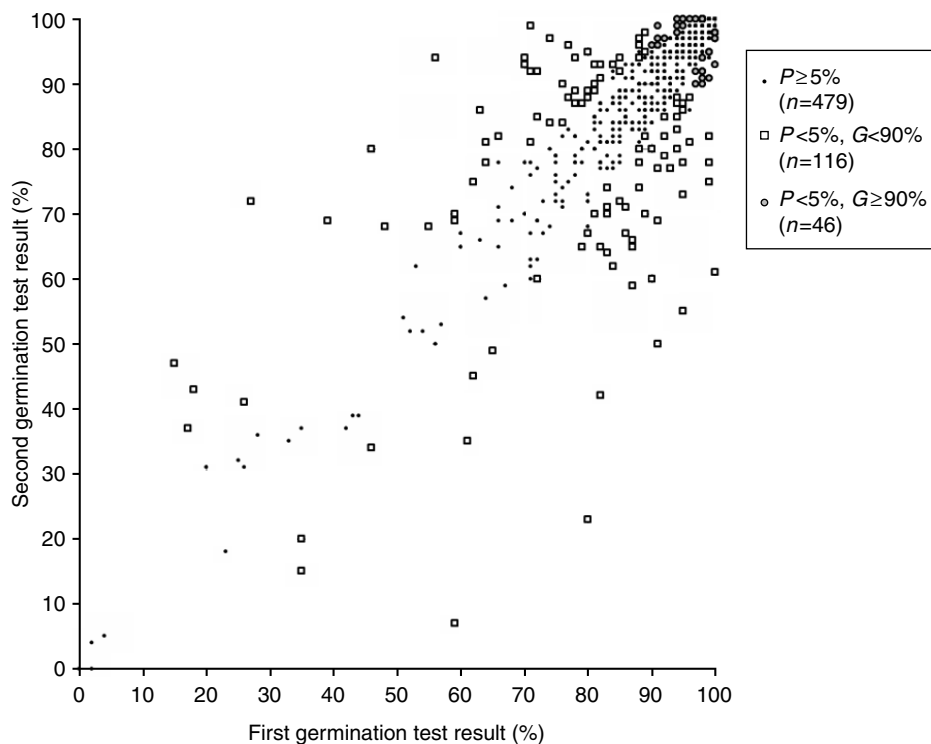


Fig. 1. Germination control tests of 641 seed lots. P indicates the probability of the results by chance alone, while n represents the number of included pairs. For data points with $P < 5\%$, a distinction is made for germination values (G) lesser than 90% and values $\geq 90\%$.

Table 1. Frequency of suspicious results using a 95% confidence interval, presented per crop, population type and testing year^a

	Number of tests	Suspicious cases (%)	Confidence interval (%)
Lettuce	145	7	3–12
Wheat	104	13	6–19
Barley	58	17	7–26
Potato	53	42	29–55
Tomato	51	10	2–18
Pepper	41	29	15–43
Crucifers	28	14	1–26
Cucumber	20	19	2–36
Allium	17	30	8–52
Faba bean	16	36	13–60
Eggplant	16	19	0–38
Spinach	15	32	8–56
Pea	14	22	0–43
Oat	12	5	0–17
Clover	10	18	0–42
Other	41	28	14–42
Cultivated	507	16	1–31
Wild	134	24	9–39
2001	84	18	9–26
2002	95	27	18–36
2003	82	10	3–16
2004	58	20	9–30
2005	42	5	0–12
2006	59	22	11–32
2007	66	14	6–22
2008	55	15	6–25
2009	37	23	9–36
2010	30	23	8–38
2011	33	26	11–41
Total	641	18	15–21

^aData were corrected based on an additive model with crop, population type and year as factors.

Discussion

The results of germination tests are of vital importance for genebank management decisions, and hence should be sufficiently reliable. Sample size is an important factor influencing reliability. For example, a recommended sample of 200 seeds (FAO/IPGRI, 1994) results in a sampling error of 2.8% at a germination level of 80%, while this value is 5.7% when only 50 seeds are used.

For the tests reported here, observed error levels were substantially higher, e.g. 8.0% at 80% germination using an average number of 181.6 seeds. This implies that a sample with a true germination of 80% can show test results with a 95% confidence interval of 62.6 to 97.4%. The causes of these observed errors could be many. Some crops appeared more error prone than others,

crop wild relatives appeared more difficult than crops, and in some testing years the discrepancies were substantially higher than in others (Table 1). Underlying causes may include dormancy, misjudgement of seedling health or unequal composition of different samples of a seed lot. Whether these factors explain the observed differences remains a subject of further study.

Outsourcing is not common practice in most other well-established genebanks, making it possible to optimally use in-house experience. Our results indicate that genebanks in general should be aware of the potential problems due to low reliability of germination data, and therefore should critically examine their test procedures and the implications of the test results for decisions on genebank operations.

Acknowledgements

The authors would like to thank Liesbeth de Groot (CGN) for preparing the data from the germination control tests. The study was part of the Programme for Statutory Research Tasks regarding Genetic Resources (WOT-03-436) and the Fundamental Research Programme on Sustainable Agriculture (KB-12-005.03-003), both funded by the Dutch Ministry of Economic Affairs, Agriculture and Innovation. The authors are grateful to Bert Visser, Chris Kik and Liesbeth de Groot for their comments on an earlier version of the manuscript, and to two anonymous reviewers for their constructive suggestions for further improvements.

References

- Food and Agriculture Organization of the United Nations/International Plant Genetic Resources Institute (FAO/IPGRI), (1994) *Genebank Standards*. Rome: FAO/IPGRI, 17 p.
- Nagel M, Vogel H, Landjeva S, Buck-Sorlin G, Lohwasser U, Scholz U and Börner A (2009) Seed conservation in *ex situ* genebanks — genetic studies on longevity in barley. *Euphytica* 170: 5–14.
- van de Wouw M, van Treuren R and van Hintum TJL (2011) Authenticity of old cultivars in genebank collections: A case study on lettuce. *Crop Science* 51: 736–746.
- van Hintum TJL, van de Wiel CMM, Visser DL, van Treuren R and Vosman B (2007) The distribution of genetic diversity in a *Brassica oleracea* gene bank collection related to the effects on diversity of regeneration, as measured with AFLPs. *Theoretical and Applied Genetics* 114: 777–786.
- Walters C, Wheeler LM and Grotenhuis JM (2005) Longevity of seeds stored in a genebank: Species characteristics. *Seed Science Research* 15: 1–20.