

CGIAR genebank viability data reveal inconsistencies in seed collection management

Fiona R. Hay^{a,b,*}, Katherine J. Whitehouse^c, Richard H. Ellis^b,
N. Ruairaidh Sackville Hamilton^{d,1}, Charlotte Lusty^e, Marie Noelle Ndjondjop^f, Daniel Tia^f,
Peter Wenzl^g, Luis G. Santos^g, Mariana Yazbek^h, Vania C.R. Azevedoⁱ, Ovais H. Peerzadaⁱ,
Michael Abberton^j, Olaniyi Oyatomi^j, Flora de Guzman^d, Grace Capilit^d, Alice Muchugi^{k,2},
Zakayo Kinyanjui^k

^a Department of Agroecology, Aarhus University, Forsøgsvej 1, 4200 Slagelse, Denmark

^b School of Agriculture, Policy and Development, University of Reading, Earley Gate, PO Box 237, Reading, RG6 6EU, UK

^c Australian Grains Genebank, Agriculture Victoria Research Division, Department of Jobs, Precincts and Regions, Private Bag 260, Horsham, Victoria, 3401, Australia

^d T.T. Chang Genetic Resources Center, International Rice Research Institute (IRRI), Los Baños, Philippines

^e Global Crop Diversity Trust, Platz Der Vereinten Nationen 7, 53113, Bonn, Germany

^f Rice Biodiversity Center for Africa, Africa Rice Center (AfricaRice), 01 BP 2551, Bouaké, Cote d'Ivoire

^g International Center for Tropical Agriculture (CIAT), Km 17, Recta Cali–Palмира CP 763537, Apartado Aéreo, 6713, Cali, Colombia

^h International Center for Agricultural Research in the Dry Areas (ICARDA), Dalia Bldg. 2nd Floor, Bashir El Kassar Street, Verdun Area, Beirut, Lebanon

ⁱ International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, 502324, Telangana, India

^j Genetic Resources Center, International Institute of Tropical Agriculture (IITA), PMB 5320, Oyo Road, Ibadan, 200001, Nigeria

^k World Agroforestry, United Nations Avenue, Gigiri PO Box 30677, Nairobi, 00100, Kenya

ARTICLE INFO

Keywords:

Agrobiodiversity
Genebank management
Genetic resources
International genebanks
Seed longevity
Viability monitoring

ABSTRACT

Genebanks underpin global food security, conserving and distributing agrobiodiversity for use in research and breeding. The CGIAR collections include >700,000 seed accessions, held in trust as global public goods. However, the role of genebanks in contributing to global food security can only be realized if collections are effectively managed. Examination of the historical viability monitoring data from seven CGIAR genebanks confirmed that high seed viability was maintained for many decades for the various crops and forage species. However, departures from optimum management procedures were revealed, and there were insufficient data gathered to derive reliable estimates of longevity needed to better forecast regeneration requirements, estimate the size of seed lots that should be stored, and optimize accession monitoring intervals.

1. Introduction

Since the advent of agriculture, greater food production has resulted from increasing cropping area, the number of crops per annum and/or crop yield per unit area. Crop food production and supply chains are internationally interdependent. Moreover, they depend upon plant genetic material that has been collected and distributed internationally: for example, “new” crops were moved across oceans by colonialists from 1492 onwards, following the return of Columbus from the New World, whilst modern plant breeding developed and distributed improved

varieties of existing crops from the late-19th/early-20th centuries onwards (Kloppenborg, 1988).

The need for diverse germplasm adapted to different environments was recognised nationally in the late-19th century by Russia and the USA, both with large land masses suitable for improved agriculture. In 1894, Russia established its Bureau of Applied Botany to collect and study crop diversity, which quickly collected germplasm from within Russia and received other material from Canada and Sweden (Loskutov, 2020). By 1921, it had become the Department of Applied Botany and Plant Breeding (later, the All-Union Institute of Plant Industry) led by N.

* Corresponding author. Department of Agroecology, Aarhus University, Forsøgsvej 1, 4200, Denmark.

E-mail address: fiona.hay@agro.au.dk (F.R. Hay).

¹ Retired.

² Current address International Livestock Research Institute (ILRI), PO Box 5689, Addis Ababa, Ethiopia.

I. Vavilov. Its scientists collected and studied plant germplasm from across five continents (Vavilov, 1997; Loskutov, 2020). Similarly, in 1898, the United States Department of Agriculture (USDA) created its Section of Seed Production and Introduction and despatched seed collectors globally (Kloppenborg, 1988; Kaplan, 1998). In that first year, new durum and bread wheat varieties were collected from Russia and introduced successfully into the USA, such that, within five years, USA wheat production had increased from 60,000 to 20 million bushels a year (Kaplan, 1998). Both countries subsequently introduced seed stores to maintain their germplasm collections. The USDA opened the National Seed Storage Laboratory in 1958 dedicated to this task as a keystone of the USDA germplasm system (James, 1972; Griesbach, 2013).

The growing importance of germplasm to improve crop varieties to support the burgeoning world population after World War 2, the simultaneous loss of that germplasm by genetic erosion, and the growing understanding of the relevance of genebanks, led to international action. The International Board for Plant Genetic Resources (IBPGR) was created in 1974 and soon released recommendations on standards for long-term genetic resources conservation (IBPGR, 1976), detailed advice on the design of long-term seed stores (Cromarty et al., 1982), and protocols to germinate samples from seeds held in long-term storage to monitor their viability (Ellis et al., 1985a, b). By the end of its first decade, IBPGR was able to celebrate a network of 113 significant plant germplasm collections, including those of CGIAR (Hanson et al., 1984).

CGIAR was founded in 1971 as a global research partnership aiming to tackle the food crisis affecting many countries in the developing world. Today, the CGIAR Centers manage some of the oldest, largest and most diverse collections of staple food crops in the world that are held in-trust as global public goods (FAO, 2010). Together, they include more than 700,000 seed accessions stored in 27 crop collections at 10 CGIAR Centers, and further collections of tissue culture and live plants (<https://www.genebanks.org/resources/annual-reports/>). The diversity in the collections underpins CGIAR's research and breeding efforts and is shared upon request to users worldwide. Thus, the CGIAR genebanks play an important part in the delivery of improved crop varieties to meet wide-ranging goals to alleviate poverty, improve food and nutrition security, and address climate change (Galluzzi et al., 2016).

To be effective, genebanks need to ensure that the seed samples representing each and every accession are in a healthy state when put into storage and maintain high viability on a long-term basis (IBPGR, 1976). However, each plant species shows individual seed behaviour and differ in the period (by decades) they remain viable in storage (Walters et al., 2005; Ellis et al., 2018, 2019; Colville and Pritchard, 2019). In theory, as soon as seed viability falls to a certain threshold (typically 85%; FAO, 2014), accession regeneration should be triggered so that new stocks are generated for storage and use (Fig. 1). 'Active management' (i.e. further viability monitoring) of the low viability lot then ceases and it should be discarded unless it is still considered suitable for distribution (or for research, for example, as discussed later, to better understand longevity in storage). In this way, for each accession, no more than one seed lot in the active collection and one seed lot in the base collection should be under active management at any one time (Fig. 1). However, early on, it was recognised that the task to monitor the considerable number of seed accessions of diverse genotypes, including crop wild relatives, was not trivial. Furthermore, it was concluded that "many genebanks continue to place seeds in storage without adequate testing and many have not properly established routine monitoring regimes" (IBPGR, 1982).

IBPGR (1985) recommended genebanks test the viability of accessions initially upon acquisition or entry from the field and, thereafter, having been kept in optimal storage conditions (-18°C , 3–7% moisture content), retest at regular intervals of 5 or 10 years. Since this advice was published, the UN Food and Agriculture Organization (FAO) has held two rounds of consultations on genebank standards, the most recent being published in 2014 (FAO, 2014). With regards to viability testing, their advice reiterates the need for initial tests (within 12 months of the

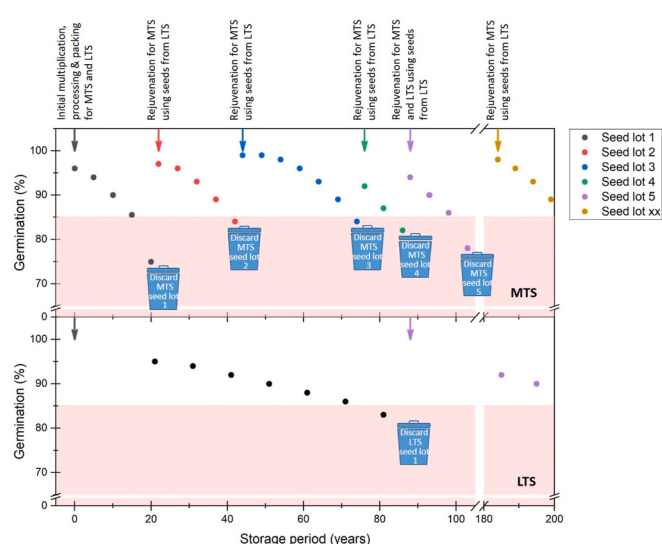


Fig. 1. Schematic showing the 'life' of a genebank accession in an example where seed lots are maintained in both medium-term storage (MTS) and long-term storage (LTS) environments, and where the genebank standard that "initial germination of a seed lot should exceed 85%" (FAO, 2014) is met. This schematic assumes that accession regeneration is not required because of loss of seed through distribution. After initial multiplication and processing for storage, seed lot 1 is divided and placed into both MTS and LTS. Entry into storage is indicated by the downward arrows, this colour corresponding to that of the test results. Viability monitoring tests are then every 5 years for seeds in MTS in this example. In accordance with "The viability threshold for regeneration or other management decision such as recollection should be 85% ..." (genebank standard 4.3.4; FAO, 2014), when the germination of seed lot 1 in MTS is $\leq 85\%$, a new seed lot is produced for the MTS. Viability monitoring of seed lot 1 in LTS should then commence, typically at 10-year intervals. It is preferable to use seeds from LTS (which are or are closer to the most original sample) for rejuvenation, though up to three cycles of rejuvenation can be made using seeds from MTS, in particular for in-breeding species (Sackville Hamilton and Chorlton, 1997). Once an accession has been 'rejuvenated', the remaining seeds of the previous seed lot may be used for immediate distribution, but should otherwise be discarded. The time scale indicated is not intended to reflect a particular crop species. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

accession being received) and subsequent retesting every 5 or 10 years. However, if the length of time it takes for the viability of a seed lot to fall to 85% can be estimated, then the monitoring test interval should be one-third of that period up to a maximum interval of 40 years (FAO, 2014).

The activities of the CGIAR genebanks have had their programmatic home in the CGIAR Genebank Platform (2017–2021), coordinated by the Global Crop Diversity Trust. Routine genebank operations, including accession management and viability monitoring, and some supporting conservation research, is organized under the Conservation Module of the Platform with an underlying principle of trying to improve operations while reducing costs. In this context, a better understanding of how to maximize the longevity of seeds before and during storage, and of the relative longevity of seed lots in storage may allow savings through the extension of viability monitoring and rejuvenation intervals. A number of genebanks have published historical viability data (e.g. Lee et al., 2013; Van Treuren et al., 2013; Yamasaki et al., 2020) and collaboration among genebanks in sharing such data has been encouraged (Solberg et al., 2020). Such a collaboration among CGIAR genebanks began in 2015, with the specific target to redefine storage periods for 20 crops by 2022 (<https://www.genebanks.org/wp-content/uploads/2019/09/Genebanks-Platform-Full-Proposal.pdf>).

This collaboration involved discussions on operations at, and extraction of viability monitoring records for seed collections in, seven

CGIAR genebanks. Theoretically, it should be possible to identify trends in the data over long time periods or even model the data following customary approaches (Ellis and Roberts, 1980; Hay et al., 2014). The results of such analysis should inform both individual genebank management and help improve genebank standards and advice to genebank managers in general. However, analysis of real genebank data is challenging for a number of reasons (Hay and Whitehouse, 2017). In many cases, rather than gaining better understanding of seed longevity in genebanks, the current study has highlighted many issues related to sustaining processes to test seeds for viability over a long time period and using viability data to manage seed collections.

2. Methods

2.1. Data

The data from the seven CGIAR genebanks were provided between 2013 and 2017 (Tables 1–4). At that time, all these genebanks used different database systems and the data recorded varied. The data requested included: accession number, taxon, type of material (e.g. traditional cultivar/landrace, advanced/improved cultivar, research/breeding material), country of origin, harvest season/year, storage type, date seeds were placed into storage, date tested, test result (% viability or % germination). ‘Storage type’ refers to whether the seeds were in medium-term storage (MTS, normally used for the active collection; seeds typically stored at 2–5°C) or in long-term storage (LTS, base collection; seeds typically stored at -20°C). Information on how seeds were handled after harvest was also collected. At the time of requesting this information, none of these genebanks routinely recorded germination test conditions within their respective databases.

The data were provided in Excel file(s), with different files and/or sheets for different crops and/or different types of storage (medium- or long-term). Data were ‘tidied’ and sorted prior to analysis. Examples of data checks made included:

- That there was consistency in format of fields (e.g. accession number and harvest year or season, where a mistake in data entry might suggest different accessions and/or seed lots);
- That dates were correctly and consistently formatted and that processing steps were in the correct sequence (for example, a viability test was not recorded as having been made before a seed lot was harvested);
- That viability test results were in the range 0–100%;
- That records were not duplicated or missing data.

Examples of manipulations that were made:

- Combining of data sets containing historical and current viability and/or rearrangement of data, so that there was only one field for each of viability result and date tested;
- Many genebanks did not record the date when seed lots were placed into storage; ‘date banked’ was therefore estimated, as a few months after the harvest year and/or season (most genebanks) or date processed (IRRI) in order to estimate how long a seed lot had been stored at the time of testing;
- Creation of a new field that combined accession number and seed lot to identify unique seed lots.

2.2. Analysis

We expected to use probit analysis to fit the Ellis-Roberts viability equation, $v = K_i - p/\sigma$ (Ellis and Roberts, 1980), to the data for seed lots that had been tested at least three times. This equation describes the linear relationship between probit viability, v , as a function of the duration of storage, p , under constant conditions (moisture content and temperature), taking into account the binomial error distribution of

germination data (Hay et al., 2014). K_i is the theoretical initial viability in probits and σ is the period it takes for viability to fall by 1 probit. This equation can thus be used to set appropriate monitoring intervals. The assumption that the storage conditions are constant cannot be tested, though given the age of some of the genebanks and location (i.e. in countries where electricity supply cannot always be guaranteed), it is perhaps likely that there were some variations in the storage temperature at least. Furthermore, not all of the genebanks store the seeds in air-tight containers (e.g. ICRISAT), and thus it would not be surprising if there was also some variation in the moisture content of the seeds during storage. Even when seed lots are reportedly dried and processed identically and then stored in air-tight containers, their moisture content may vary to some degree during storage (Hay et al., 2013; S. Timple and F.R. Hay, unpublished data from the IRRI genebank), perhaps because the conditions under which seed lots are packed or where containers are opened for sampling are not effectively controlled. However, more fundamentally, the aggregate dataset only had sequential data for a small proportion of seed lots, the most comprehensive set being from IRRI. Probit analysis was used for rice (*Oryza glaberrima* Steud. and *O. sativa* L.) data from IRRI and for *Phaseolus vulgaris* L. data from CIAT, for seed lots with at least four viability results. The effect of constraining the estimate of σ to a common value for all the seed lots was also evaluated through approximate F-tests (see e.g. Ellis et al., 2018, 2019).

Various other methods of data interrogation and presentation were explored for the data from other genebanks. Specifically, to be able to give a measure of the potential longevity of seeds for a particular crop/species in genebank storage, we determined the longest estimated storage period (based on the storage periods calculated as described above) where a viability result of 95% or higher had been reported.

3. Results

The entire data set comprised 874,427 observations (viability test results) of 822 different species (Tables 1–4). Some species are conserved in more than one genebank. For example, the two cultivated rice species (*Oryza glaberrima* and *O. sativa*) are conserved at both AfricaRice and IRRI, and chickpea at both ICARDA and ICRISAT. The earliest seed lots covered by the data were harvested in the late-1970s (Table 1). The earliest monitoring tests covered by the data were made in 1989 (ICRISAT), 1990 (ICARDA), 1991 (IRRI), 1996 (World Agroforestry Centre), 1997 (CIAT), 2004 (IITA, not including two results apparently from 1987) and 2014 (AfricaRice; date of initial test not recorded). The estimated storage periods spanned up to 37 years (Tables 1–4; Fig. 2), though for some crops (species)/genebanks it was much less.

While most of the viability data was gathered through germination tests, the noticeable exception to this was the forage collections at CIAT (Tables 3 and 4), particularly those of Poaceae where the result of a tetrazolium test (a seed embryo staining test to identify viable tissue; ISTA, 2021) was the only available estimate of viability for almost half of the seed lots (Table 4). The viability range covered by the data varied depending on crop and genebank (Appendix Figs. A1–A4). For example, while most of the data for *Oryza glaberrima* and *O. sativa* were above 85% for the collections at both AfricaRice and IRRI (Appendix Fig. A1), the germination data for the collections at IITA spanned the entire range, 0–100% (Appendix Fig. A2). Furthermore, in the case of the data from IITA, a substantial proportion (45 and 54% for *Vigna subterranea* (L.) Verdc. and *V. unguiculata* (L.) Walp., respectively) of germination results was $\leq 85\%$. In contrast, the data for some of the collections covered a narrower range. For example, there were very few observations below 50% for the *Triticum aestivum* data from ICARDA. It was not possible to determine why there was an apparent cut-off in the viability results recorded, not least since the ICARDA genebank has relocated since the data were collected, but it cast doubt on the validity of using the data for further analysis. Other anomalies in the data were also identified. For example, there were 94 results of 83% germination in the first available

Table 1

Summary of the historical viability monitoring data for cultivated crops included in the analysis to assess seed longevity in medium- and long-term genebank storage (MTS and LTS, respectively).

Center and date data provided	Comments on data	Crop	Species	Viability threshold	Storage environment (s)	Date range covered by data	Number of seeds tested	Number of observations with date and germination data	Number of accessions represented in data	Number of seed lots represented in data	Maximum storage period (years)	Maximum storage period with $\geq 95\%$ viability (years)
AfricaRice May 2016	- MTS data comprises just one result for most accessions. - Some MTS seed lots with more than one result may be duplicated entries in some cases, but there are also instances when more than one result (different % values) was recorded for tests made at the same time. - LTS data includes initial and most recent germination result. - LTS initial test data does not have date of testing.	Rice	<i>Oryza glaberrima</i>	85%	MTS: seeds dried at 15%RH, 15 °C and then stored at 2–4 °C	Seed lots from 2004 to 2015	2 × 50	728	604	604	11.5	11.3
					LTS: seeds dried at 15%RH, 15 °C and then stored at –20 °C	Seed lots from 1981 to 1996		1753	1524	1529	34.5	34.5
			<i>Oryza sativa</i>	85%	MTS: seeds dried at 15%RH, 15 °C and then stored at 2–4 °C	Seed lots from 2009 to 2014	2 × 50	1181	1133	1133	6.9	6.7
					LTS: seeds dried at 15%RH, 15 °C and then stored at –20 °C	Seed lots from 1982 to 1996		2790	2750	2787	33.5	33.5
IRRI October 2013 (<i>O. sativa</i> data described in Hay et al., 2015)	- Date seeds placed into storage estimated from processing date. - More than one germination result available for many seed lots.	Rice	<i>Oryza glaberrima</i>	85%	MTS: seeds dried at 15%RH, 15 °C and then stored at 2–4 °C.	Seed lots from 1984 to 2003	2 × 100	2501	797	798	24.6	24.6
					LTS: seeds dried at 15%RH, 15 °C and then stored at –20 °C (–10 °C prior to 1993)	Seed lots from 1984 to 2003	2 × 50	1534	757	757	28.8	22.6
			<i>Oryza sativa</i>	85%	MTS: seeds dried at 15%RH, 15 °C and then stored at 2–4 °C.	Seed lots from 1979 to 2003	2 × 100	254,654	69,232	71,855	31.9	30.0
					LTS: seeds dried at 15%RH, 15 °C and then stored at –20 °C (–10 °C prior to 1993)		2 × 50	149,165	69,953	70,918	32.5	32.3
IITA May 2016	- Only initial and most recent monitoring test results stored in genebank inventory. - Initial germination result and/or initial test date not available for all seed lots. - Harvest date not available for all seed lots. - Initial germination expressed as a proportion; most recent germination expressed as %.	Bambara groundnut	<i>Vigna subterranea</i>	80–85%	MTS: seeds dried at 17 °C, 14–15% RH and then stored at 5 °C	Seed lots from 1987 to 2015	1 × 20	5330	1679	3388	27.6	27.2
					LTS: seeds dried at 17 °C, 14–15% RH and then stored at –20 °C	Seed lots from 1985 to 2015		4283	1515	2904	28.6	27.8
		Cowpea	<i>Vigna unguiculata</i>	80–85%	MTS: seeds dried at 17 °C, 14–15% RH and then stored at 5 °C	Seed lots from 1978 to 2017	1 × 20	46,690	14,213	31,938	33.8	32.8
					LTS: seeds dried at 17 °C, 14–15% RH and then stored at 5 °C	Seed lots from		17,457	7579	8908	34.8	28.0

(continued on next page)

Table 1 (continued)

Center and date data provided	Comments on data	Crop	Species	Viability threshold	Storage environment (s)	Date range covered by data	Number of seeds tested	Number of observations with date and germination data	Number of accessions represented in data	Number of seed lots represented in data	Maximum storage period (years)	Maximum storage period with $\geq 95\%$ viability (years)
ICRISAT November 2016–January 2017	- Some harvest bulks were split into different containers and the same initial germination result entered for the sub-lots; sub-lots monitored independently. - Only most recent germination result stored in operational database; historical data downloaded from archive. - Initial test not routinely made. - Some accessions or seed lots only in one environment. - Only one or two results available for most seed lots x storage environment. - Recent change: monitoring intervals now varied depending on initial germination result. - MTS second result often the same or very similar to the first result, regardless of the value of the first result.	Maize	<i>Zea mays</i>	80–85%	RH and then stored at -20°C MTS: seeds dried at 17°C , 14–15% RH and then stroed at 5°C LTS: seeds dried at 17°C , 14–15% RH and then stroed at -20°C	1978 to 2011 Seed lots from 2004 to 2015 Seed lots from 2008 to 2016	1 × 20	3976	770	2222	11.3	11.3
								2616	1256	1407	11.5	11.4
		Chickpea	<i>Cicer arietinum</i>	85%	MTS: seeds in non-air-tight containers at 30% RH, 5°C . LTS: seeds dried at 15% RH, 15°C and then stored at -20°C .	Seed lots from 1974 to 2015	50	22,312	20,173	20,173	35.3	35.3
								22,918	17,319	17,319	25.1	25.1
		Groundnut	<i>Arachis hypogaea</i>	80%	MTS: seeds in non-air-tight containers at 30% RH, 5°C . LTS: seeds dried at 15% RH, 15°C and then stored at -20°C .	Seed lots from 1980 to 2015	2 × 25	19,197	14,743	14,743	33.3	33.3
								15,066	14,039	14,039	24.7	24.4
		Pearl millet	<i>Pennisetum glaucum</i>	85%	MTS: seeds in non-air-tight containers at 30% RH, 5°C . LTS: seeds dried at 15% RH, 15°C and then stored at -20°C .	Seed lots from 1973 to 2015	2 × 50	22,925	22,250	22,250	36.6	36.1
								22,156	21,956	21,956	25.7	25.7
		Pigeon pea	<i>Cajanus cajan</i>	85%	MTS: seeds in non-air-tight containers at 30% RH, 5°C . LTS: seeds dried at 15% RH, 15°C and then stored at -20°C .	Seed lots from 1975 to 2015	2 × 50	13,322	12,861	12,861	36.5	32.9
								12,349	12,033	12,033	23.8	23.8
		Sorghum	<i>Sorghum bicolor</i>	85%	MTS: seeds in non-air-tight containers at 30% RH, 5°C . LTS: seeds dried at 15% RH, 15°C and then stored at -20°C .	Seed lots from 1975 to 2015		40,288	39,445	39,445	37.5	32.3
								38,798	38,242	38,242	25.4	25.4

(continued on next page)

Table 1 (continued)

Center and date data provided	Comments on data	Crop	Species	Viability threshold	Storage environment (s)	Date range covered by data	Number of seeds tested	Number of observations with date and germination data	Number of accessions represented in data	Number of seed lots represented in data	Maximum storage period (years)	Maximum storage period with $\geq 95\%$ viability (years)
ICARDA April 2016	• Separate database for each crop. • No initial testing for harvest years covered by data set. • Most seed lots have only one result. • Apparent censoring of data showing <50% viability	Chickpea	<i>Cicer arietinum</i>	85%	MTS: seeds dried at 15% RH, 15 °C and then stored at 2 °C.	Seed lots from 1986 to 2010	2 × 50	12,832	11,223	12,799	20.4	16.3
		Wheat	<i>Triticum aestivum</i> subsp. <i>aestivum</i>	85%	MTS: seeds dried at 15% RH, 15 °C and then stored at 2 °C.	Seed lots from 1974 to 2011	2 × 50	15,110	13,138	14,876	32.0	27.4
			<i>Triticum turgidum</i> subsp. <i>durum</i>	85%	MTS: seeds dried at 15% RH, 15 °C and then stored at 2 °C.	Seed lots from 1973 to 2010	2 × 50	17,098	15,538	16,993	31.3	31.3
CIAT September 2017, January 2019	• Seed lots only stored in long-term conditions.	Bean	<i>Phaseolus vulgaris</i>	85%	LTS: three drying steps comprising 1.20 °C/35% RH for 3–4 days; 2.20 °C/20% RH 1 week; 3.15–20 °C/10% RH 3–4 weeks. Seeds vacuum packed and stored at –18 °C.	Seed lots from 1979 to 2017	1 × 50 (2 × 50 in 1990s)	59,640	25,987	32,272	27.2	27.2

NA = not available.

Table 2

Summary of viability monitoring data provided by the World Agroforestry genebank in May 2019. Seed lots were harvested between 1980 and 2017, although some of the oldest collections were donated and exact processing and storage conditions were not recorded. Subsequently seeds were dried at 15% RH and 15 °C and then stored hermetically in aluminium foil bags at 5 °C. Germination protocols and number of seeds tested are specific to each species.

Family	Genus	Number of species†	Number of observations	Number of accessions represented in data	Number of seed lots represented in data	Proportion of seed lots only tested once	Maximum estimated storage period (years)	Maximum storage period with ≥95% viability (years) ‡	Maximum storage period with ≥50% viability (years) ‡	Maximum reported germination (%)
Anacardiaceae	<i>Sclerocarya</i> Hochst.	1	2	1	1	0	2.2	NA	NA	26
Annonaceae	<i>Annona</i> L.	1	1	1	1	100	4.0	NA	NA	0
Asclepiadaceae	<i>Calotropis</i> R. Br.	1	43	43	43	100	1.1	NA	NA	44
Asteraceae	<i>Tithonia</i> Desf. ex Juss.	1	5	3	3	33.3	17.0	NA	NA	7
Betulaceae	<i>Alnus</i> Mill.	1	1	1	1	100	6.5	NA	NA	10
Bignoniaceae	<i>Markhamia</i> Seem. ex Baill.	1	2	2	2	100	9.5	NA	NA	20
Bombacaceae	<i>Adansonia</i> L.	1	171	129	129	74.4	12.0	0	6.6	100
Boraginaceae	<i>Cordia</i> L.	1	2	1	1	0	3.1	NA	2.4	57
Casuarinaceae	<i>Casuarina</i> L.	1	3	2	2	50	11.4	NA	NA	5
Cupressaceae	<i>Juniperus</i> L.	1	1	1	1	100	10.6	NA	NA	0
Euphorbiaceae	<i>Bridelia</i> Willd.	1	1	1	1	100	10.5	NA	NA	0
	<i>Croton</i> L.	1	2	2	2	100	11.4	NA	NA	10
	<i>Jatropha</i> L.	1	2	1	1	0	6.2	NA	NA	26
Fabaceae	<i>Acacia</i> Mill.	17	447	270	270	61.1	22.3	10.7	22.3	100
	<i>Acrocarpus</i> Wight ex Arn.	1	1	1	1	100	4.1	NA	NA	8
	<i>Azizelia</i> Sm.	1	3	1	1	0	3.2	NA	NA	45
	<i>Albizia</i> Benth.	10	88	51	51	39.2	11.8	NA	11.8	90
	<i>Cajanus</i> Adans.	1	1	1	1	100	9.0	NA	NA	0
	<i>Calliandra</i> Benth.	2	138	98	98	73.5	14.4	NA	12.1	93
	<i>Crotalaria</i> L.	4	8	5	5	20	12.2	NA	11.0	50
	<i>Cytisus</i> Desf.	1	25	19	19	68.4	2.2	NA	2.2	52
	<i>Dalbergia</i> L. f.	1	1	1	1	100	6.6	NA	6.6	78
	<i>Delonix</i> Raf.	1	9	8	8	87.5	7.1	5.1	7.1	95
	<i>Enterolobium</i> Mart.	1	3	3	3	100	18.2	NA	NA	35
	<i>Erythrophleum</i> Afzel. ex R. Br.	1	4	3	3	33.3	9.1	NA	NA	22
	<i>Faidherbia</i> A. Chev.	1	188	150	150	82.7	16.8	13.4	16.8	100
	<i>Gliricidia</i> Kunth	2	40	23	23	78.3	9.1	NA	5.2	94
	<i>Hardwickia</i> Roxb.	1	1	1	1	100	8.9	NA	NA	0
	<i>Hesperalbizia</i> Barneby & J. W. Grimes	1	2	1	1	0	13.1	NA	13.1	60
	<i>Hybosema</i> Harms	1	2	1	1	0	2.6	NA	NA	45
	<i>Leucaena</i> Benth.	20	501	389	389	77.4	14.9	14.9	14.9	100
	<i>Macroptilium</i> (Benth.) Urb.	1	1	1	1	100	9.0	NA	NA	9
	<i>Milletia</i> Wight & Arn.	2	14	13	13	92.3	15.0	NA	NA	0
	<i>Mimosa</i> L.	1	1	1	1	100	0	NA	NA	4
	<i>Mucuna</i> Adans.	1	2	1	1	0	9.9	NA	NA	5
	<i>Parkinsonia</i> L.	1	1	1	1	100	4.2	NA	NA	7
	<i>Pericopsis</i> Thwaites	1	1	1	1	100	9.0	NA	NA	0
	<i>Philenoptera</i> Hochst. ex A. Rich.	1	1	1	1	100	15.1	NA	NA	0
	<i>Prosopis</i> L.	3	7	6	6	83.3	10.2	NA	NA	18
	<i>Senna</i> Mill.	2	9	4	4	0	9.0	NA	1.9	82
	<i>Sesbania</i> Adans.	3	85	55	55	60	18.1	17.7	17.7	100
	<i>Tamarindus</i> L.	1	3	3	3	100	9.0	1.8	7.9	100
	<i>Tephrosia</i> Pers.	2	29	4	4	0	11.1	3.6	8.6	98
	<i>Tetrapleura</i> Benth.	1	2	2	2	100	15.0	NA	NA	1
Flacourtiaceae		1	1	1	1	100	3.2	NA	NA	22

(continued on next page)

Table 2 (continued)

Family	Genus	Number of species†	Number of observations	Number of accessions represented in data	Number of seed lots represented in data	Proportion of seed lots only tested once	Maximum estimated storage period (years)	Maximum storage period with ≥95% viability (years) ‡	Maximum storage period with ≥50% viability (years) ‡	Maximum reported germination (%)
	<i>Dovyalis</i> E. Mey. ex Arn.									
Lamiaceae	<i>Tectona</i> L. f.	1	1	1	1	100	0	NA	NA	1
	<i>Vitex</i> L.	2	6	5	5	80	6.0	NA	0.8	87
Malvaceae	<i>Abutilon</i> Mill.	2	4	2	2	50	13.2	NA	13.2	75
	<i>Grewia</i> L.	1	1	1	1	100	9.1	9.1	9.1	98
Meliaceae	<i>Melia</i> L.	1	2	2	2	100	6.0	NA	6.0	50
Moraceae	<i>Milicia</i> Sim	1	1	1	1	100	4.3	NA	NA	0
Moringaceae	<i>Moringa</i> Adans.	3	12	7	7	57.1	15.0	NA	6.7	93
Myrtaceae	<i>Eucalyptus</i> L'Hér.	2	2	2	2	100	6.9	6.9	6.9	100
Oleaceae	<i>Olea</i> L.	1	1	1	1	100	3.7	NA	NA	10
Passifloraceae	<i>Passiflora</i> L.	1	2	2	2	100	7.7	NA	NA	20
Podocarpaceae	<i>Afrocarpus</i> (J. Buchholz & N. E. Gray) C. N. Page	1	1	1	1	100	8.3	NA	NA	0
Proteaceae	<i>Grevillea</i> R. Br. ex Knight	1	1	1	1	100	3.8	NA	NA	4
Rhamnaceae	<i>Ziziphus</i> Mill.	1	3	2	2	50	10.5	NA	10.5	77
Rubiaceae	<i>Calycophyllum</i> DC.	1	34	34	34	100	18.1	18.0	18.0	100
	<i>Cinchona</i> L.	1	1	1	1	100	6.9	NA	NA	20
	<i>Vangueria</i> Juss.	1	1	1	1	100	7.7	NA	NA	0
Solanaceae	<i>Solanum</i> L.	1	1	1	1	100	4.5	NA	4.5	90
Sterculiaceae	<i>Guazuma</i> Mill.	1	101	64	64	51.6	16.0	NA	14.9	56
Theaceae	<i>Schima</i> Reinw. ex Blume	1	1	1	1	100	7.6	NA	NA	0
	OVERALL	125	2031	1438	1438	70.6	22.3			

Table 3

Summary of observations made about CIAT's Fabaceae forage collection. For processing and storage conditions, see information in Table 1 for *Phaseolus vulgaris* seeds stored at CIAT.

Genus	Number of species ^a	Number of observations	Number of accessions represented in data	Number of seed lots represented in data	Proportion of seed lots only tested with tetrazolium	Proportion of seed lots only tested once	Maximum estimated storage period	Maximum storage period with ≥95% viability (years) ^b
<i>Abrus</i> Adans.	2	73	42	51	0	76.5	26.2	6.7
<i>Acacia</i> Mill.	6	40	16	21	0	28.6	26.2	13.0
<i>Aeschynomene</i> L.	32	2441	1172	1553	0	54.9	21.1	17.8
<i>Albizia</i> Benth.	1	1	1	1	0	100.0	8.7	NA
<i>Alysicarpus</i> Neck. ex Desv.	10	632	249	292	0	25.7	31.1	26.2
<i>Arachis</i> L.	25	231	111	179	31.8	77.1	14.0	<1
<i>Atylosia</i> Wight & Arn.	1	1	1	1	0	100.0	2.2	NA
<i>Bauhinia</i> L.	1	9	4	5	0	40.0	17.1	12.0
<i>Cajanus</i> Adans.	5	293	130	195	0	32.8	19.2	14.6
<i>Calliandra</i> Benth.	3	37	23	26	0	57.7	15.0	8.4
<i>Calopogonium</i> Desv.	4	1242	538	725	0	49.2	31.1	27.2
<i>Camptosema</i> Hook. & Arn.	4	9	6	7	0	85.7	14.0	0
<i>Canavalia</i> Adans.	13	528	207	356	0	59.6	18.1	18.1
<i>Cassia</i> L.	1	2	1	1	0	0.0	7.5	7.5
<i>Centrosema</i> (DC.) Benth.	34	5704	2772	3630	0	58.6	35.1	24.1
<i>Chaetocalyx</i> DC.	3	4	3	4	0	100.0	5.0	5.0
<i>Chamaecrista</i> (L.) Moench	17	730	395	479	0	51.1	18.2	11.9
<i>Christia</i> Moench	2	23	13	14	0	35.7	9.6	8.8
<i>Clitoria</i> L.	5	258	151	189	0.5	67.7	14.8	13.0
<i>Codariocalyx</i> Hassk.	2	108	36	44	0	31.8	27.1	27.1
<i>Coursetia</i> DC.	1	29	12	19	0	63.2	16.1	NA
<i>Cratylia</i> Mart. ex Benth.	1	141	50	70	1.4	55.7	23.2	17.0
<i>Crotalaria</i> L.	25	794	268	471	0	54.6	27.2	16.3
<i>Dalea</i> L. Mill.	2	10	3	8	0	75.0	9.8	0
<i>Dendrolobium</i> (Wight & Arn.) Benth.	4	95	53	67	0	62.7	25.1	12.2
<i>Desmanthus</i> Willd.	13	644	325	372	0	36.8	21.9	18.2
<i>Desmodium</i> Desv.	61	6583	3236	3979	0.1	50.0	31.1	25.1
<i>Dicerma</i> DC.	1	1	1	1	0	100.0	1.3	NA
<i>Dichrostachys</i> (DC.) Wight & Arn.	1	2	1	1	0	0.0	11.7	NA
<i>Dioclea</i> Kunth	12	449	193	235	0	44.3	27.2	23.2
<i>Dolichopsis</i> Hassl.	2	9	7	9	0	100.0	3.5	3.5
<i>Dolichos</i> L.	4	27	14	16	0	43.8	19.2	11.6
<i>Dunbaria</i> Wight & Arn.	5	23	13	15	0	73.3	15.0	15.0
<i>Dysolobium</i> (Benth.) Prain	3	6	5	5	0	80.0	5.5	NA
<i>Entada</i> Adans.	1	1	1	1	0	100.0	3.1	NA
<i>Eriosema</i> (DC.) Desv.	5	54	25	29	0	41.4	18.1	11.1
<i>Erythrina</i> L.	3	7	3	3	0	33.3	12.8	NA
<i>Flemingia</i> Roxb. ex Rottler	6	410	173	231	0	48.5	29.1	16.0
<i>Galactia</i> P. Browne	12	1189	550	916	0	74.6	25.1	23.0
<i>Gliricidia</i> Kunth	1	13	11	12	0	91.7	12.7	0
<i>Glycine</i> Willd.	3	13	6	9	0	55.6	11.2	11.2
<i>Hegnera</i> Schindl.	1	3	2	3	0	100.0	1.2	NA
<i>Indigofera</i> L.	13	515	179	236	0.8	33.5	21.1	17.4
<i>Lablab</i> Adans.	2	275	156	189	0	65.6	16.0	12.0
<i>Leucaena</i> Benth.	17	452	182	335	0	71.6	27.2	20.9
<i>Lotononis</i> (DC.) Eckl. & Zeyh.	1	6	1	2	0	50.0	15.1	NA

(continued on next page)

Table 3 (continued)

Genus	Number of species ^a	Number of observations	Number of accessions represented in data	Number of seed lots represented in data	Proportion of seed lots only tested with tetrazolium	Proportion of seed lots only tested once	Maximum estimated storage period	Maximum storage period with $\geq 95\%$ viability (years) ^b
<i>Macroptilium</i> (Benth.) Urb.	10	2003	1046	1399	0	63.8	20.1	18.3
<i>Macrotyloma</i> (Wight & Arn.) Verdc.	5	59	35	37	0	51.4	21.5	21.5
<i>Medicago</i> L.	1	5	2	2	0	0.0	10.5	9.2
<i>Mimosa</i> L.	12	55	23	27	0	40.7	31.1	10.4
<i>Mucuna</i> Adans.	2	48	25	28	3.6	60.7	16.7	16.0
<i>Neonotonia</i> J. A. Lackey	1	170	68	87	0	41.4	21.1	17.2
<i>Neptunia</i> Lour.	2	5	3	3	0	66.7	21.6	21.6
<i>Pachecoa</i> Standl. & Steyerf.	1	1	1	1	0	100.0	−0.2	NA
<i>Pachyrhizus</i> Rich. ex DC.	2	20	12	20	0	100.0	21.3	16.3
<i>Parkinsonia</i> L.	1	2	2	2	0	100.0	5.2	5.2
<i>Parosela</i> Cav.	1	4	1	2	0	50.0	13.8	0
<i>Periandra</i> Mart. ex Benth.	4	30	21	22	0	63.6	13.8	13.0
<i>Phyllodium</i> Desv.	4	255	137	165	0	49.1	18.9	18.9
<i>Piptadenia</i> Benth.	1	1	1	1	0	100.0	4.0	4.0
<i>Prosopis</i> L.	1	18	9	11	0	54.5	15.1	13.4
<i>Pseudarthria</i> Wight & Arn.	2	128	70	86	0	53.5	23.2	8.9
<i>Psophocarpus</i> Neck. ex DC.	2	9	3	5	0	60.0	13.8	13.8
<i>Pueraria</i> DC.	5	556	239	405	0	68.6	26.2	18.9
<i>Pycnospora</i> R. Br. ex Wight & Arn.	1	74	30	32	0	18.8	15.9	15.9
<i>Rhynchosia</i> Lour.	12	715	362	401	0	48.1	32.1	23.6
<i>Senna</i> Mill.	13	43	26	31	3.2	67.7	11.7	6.6
<i>Sesbania</i> Adans.	15	170	61	87	0	43.7	27.2	20.6
<i>Shuteria</i> Wight & Arn.	1	1	1	1	0	100.0	16.0	16.0
<i>Soemmeringia</i> Mart.	1	5	2	2	0	0.0	9.6	7.8
<i>Stylosanthes</i> Sw.	29	7323	3750	4875	0	53.9	34.1	21.0
<i>Tadehagi</i> H. Ohashi	2	190	107	132	0	56.8	15.6	15.6
<i>Tephrosia</i> Pers.	16	357	145	205	0	50.2	18.7	18.7
<i>Teramnus</i> P. Browne	5	891	374	498	0.2	45.4	25.2	20.6
<i>Teyleria</i> Backer	1	2	1	1	0	0.0	6.9	0.3
<i>Trifolium</i> L.	3	6	3	3	0	33.3	16.5	16.5
<i>Uraria</i> Desv.	4	276	162	183	0	56.3	16.6	16.6
<i>Vigna</i> Savi	35	1792	974	1259	0	68.0	21.9	16.5
<i>Zapoteca</i> H. M. Hern.	2	8	4	6	0	66.7	16.7	NA
<i>Zornia</i> J. F. Gmel.	13	1665	890	958	0	37.9	35.1	17.0
<i>Leguminosa</i> indet. faboideae	1	1	1	1	0	100.0	−0.5	
OVERALL	582	41,005	19,932	25,985	0.3	54.8	35.1	

^a Accessions not identified to species level counted as 1; hybrids and *sp. nov.* also counted as separate species in this tally.

^b NA indicates that the maximum germination recorded was <95%; entries with 0 indicate that the result was the initial test result (with estimated storage period <0 years).

Table 4

Summary of observations made about CIAT's Poaceae forage collection. For processing and storage conditions, see information in Table 1 for *Phaseolus vulgaris* seeds stored at CIAT.

Genus	Number of species ^a	Number of observations	Number of accessions represented in data	Number of seed lots represented in data	Proportion of seed lots only tested with tetrazolium	Proportion of seed lots only tested once	Maximum estimated storage period (years)	Maximum storage period with $\geq 95\%$ viability (years) ^b
<i>Andropogon</i> L.	2	84	52	59	59.3	71.2	22.2	17.1
<i>Axonopus</i> P. Beauv.	2	12	12	11	81.8	90.9	11.0	5.0
<i>Brachiaria</i> (Trin.) Griseb.	19	1362	500	704	62.5	50.7	20.2	17.1
<i>Cenchrus</i> L.	4	35	20	31	64.5	87.1	18.7	8.9
<i>Chloris</i> Sw.	4	40	13	20	15.0	50.0	16.1	11.1
<i>Chrysopogon</i> Trin.	1	2	2	2	100.0	100.0	3.1	NA
<i>Coix</i> L.	1	1	1	1	0.0	100.0	0.4	NA
<i>Cynodon</i> Rich.	1	3	2	2	100.0	50.0	8.1	NA
<i>Dactyloctenium</i> Willd.	2	5	2	2	50.0	0.0	11.1	0
<i>Dichanthium</i> Willemet	2	3	2	2	50.0	50.0	9.3	NA
<i>Digitaria</i> Haller	2	4	3	4	25.0	100.0	0	NA
<i>Echinochloa</i> P. Beauv.	5	22	12	17	64.7	82.4	15.2	3.8
<i>Enteropogon</i> Nees	1	3	1	1	0.0	0.0	14.2	14.2
<i>Eragrostis</i> Wolf	6	44	18	23	47.8	43.5	15.2	15.2
<i>Hyparrhenia</i> Andersson ex E. Fourn.	9	55	33	43	58.1	81.4	28.5	11.2
<i>Leptochloa</i> P. Beauv.	1	3	1	2	100.0	50.0	6.1	0
<i>Melinis</i> P. Beauv.	1	21	10	10	30.0	30.0	15.0	NA
<i>Panicum</i> L.	9	852	441	738	46.3	86.9	25.2	20.3
<i>Paspalum</i> L.	16	93	66	80	51.3	86.3	21.2	2.2
<i>Pennisetum</i> Rich.	5	40	23	27	74.1	63.0	20.7	0
<i>Rottboellia</i> L. f.	1	1	1	1	100.0	100.0	-0.3	NA
<i>Secale</i> L.	1	4	1	2	0.0	50.0	11.1	11.1
<i>Setaria</i> P. Beauv.	3	16	14	13	38.5	76.9	17.7	NA
<i>Sorghum</i> Moench	3	8	4	5	60.0	60.0	15.2	8.6
<i>Sporobolus</i> R. Br.	1	1	1	1	100.0	100.0	2.1	NA
<i>Urochloa</i> P. Beauv.	2	8	4	6	50.0	66.7	11.3	NA
OVERALL	104	2722	1239	1807	54.3	70.0	28.5	

^a Accessions not identified to species level counted as 1; hybrids also counted as separate species in this tally.

^b NA indicates that the maximum germination recorded was $<95\%$; entries with 0 indicate that the result was the initial test result (with estimated storage period <0 years).

test result for pearl millet seed lots in MTS at ICRISAT, all of which provided the identical result in the second viability test 14–15 years later (Appendix Fig. B1). Another issue was that for many of the seed lots, only two results were available, the initial result and the most recent result. Moreover, the difference between initial and most-recent viability test results was highly variable, and in many comparisons, was in fact positive (Appendix Figs B1, B2). Whilst re-assuring for the maintenance of seed accession viability during long-term storage, it was difficult to be consistent in the approach to data analysis and understanding seed storage longevity due to this variation in the nature of the data.

Most of the viability data considered was collected when seeds lots had been in storage for less than 12.5 years (Fig. 2), due in part to high rates of regeneration over the last 10–15 years. The data also showed, on occasion, that several different seed lots of the same accession were being actively managed (viability monitored) simultaneously (see as examples, Appendix Fig. B5 and B6). For some collections, the proportion of results $\leq 85\%$ relative to the proportion of results $>85\%$ increased the longer the estimated storage period, for example rice seeds in MTS at IRRI and seeds of *Vigna* spp. at IITA. Nonetheless, clearly, many seed lots were able to maintain high viability for at least three decades in MTS and LTS (Tables 1–4; Fig. 2; Appendix Figs A1–A4).

Given the size of the data set and the length of time that the genebanks have been in operation, there were relatively few seed lots with a series of viability results. Conversely, for collections where such data was available, the number of seed lots covered was large, for example

there were 209 *Phaseolus vulgaris* seed lots in LTS at CIAT with at least four viability test results each. In the case of these *P. vulgaris* seed lots' results, probit analysis was applied to model seed survival and it was possible to constrain the data for the different seed lots to a common slope, however, the slope was positive (i.e. viability seemingly increased during storage). In similar analyses of data for IRRI's *Oryza glaberrima* and *O. sativa* accessions (with at least four or at least five sequential observations, respectively), it was not possible to accept the common slope model for either species.

4. Discussion

4.1. Insights from the data

Examination of historical viability monitoring data from genebanks provides insights into how each genebank has been managed over time, and their respective priorities and protocols. The data also exposes several issues that a genebank is confronted with in relation to the effective and efficient management of collections. The first issue this study highlights is the importance of a robust information management system (Fu, 2017) and of effective and consistent documentation processes. Genebank data management systems and their constitutive data have, in the past, developed organically. Many started as paper-based systems and, after conversion to computer-based systems, may have had multiple iterations (including spreadsheets) over the years; possibly with more than one system 'live' at the same time within one genebank.

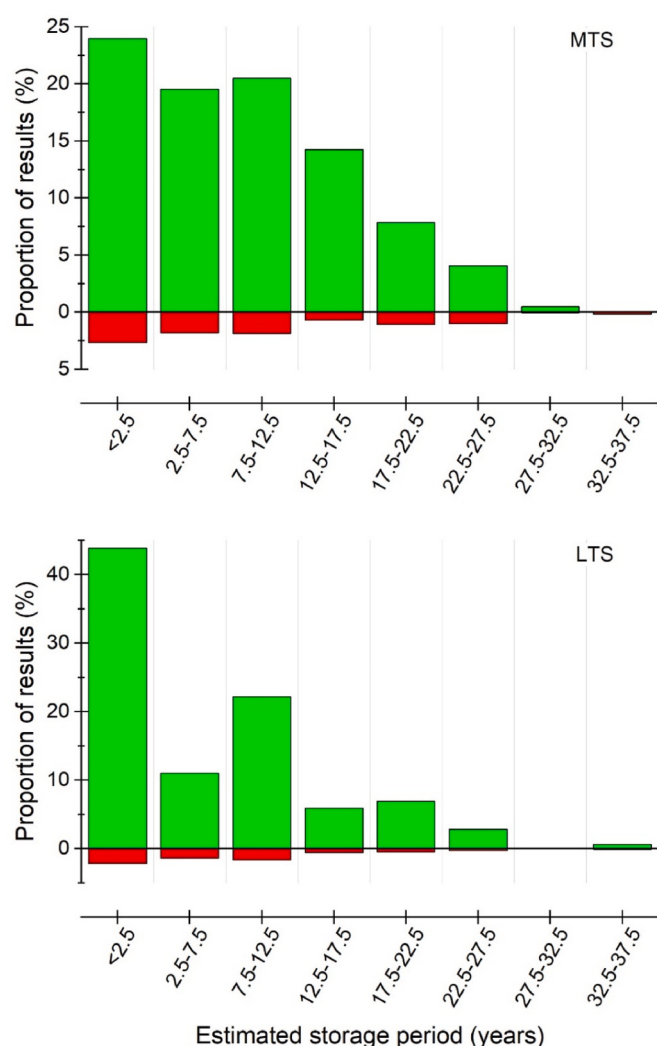


Fig. 2. The distribution of the historical viability test results for seeds of cultivated crops and forages over different periods of storage from seven CGIAR genebanks in medium-term storage (MTS) or long-term storage (LTS). Green columns show the proportion of results >85% and red columns the proportion ≤85%. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Loss of data may have occurred during handovers in management, when upgrading data management systems, from loss of original paper records, or when transferring data from paper records to a database. Some of the more obvious data errors we observed, such as dates that were not in sequence with seed lot processing steps, are now unlikely as more processes are automated and/or involve direct digital data entry (Hay and Sershen, 2021).

Although levels of data and quality of processes have generally increased over time, none of the databases included all the fields that are required to be able to derive reliable estimates of seed longevity. For example, most genebanks have not recorded the date seed lots were placed into storage; it was necessary to estimate this date based on harvest season or processing date. It would also be better to record all the viability data, including viability, sample size, date of sampling (removal from storage), and the germination test (and dormancy-breaking, if any) conditions used. Only a single date was recorded for viability tests, which may have been the date seeds were set to germinate (typically one day after removal from storage to allow seed containers to equilibrate to room temperature) or the date of scoring. Hence, the 'storage periods' that we calculated to assess seed longevity are

estimates with a number of likely, albeit small, errors relative to the total length of storage. Ellis et al. (2018, 2019) solved a similar problem by only quantifying storage periods in integers of years.

Some genebanks conserve very many species and considerable diversity within each species (Tables 2–4). Even if they tend to use a standard germination protocol for a particular species, that protocol may not be optimal for all accessions within a species (e.g. Salazar et al., 2020). It is important that information on dormancy-breaking treatments and test conditions, including the number of seeds tested, can be accessed, either within the database, for each viability test, or by referring to the standard operating procedure that was current at the time of testing. The number of seeds tested is important if it is necessary to make statistical tests (Hay et al., 2014), although such analyses are not common in daily genebank management. However, it is somewhat alarming that critical management decisions have potentially been made based on testing samples as small as 20 seeds in some cases (Table 1; Box 1). As in many aspects of genebank management, there is a balancing act: using more seeds for each viability monitoring test to improve accuracy versus increasing consumption of seeds in viability monitoring and hence, in the longer term, the rate of regeneration with the knock-on consequences of that for the conservation of genetic diversity.

This study has also shown that the data recorded by genebanks are not as valuable as expected in terms of its utility to better understand and quantify seed longevity in long-term genebank storage conditions. For most species × genebank × storage environment, the aggregated data did not reflect the expected sigmoidal survival curve (Hay et al., 2014). There are a number of reasons for this:

- (1) Most genebanks have a viability threshold of 85% for crop accessions and a lower threshold for wild species (FAO, 2014). It is not expected that a seed lot will continue to be monitored once viability declines below the threshold (Fig. 1; Whitehouse et al., 2020) because the original seed lot is replaced by a new seed lot. This results in a scatter graph of viability versus storage period with a high density of observations >85% and few <85%.
- (2) Germination test results are not always reliable. This may be in relation to overcoming dormancy, resulting in, for example, as seen for *Vigna subterranea* and *V. unguiculata* (Fig. A2), 0–100% viability regardless of length of time in storage, or in an increase in germination the longer the seeds had been in storage, as seen for *Phaseolus vulgaris* (see below). The latter pattern, ascribed to loss of hardseededness, was also shown in 12 genera of Fabaceae in the ILRI genebank (Ellis et al., 2018). Most genebanks will use the dormancy breaking treatments and germination conditions set by the International Seed Testing Association (ISTA, 2021), if they are available for the species concerned. However, the germination protocols may not be suitable for all the diverse material in a genebank (Ellis et al., 1985a, b). Furthermore, the reliability of germination results may be compromised if they are conducted by staff who have not been sufficiently trained in conducting and evaluating germination tests.
- (3) Human errors and biases, and/or apparent selective recording of data. There were a number of instances where the data suggested something that is not possible based on biological and statistical principles.

The viability data should be used to drive decisions about which accessions are available for distribution or require regeneration due to declining viability. However, there is evidence to suggest that viability data has not always been used to manage collections. For example, cases where seed lots were placed into long-term storage even if initial viability was below the threshold, or where seed lots continued to be stored and tested after viability had declined below the threshold. In relation to the former, best practice would be to validate the test result, regenerate again if necessary and then discard the low viability seed lot.

Box 1**Number of seeds to use for a viability monitoring test**

Germination tests are conducted on samples of seeds drawn at random from the main bulk (or perhaps, as pre-prepared samples in their own packet) and their accuracy depends on the number of seeds tested (Ellis et al., 1985a). Official seed testing of commercial seed lots use samples of 400 seeds each (ISTA, 2021), but this is too many for most genebanks. According to FAO (2014), “Sample sizes for viability monitoring will inevitably be dependent upon the size of the accession but should be maximized to achieve statistical certainty. However, the sample size should be minimized to avoid wasting seed.” The ‘statistical certainty’ relates to how close the result is to the actual viability. The relationship between the probability of a test on a sample of seeds passing a threshold value, such as that at which regeneration is required, and the true value of accession viability is called the Operating Characteristic. The curve is steeper, and so decisions subject to less error, the greater the number of seeds sampled and tested (Ellis et al., 1985a). Appendix Fig. B4 provides the Operating Characteristic for the AfricaRice monitoring tests (100-seed samples; 85% viability regeneration standard) presented in Appendix Figs. B2 – B3, which shows that most of the variation detected during long-term storage was due to random sampling error.

In the study reported in this paper, the number of seeds used for a viability test ranged from 20 to 200 (Tables 1–4). Genebanks do not routinely do any statistical tests on viability data, with the exception perhaps of tolerance testing (to ensure that the variation in results from a sample of seeds sown across multiple sowing units (e.g. Petri dishes) could be obtained through random sampling error), or express any degree of certainty regarding the viability result. Nonetheless, the value of a sample size as low as 20, is somewhat questionable. For example, if the real viability of a seed lot is 90% and 20 seeds are sown, the probability of 17 seeds (85%) or fewer germinating, is 0.32 i.e. in approximately a third of tests, a decision to rejuvenate might be taken unnecessarily. Conversely, if 100 seeds are tested, that probability reduces to 0.07. These calculations are based on the binomial distribution and it does not matter whether the sample is subdivided between sowing units. However, sowing a sample across experimental units reduces the risk of losing a result for example, if one unit is dropped and therefore has to be discarded. The other reason for germinating seeds across multiple units is that the seeds can be spaced out, which may be better for germination and for scoring the germinated seeds.

To solve the dilemma of too many valuable seeds being used in genebank viability monitoring tests, sequential testing procedures were proposed (Ellis et al., 1985a), but have not been widely adopted. CIAT has been using sequential testing for the bean collection for several years, but were not able to store the sequential test results within their database in a meaningful way.

Continuing to actively manage a seed lot with low initial viability as the only representative sample of an accession, especially if the accession is genetically-heterogeneous, raises questions about the effective conservation (and distribution) of genotypes. Of course, in species where it is difficult to produce seeds and/or seed lots with high initial quality, or for which reliable testing protocols are not available, it may be necessary to accept a lower viability threshold (together with the inherent implications that has in relation to what is being conserved).

We were also able to identify accessions that were represented by multiple seed lots within a collection $\times$ storage environment, despite seemingly high levels of viability (see examples in Appendix Figs B5, B6). A range of reasons were given for this including (and not necessarily independent of each other) that: (i) it was not possible to collect ‘sufficient’ quantity of seeds at one harvest; (ii) some seed lots did not meet phytosanitary requirements; and (iii) there was a question over the identity of seed lots. Actively maintaining multiple seed lots of the same accession within a collection (i.e. within the active or base collections) should be avoided, unless for a justified reason, since it affects costs and risks of error, as well as taking up physical space in cold storage. In some cases, it may be necessary to set up a strategy to remove multiple seed lots from a collection, once they have accumulated (Hay and Whitehouse, 2017; Box 2).

4.2. Findings related to seed longevity

Different studies have taken different approaches to analysing genebank viability data, perhaps depending on the main driver for such studies: to better manage a genebank or to better understand seed longevity. Genebanks are concerned about how many seed lots need to be regenerated in any particular year. Hence, modelling the proportion of seed lots that are above/below the viability threshold as storage period increases (Fig. 2; Agacka et al., 2013, 2014; Hay et al., 2015; Yamasaki et al., 2020) could be considered more meaningful than attempting to model the behaviour of individual seed lots. Hay et al. (2015) were able to identify seed production years or countries of origins that resulted in more seed lots with poorer storability. Yamasaki et al. (2020) used time-to-event analysis to model the proportion of seed

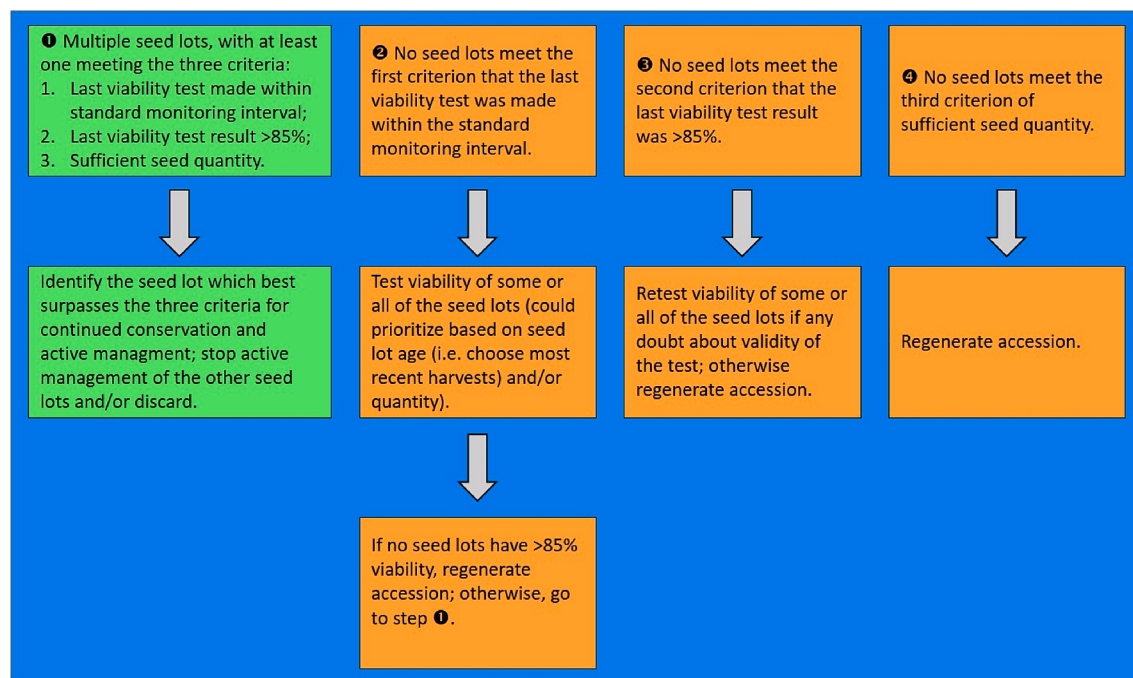
lots reaching the regeneration threshold and thence calculated a recommended monitoring interval as one-third of the ‘median survival time’, based on fitting a Weibull distribution. These monitoring intervals, however, are specific to those crops and those storage conditions. The results are also difficult to compare with other studies where a more defined parameter is used as a measure of longevity, most commonly the period for viability to fall to 50% (p_{50}).

In most of the seed longevity literature, probit analysis has been used to estimate p_{50} (e.g. Probert et al., 2009; Merritt et al., 2014; Desheva, 2016). This has the advantage of accounting for the binomial error distribution of germination data, i.e., that each individual seed will either germinate or not germinate and the error associated with each germination result depends on the number of seeds tested (Hay et al., 2014). However, the estimated parameters are less reliable if there are not many observations for each seed lot between ca. 85 and 15%. Both of these constraints are typical of genebank data (assuming dormancy is not an issue). In the data set considered in our study, for many seed lots, only one or two viability results were available. Ellis et al. (2018, 2019) applied probit analysis to the viability monitoring data for seed lots of forage species held in medium-term storage at the International Livestock Research Institute (ILRI), restricting the analysis to those seed lots with at least two results. For many of these forage accessions, viability monitoring had continued even though viability was apparently low (Ellis et al., 2018, 2019), with the purpose of gaining more information on seed longevity, even though this is not efficient from a genebank management perspective (Whitehouse et al., 2020). It was possible to constrain the data to a survival curve with a common slope within some of the species/genera in these studies and hence, Ellis et al. (2018, 2019) were able to derive estimates of longevity. However, in some cases, the trend in viability showed a positive response over time, due to loss of dormancy and/or improvement in testing procedures. This was also observed for seeds of *Phaseolus vulgaris* in our study. In the case of the rice collections at IRRI, applying probit analysis to the data for seed lots with at least four or at least five sequential observations, it was not possible to constrain the slope to a common estimate within either of the two species. This may have been because many of these seed lots were not yet showing significant decline in viability and/or that the trends

Box 2**Minimizing the active management of more than one seed lot per accession x collection**

Many genebanks use standard plot sizes and/or number of plants when regenerating seeds of a particular crop/species for genebank storage. These standards may have been determined based on 'typical' accessions or be set due to logistical constraints. However, if an accession produces fewer seeds than the 'typical' accession, or the regeneration is impacted by biotic or abiotic stress, it may be inevitable that the resulting seed lot is inadequate to meet thresholds for conservation, but is nonetheless processed for storage to avoid 'waste'. The accession may also be regenerated again resulting in the active management of more than one seed lot per accession. One solution would be to use more plants for regeneration. This may mean that it is not possible to regenerate as many accessions in the available area of land and that consequently, some accessions become unavailable. However, in the longer term, there should be a more efficient strategy in terms of preventing the buildup of multiple seed lots per accession that each require independent active management. It is also more effective in terms of long-term conservation of biodiversity.

Where multiple seed lots of an accession have accrued, a rationalization strategy may be required to remove seed lots from active management. Bulking seed lots from different seasons is rarely acceptable, as the subpopulations of seeds from different harvests will be losing percentage viability at different rates. An outline for making rationalization decisions is shown below.



across diverse accessions were just too variable. Plotting the initial and current germination data against storage period depending on the initial result for rice seeds in long-term storage at AfricaRice, clearly illustrates the considerable variation in the apparent change in viability between two observations (Appendix Fig. B2). Much of this variation will be due to random sampling error; however, Whitehouse et al. (2018) reported that the slope of the survival curve could vary by up to 7-fold across different seed lots of *O. sativa* stored under identical controlled experimental storage conditions, and Lee et al. (2019) were able to identify QTLs associated with the slope of survival curves within a diverse panel of Indica Group *O. sativa* accessions. Thus, in rice at least, there does appear to be considerable variation in survival curve slopes for seeds stored in the same environment; this makes it harder to set appropriate monitoring intervals for an accession, based on the initial germination result and predictions using the viability equation.

Since we were not able to broadly apply probit analysis across the data set, to estimate potential longevity, we calculated the maximum estimated storage periods with ≥ 95 viability (Tables 1–4). These estimates suggest that, as a whole, viability was largely maintained during storage in these genebanks, at least up to the maximum periods covered here. Evidence that high viability is being maintained for 3–4 decades is encouraging, but we cannot conclude how much longer high viability

might be maintained, or how quickly it might be lost once decline is first detected. Van Treuren et al. (2018) reported unexpectedly rapid loss of viability over six years for wheat seed lots that had been in medium-term storage at 4 °C for 25–33 years. Thus, while it is reasonable to extend monitoring intervals for seed lots that are less than a few decades old, to reduce workload and consumption of seeds for germination testing, monitoring intervals should not change or should even be shortened for the oldest seed lots, to minimize the risk of losing genotypes. It is important to emphasize with these 'storability periods', that we have not done any averaging across seed lots within a collection $\times$ storage environment. A maximum estimated storability period may not be achievable for all seed lots of a particular species stored under the same conditions since longevity is influenced by many different factors, not least those that impact the 'initial' quality of a seed lot when it is first placed into storage, as discussed for example by Probert et al. (2007) and Kameswara Rao et al. (2017), however it perhaps does give something to 'aim for'. Furthermore, the results are comparable with, or even provide better seed storage survival than, historical viability data published by other genebanks. For example, Walters et al. (2005) reported mean initial and final (after 34 years) germination percentages of 780 accessions of *Arachis hypogaea* of 89 and 6% (stored at -18 °C for 15–19 years, and before that, at 5 °C); this compares with a mean germination

of 97% for *A. hypogaea* seeds after 22.8–24.7 years in LTS at ICRISAT. Similarly, Walters et al. (2005) reported mean initial and final (after 28.2 years) germination percentages of 84 accessions of *Oryza sativa* of 90 and 54%; for the data from IRRI, there was a mean germination of 84% after 27.7–32.5 years for seeds in LTS (234 seed lots) and of 94% after 27.5–31.9 years in MTS (2048 seed lots). This seemingly greater longevity in MTS cf. LTS (Appendix Fig. A1) emphasizes the importance of continuing to test a selected subset of older samples after their viability has declined below the viability threshold, to gather additional data over the part of the seed survival curve where viability is between 85 and 15% and hence verify longevity periods.

4.3. Thinking to the future

For too long, managing a genebank has been about negotiating trade-offs among a range of factors that limit the ability to manage collections using optimum procedures; this study has highlighted how this has limited consistent data gathering to support decision-making and long-term conservation. Since the initiation of the Genebank CGIAR Research Program in 2012, a coordinated effort has established common performance targets and quality management systems for all CGIAR genebanks. More recently, the Crop Trust has conducted audits to review and validate genebank standard operating procedures. Indeed, several of the issues highlighted by our study were identified as areas of concern in some of those audits. Further, the Seed Quality Management (SQM) project (Hay et al., 2021), of which this study is part, has endeavoured to improve operations through conservation research, for example in relation to harvesting and drying procedures (e.g. Jones et al., 2020) and germination protocols (Salazar et al., 2020). The need for more basic research on germination protocols is exemplified by the historical viability data for the forage collections at CIAT and the tree collection at World Agroforestry (Tables 2–4).

Seed viability is the basic parameter by which the staff of a seed genebank know they are fulfilling their role. While viability can be maintained year on year through planting out failing seed lots or asking for new seeds from a provider, genebanks provide the storage conditions that farmers, breeders or those involved in *in situ* conservation cannot. It therefore behoves those with a responsibility for long-term conservation to find the best possible ways to maintain the collections under their management in the best possible conditions for the longest possible time, incurring the least possible costs. This study highlights the following needs for the management of seed collections:

- Adequate and consistent data and effective data management systems designed for long-term data gathering, storage and analysis;
- Tools to facilitate the oversight and forward planning of collection management at a higher level, allowing managers to analyse the age of seed lots, patterns of their viability in storage, user demand and other factors that should feed into both annual planning of genebank activities such as regeneration but also the adaption of processes and monitoring regimes according to evident needs;
- Documented and regularly audited and reviewed processes that capture any necessary temporary deviations and facilitate staff succession;
- A culture, capacity and community that encourages active critical review and refinement of genebank operations and specialised research tailored to specific crops, collections and circumstances.

Most importantly, these actions should continue for as long as CGIAR

is responsible for managing the in-trust collections. CGIAR is reforming the governance of the 15 research institutes in its consortium, encouraging the Centers to align and share capacity, facilities and staff. While the mechanisms of the reform will, no doubt, bring challenges over the next few years, there is an opportunity for this important international community of genebanks to build further on previous collaboration and bring their combined capacity to drive innovation and improvement in germplasm conservation.

5. Conclusions

CGIAR genebanks have amassed a large quantity of viability monitoring data over the last few decades. However, although it is clear that high viability can be maintained over 30 years or longer for some seed lots of many of the crops they conserve, it is nonetheless difficult to make reliable estimates of seed lot longevity that would be meaningful in terms of revising monitoring intervals and predicting future levels of accession rejuvenation. Overall, the data suggest that a 'steady state' of operations (Koo et al., 2003) has not been reached, with genebanks having faced considerable constraints; management focus has varied depending on the uppermost priorities at different times. We emphasize, that we do not believe that these constraints have resulted in loss of unique germplasm, rather that the efficiencies that could have been made as a result of seed longevity analyses such as these are restricted by the paucity of reliable historical data. Despite this, seed genebanking, and specifically the overall framework of how collections should be stored and managed, remains the most effective way of ensuring the availability of viable crop germplasm for future generations as a means of contributing to global food security.

Funding

This study was undertaken as part of the Seed Quality Management (SQM) activity of the CGIAR Genebank Platform. We would like to thank all funders who have supported the CGIAR Genebanks over the last 4-5 decades. In particular, SQM received financial support in 2020 from the German Federal Ministry for Economic Cooperation and Development (BMZ) commissioned and administered through the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) Fund for International Agricultural Research (FIA), grant number: 81235816.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Fiona R. Hay reports financial support was provided by CGIAR. Charlotte Lusty reports financial support was provided by CGIAR. Fiona R. Hay reports a relationship with CGIAR that includes: consulting or advisory and travel reimbursement. Charlotte Lusty reports a relationship with CGIAR that includes: consulting or advisory and travel reimbursement. Corresponding author previously worked at a CGIAR center. Coauthor manages the genebank platform but is employed by the Crop Trust (CL).

Acknowledgements

We acknowledge all past and present staff at the respective genebanks who are/were involved in viability testing and recording/storage of viability data.

Appendix A. Plots of germination test results vs. estimated storage period based on the historical viability monitoring data for the indicated seed genebank collections

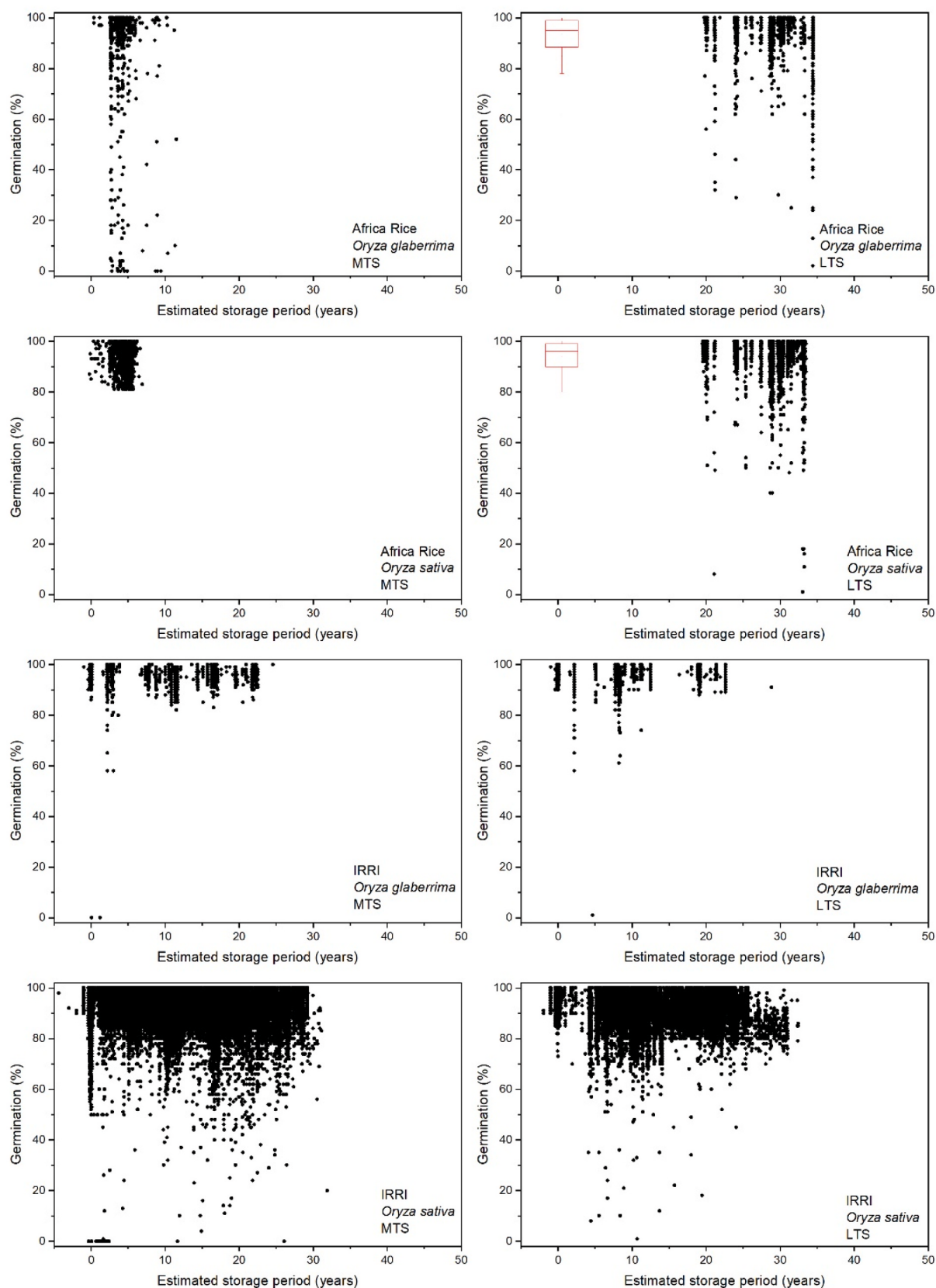


Fig. A1. Plots of germination vs. estimated storage period based on the historical viability monitoring data for *Oryza glaberrima* and *O. sativa* seeds in medium-term storage (MTS) and long-term storage (LTS) at AfricaRice and the International Rice Research Institute (IRRI). The box plots shown for seeds in LTS at AfricaRice show the initial results, with the box spanning the 25th and 75th percentiles, the median represented by the line within the box, and the whiskers extending to the 5th and 95th percentiles.

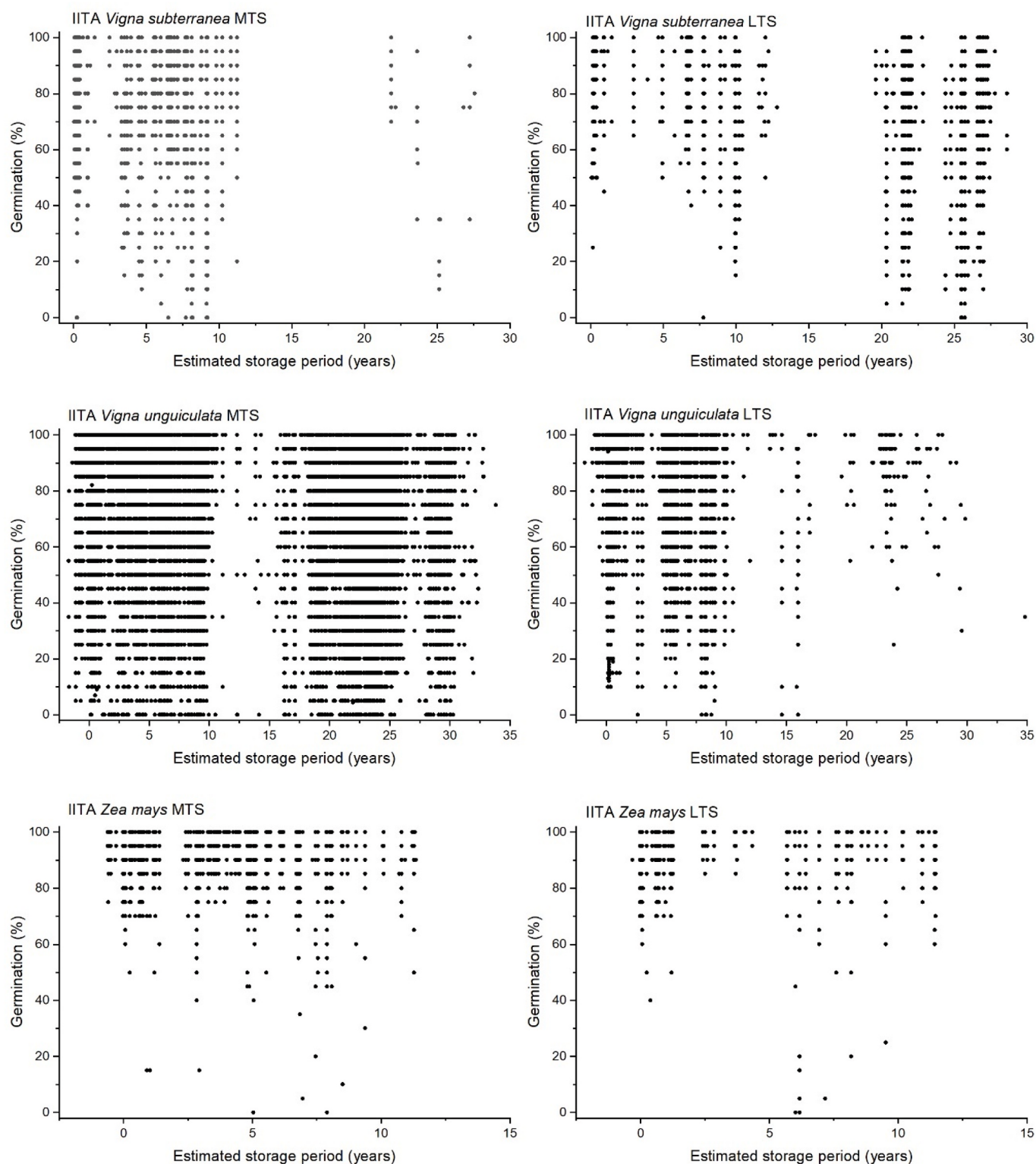


Fig. A2. Plots of germination vs. estimated storage period based on the historical viability monitoring data for Bambara groundnut (*Vigna subterranea*), cowpea (*V. unguiculata*) and maize (*Zea mays* L.) seeds in medium-term storage (MTS) and long-term storage (LTS) at the International Institute of Tropical Agriculture (IITA).

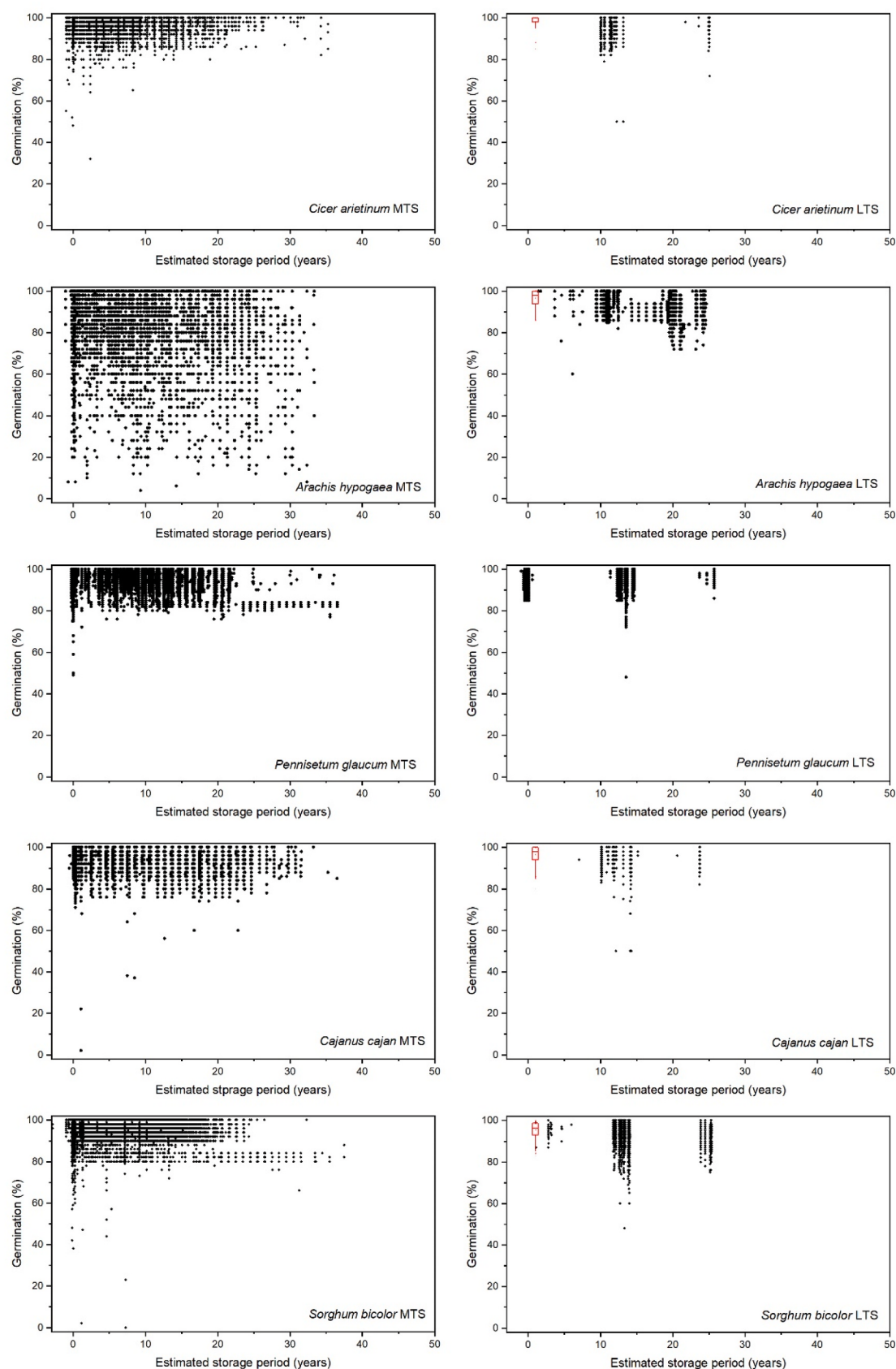


Fig. A3. Plots of germination vs. estimated storage period based on the historical viability monitoring data for chickpea (*Cicer arietinum* L.), groundnut (*Arachis hypogaea* L.), pearl millet (*Pennisetum glaucum* L.) R. Br. (synonym of *Cenchrus americanus* (L.) Morrone), pigeon pea (*Cajanus cajan* (L.) Huth) and sorghum (*Sorghum bicolor* L. Moench) seeds in medium-term storage (MTS) and long-term storage (LTS) at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

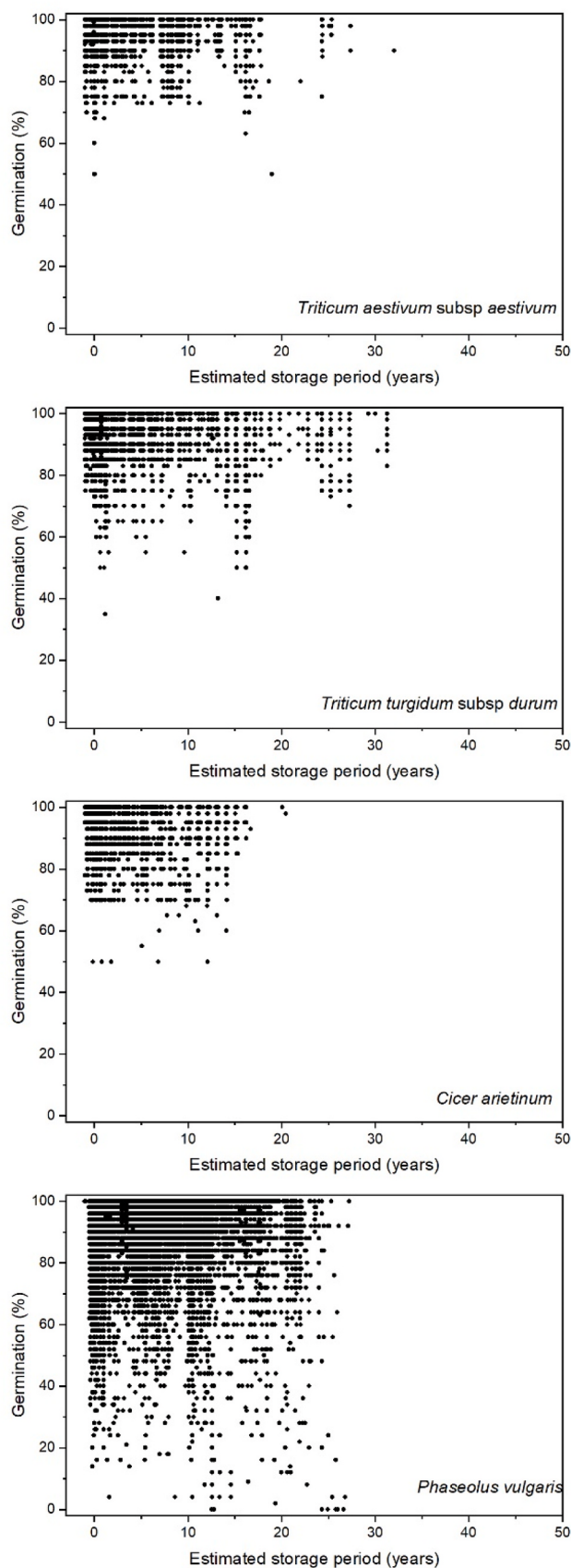


Fig. A4. Plots of germination vs. estimated storage period based on the historical viability monitoring data for wheat (*Triticum aestivum* L. subsp. *aestivum* and *T. turgidum* L. subsp. *durum* (Desf.) van Slageren) and chickpea (*Cicer arietinum* L.) in long-term storage (LTS) at the International Center for Agricultural Research in the Dry Areas (ICARDA), and bean (*Phaseolus vulgaris* L.) in LTS at the International Center for Tropical Agriculture (CIAT).

Appendix B Plots showing some of the detail of the historical viability monitoring data from some of the genebanks

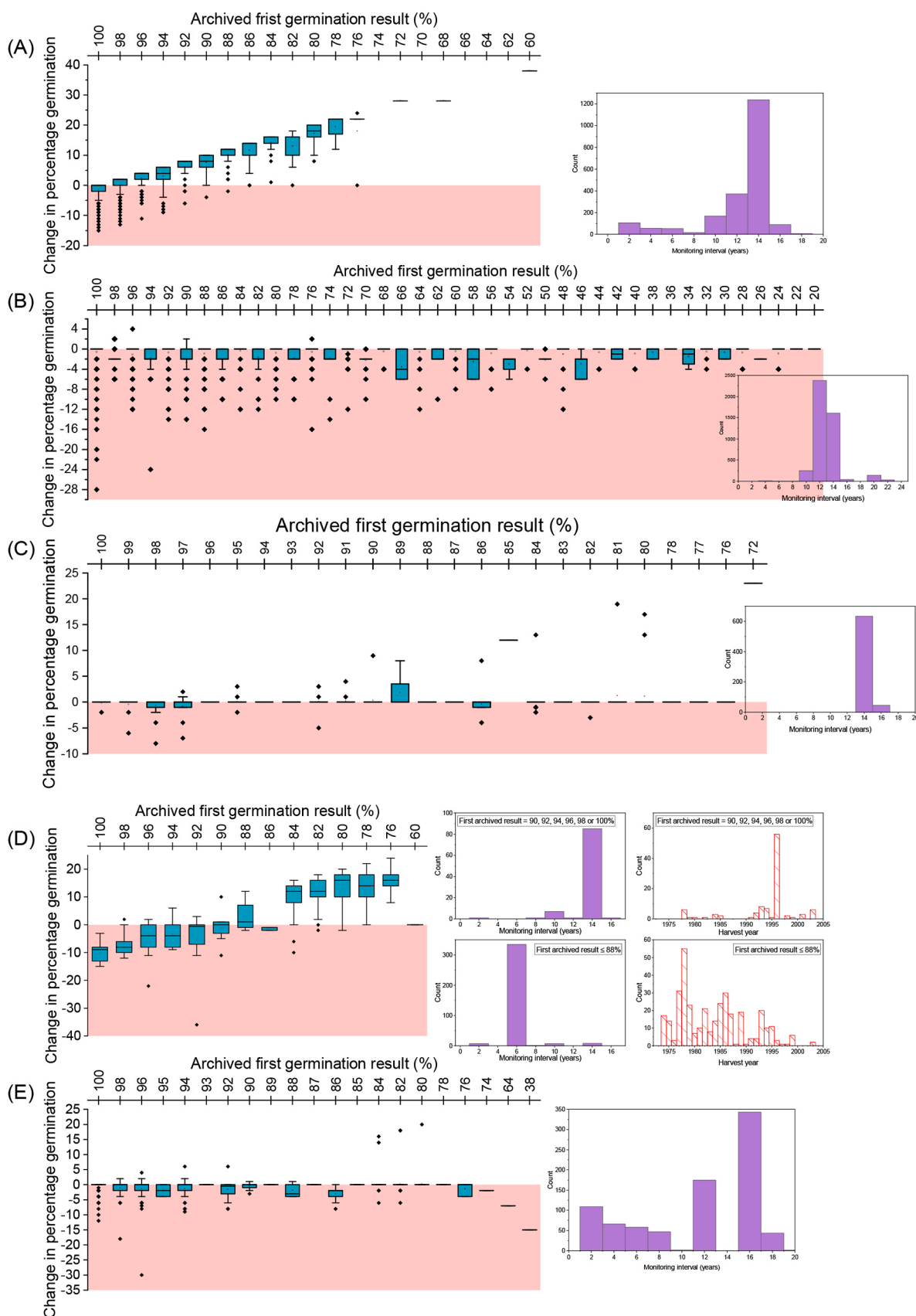


Fig. B1. Change in percentage germination between the first and second germination results for (A) chickpea, (B) groundnut, (C) pearl millet, (D) pigeon pea and (E) sorghum seed lots in medium-term storage at ICRISAT. Also shown are the length of the monitoring interval and, for pigeon pea, the length of the monitoring interval and harvest year depending on the result of the first available germination result.

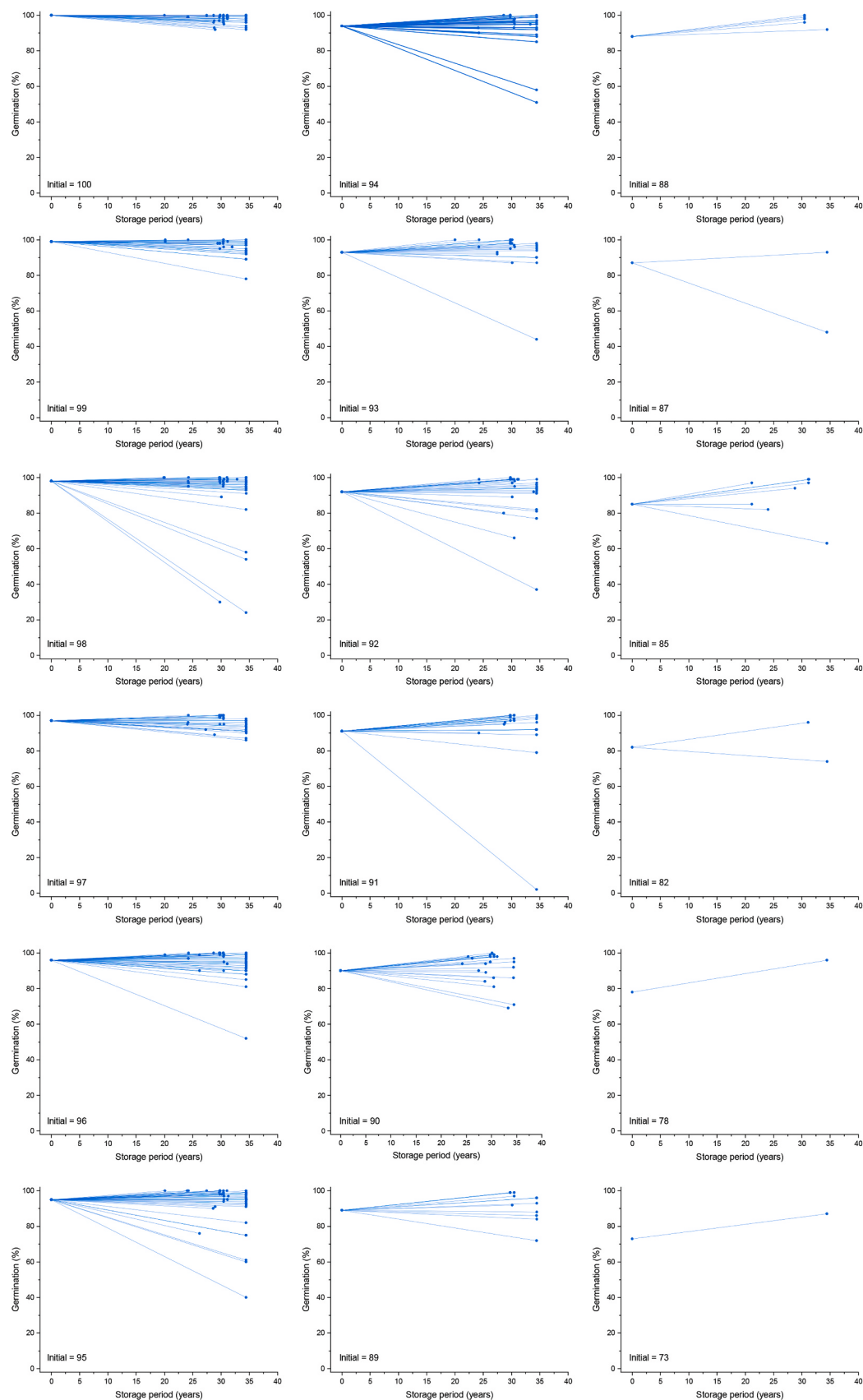


Fig. B2. Examples of the observed change in viability between the initial and most recent viability results for accessions of *Oryza glaberrima* in long-term storage at AfricaRice. The graphs are sorted according to the initial viability result, starting with all those accessions originally providing 100% in the top left, then in order of % points downwards and then right.

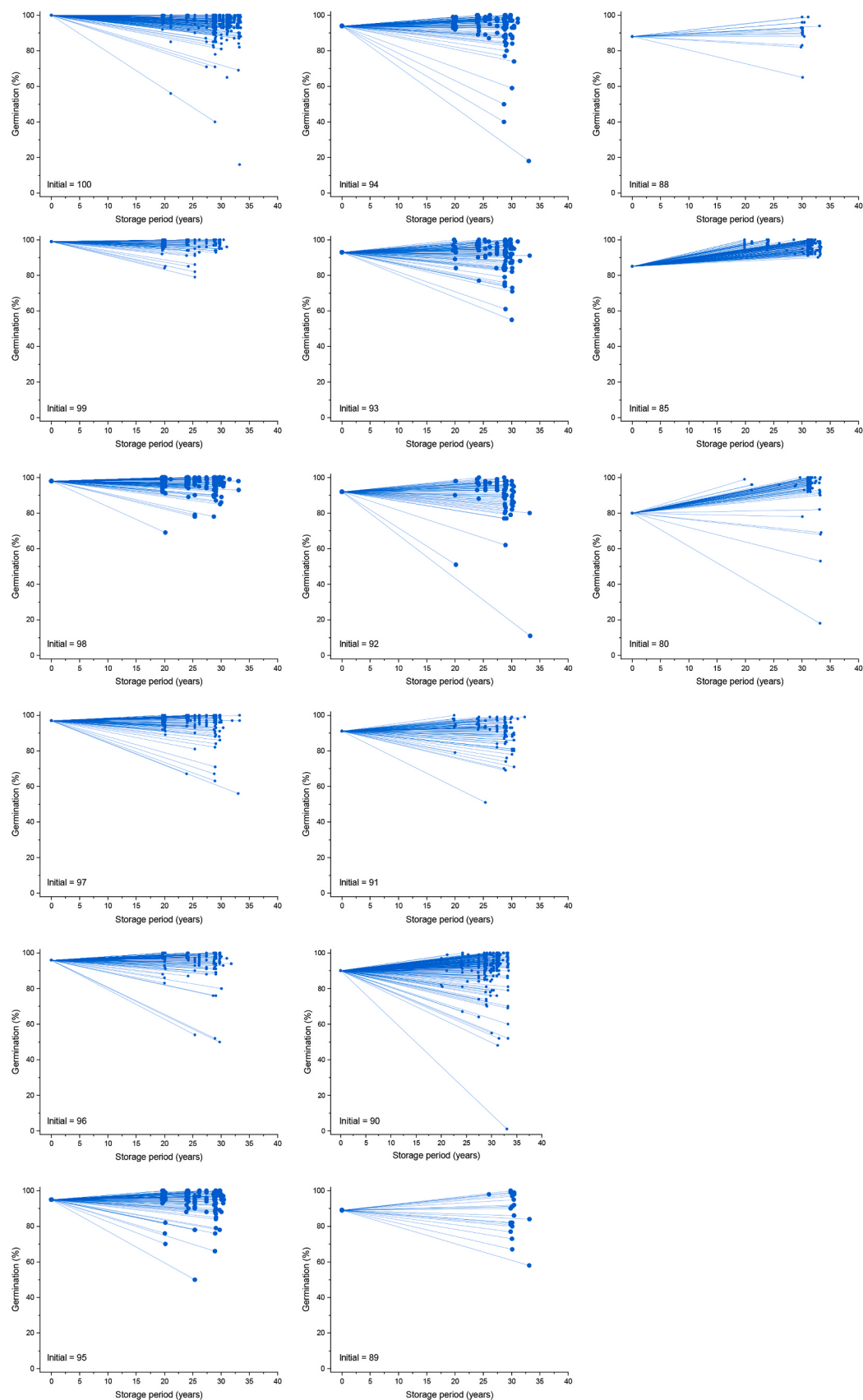


Fig. B3. Examples of the observed change in viability between the initial and most recent viability results for accessions of *Oryza sativa* in long-term storage at AfricaRice. The graphs are sorted according to the initial viability result, starting with all those accessions originally providing 100% in the top left, then in order of % points downwards and then right.

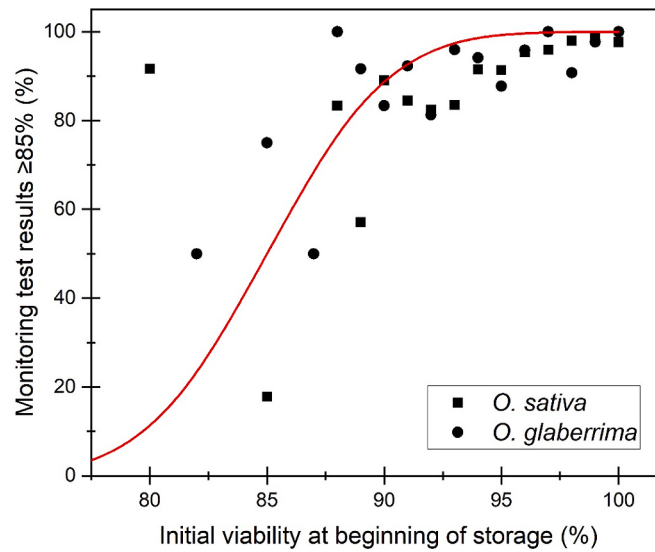


Fig. B4. Comparison of the proportion of monitoring tests which provided $\geq 85\%$ germination for accessions of *Oryza glaberrima* (●) and *Oryza sativa* (■) maintained in AfricaRice's long-term seed store for between 7124 and 12,225 days in relation to the germination of each accession upon entry into store (from Appendix Figs B2 – B3) with the Operating Characteristic (solid curve) for 100-seed tests with an 85% test boundary expected from random sampling error alone. If one assumes that the initial viability reported in Appendix Figs. B2 – B3 is the true value and that no seed deterioration occurred during storage then the symbols would be located on the curve shown. Hence there is evidence of some loss in viability in this store over two to three decades, but most of the variation detected within these accessions during storage is explained by random sampling error.

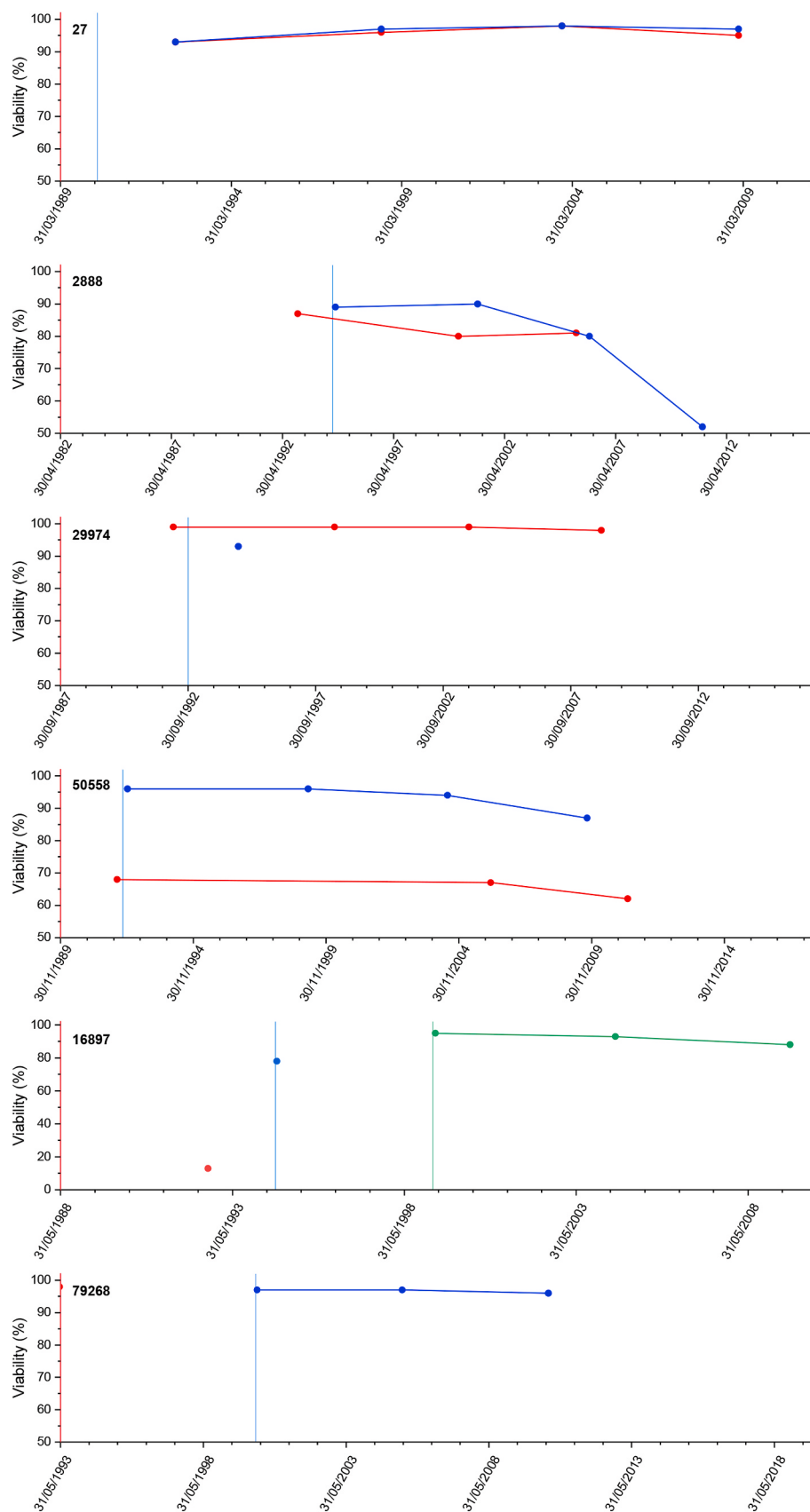


Fig. B5. Examples of six IRRI accessions (number in top left of each graph) of *Oryza sativa* for which the data set included results for more than one seed lot. The data shown are for seed lots in MTS. The coloured vertical lines indicate when seed lots were produced and the symbol or symbol-line plots with the same colour show the viability result(s) for that seed lot.

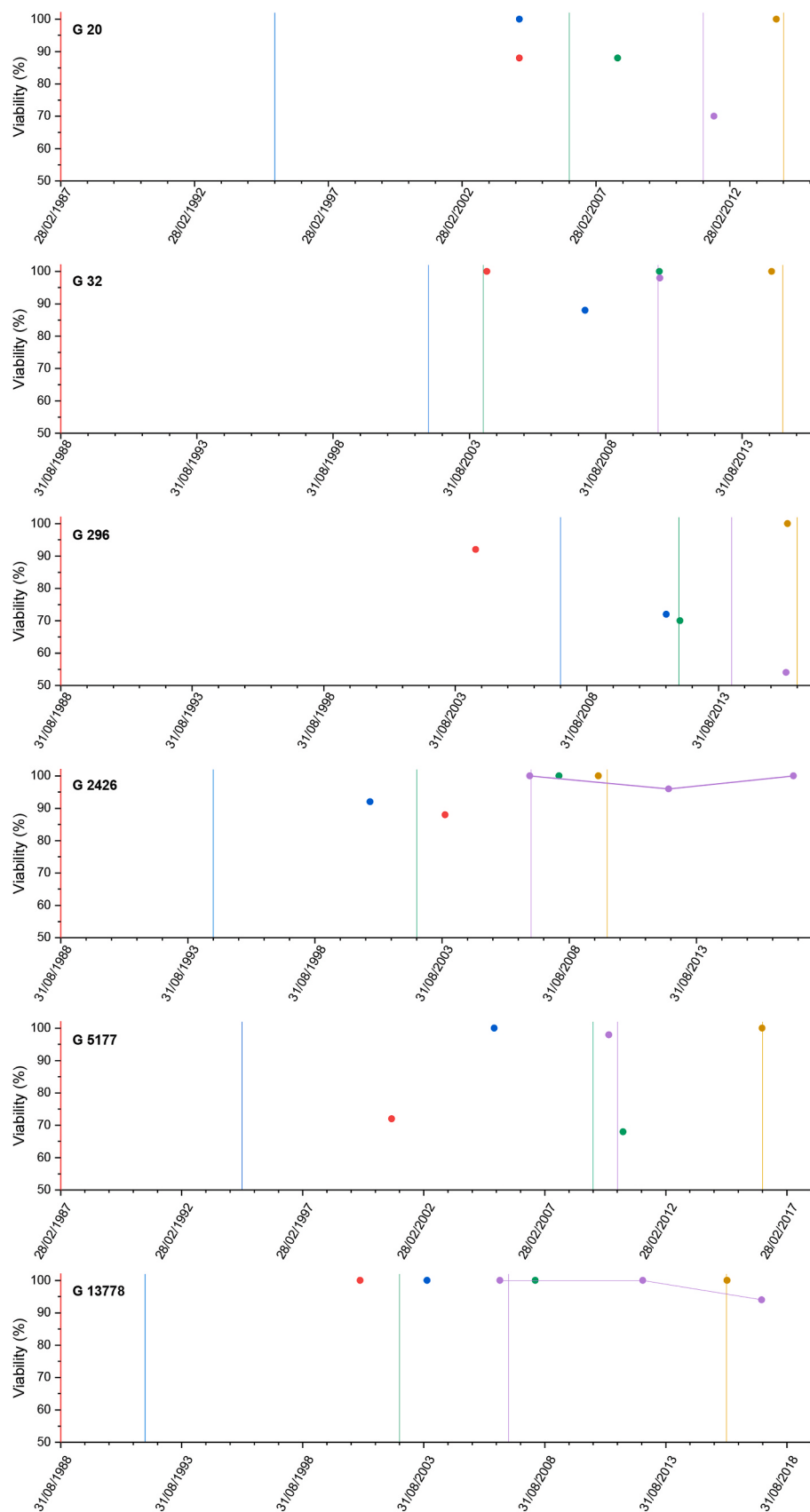


Fig. B6. Examples of six CIAT accessions (number in top left of each graph) of *Phaseolus vulgaris* for which the data set included results for more than one seed lot. The data shown are for seed lots in LTS. The coloured vertical lines indicate when seed lots were produced and the symbol or symbol-line plots with the same colour show the viability result(s) for that seed lot.

References

- Agacka, M., Depta, A., Börner, M., Doroszewska, T., Hay, F.R., Börner, A., 2013. Viability of *Nicotiana* spp. seeds stored under ambient temperature. *Seed Sci. Technol.* 41, 474–478. <https://doi.org/10.15258/sst.2013.41.3.15>.
- Agacka, M., Laskowska, D., Doroszewska, T., Hay, F.R., Börner, A., 2014. Longevity of *Nicotiana* seeds conserved at low temperatures in *ex situ* genebanks. *Seed Sci. Technol.* 42, 355–362. <https://doi.org/10.15258/sst.2014.42.3.05>.
- Colville, L., Pritchard, H.W., 2019. Seed life span and food security. *New Phytol.* 224, 557–562. <https://doi.org/10.1111/nph.16006>.
- Cromarty, A.S., Ellis, R.H., Roberts, E.H., 1982. The Design of Seed Storage Facilities for Genetic Conservation. International Board for Plant Genetic Resources, Rome, Italy.
- Desheva, G., 2016. The longevity of crop seeds stored under long-term condition in the national gene bank of Bulgaria. *Agric. For.* 62, 90–100.
- Ellis, R.H., Roberts, E.H., 1980. Improved equations for the prediction of seed longevity. *Ann. Bot.* 45, 13–30. <https://doi.org/10.1093/oxfordjournals.aob.a085797>.
- Ellis, R.H., Hong, T.D., Roberts, E.H., 1985a. Handbook of Seed Technology for Genebanks. Volume I. Principles and Methodology. International Board for Plant Genetic Resources, Rome, Italy.
- Ellis, R.H., Hong, T.D., Roberts, E.H., 1985b. Handbook of Seed Technology for Genebanks. Volume II. Compendium of Specific Germination Information and Test Recommendations. International Board for Plant Genetic Resources, Rome, Italy.
- Ellis, R.H., Nasehzadeh, M., Hanson, J., Woldemariam, Y., 2018. Medium-term seed storage of 50 genera of forage legumes and evidence-based genebank monitoring intervals. *Genet. Resour. Crop Evol.* 65, 607–623. <https://doi.org/10.1007/s10722-017-0558-5>.
- Ellis, R.H., Nasehzadeh, M., Hanson, J., Ndiwa, N., Woldemariam, Y., 2019. Medium-term seed storage of diverse genera of forage grasses, evidence-based genebank monitoring intervals, and regeneration standards. *Genet. Resour. Crop Evol.* 66, 723–734. <https://doi.org/10.1007/s10722-019-00748-y>.
- FAO, 2010. The Second Report on the State of the World's Plant Genetic Resources for Food and Agriculture. FAO, Rome.
- FAO, 2014. Genebank Standards for Plant Genetic Resources for Food and Agriculture, Rev. FAO, Rome.
- Fu, Y.-B., 2017. The vulnerability of plant genetic resources conserved *ex situ*. *Crop Sci.* 57, 2314–2328. <https://doi.org/10.2135/cropsci2017.01.0014>.
- Galluzzi, G., Halewood, M., López Noriega, I., Vernooij, R., 2016. Twenty-five years of international exchanges of plant genetic resources facilitated by the CGIAR genebanks: a case study on global interdependence. *Biodivers. Conserv.* 25, 1421–1446. <https://doi.org/10.1007/s10531-016-1109-7>.
- Griesbach, R.J., 2013. 150 Years of Research at the United States Department of Agriculture: Plant Introduction and Breeding. U.S. Department of Agriculture, Agricultural Research Service, Washington, DC.
- Hanson, J., Williams, J.T., Freund, R., 1984. Institutes Conserving Crop Germplasm: the IBPGR Global Network of Genebanks. International Board for Plant Genetic Resources, Rome, Italy.
- Hay, F.R., Serksen, 2021. New technologies to improve the *ex situ* conservation of plant genetic resources. In: Dulloo, E. (Ed.), *Plant Genetic Resources: A Review of Current Research and Future Needs*. Burleigh Dodds Science Publishing, Cambridge, UK.
- Hay, F.R., Whitehouse, K.J., 2017. Rethinking the approach to viability monitoring in seed genebanks. *Conserv. Physiol.* 5, cox009. <https://doi.org/10.1093/conphys/cox009>.
- Hay, F.R., de Guzman, F., Ellis, D., Makahiya, H., Borromeo, T., Sackville Hamilton, N.R., 2013. Viability of *Oryza sativa* L. seeds stored under genebank conditions for up to 30 years. *Genet. Resour. Crop Evol.* 60, 275–296. <https://doi.org/10.1007/s10722-012-9833-7>.
- Hay, F.R., Mead, A., Bloomberg, M., 2014. Modelling seed germination in response to continuous variables: use and limitations of probit analysis and alternative approaches. *Seed Sci. Res.* 24, 165–186. <https://doi.org/10.1017/S096025851400021X>.
- Hay, F.R., de Guzman, F., Sackville Hamilton, N.R., 2015. Viability monitoring intervals for genebank samples of *Oryza sativa*. *Seed Sci. Technol.* 43, 218–237. <https://doi.org/10.15258/sst.2015.43.2.17>.
- Hay, F.R., Lee, J.-S., Lusty, C., 2021. The international genebanks of the CGIAR and seed quality management. *Seed Test. Int.* 161, 10–13.
- IBPGR, 1976. Report of IBPGR Working Group on Engineering, Design and Cost Aspects of Long-Term Seed Storage Facilities. International Board for Plant Genetic Resources, Rome, Italy.
- IBPGR, 1982. IBPGR Ad Hoc Advisory Committee on Seed Storage: Report of the First Meeting. International Board for Plant Genetic Resources, Rome, Italy.
- IBPGR, 1985. IBPGR Ad Hoc Advisory Committee on Seed Storage: Report of the Third Meeting. International Board for Plant Genetic Resources, Rome, Italy.
- ISTA, 2021. International Rules for Seed Testing. International Seed Testing Association, Bassersdorf, Switzerland.
- James, E., 1972. Organisation of the United States national seed storage laboratory. In: Roberts, E.H. (Ed.), *Viability of Seeds*. Chapman and Hall, London, pp. 397–404.
- Jones, T.M., Oyatomi, O.A., Aikinyele, B.O., Odiyi, A.C., Abberton, M.T., Hay, F.R., 2020. Effect of high temperature drying on seed longevity of Bambara groundnut (*Vigna subterranea*) accessions. *Seed Sci. Technol.* 48, 413–418. <https://doi.org/10.15258/sst.2020.48.3.10>.
- Kameswara Rao, N., Dulloo, M.E., Engels, J.M.M., 2017. A review of factors that influence the production of quality seed for long-term conservation in genebanks. *Genet. Resour. Crop Evol.* 64, 1061–1074. <https://doi.org/10.1007/s10722-016-0425-9>.
- Kaplan, J.K., 1998. Conserving the world's plants. *Agr. Res.* 46, 4–9.
- Kloppenborg, J.R., 1988. First the Seed: The Political Economy of Plant Biotechnology 1492–2000. Cambridge University Press, Cambridge.
- Koo, B., Pardey, P.G., Wright, B.D., 2003. The economic costs of conserving genetic resources at the CGIAR centres. *Agric. Econ.* 29, 287–297. [https://doi.org/10.1016/S0169-5150\(03\)00056-2](https://doi.org/10.1016/S0169-5150(03)00056-2).
- Lee, H.-S., Jeon, Y.-A., Lee, Y.-Y., Lee, S.-Y., Kim, Y.-G., 2013. Comparison of seed viability among 42 species stored in a genebank. *Korean J. Crop Sci.* 58, 432–438. <https://doi.org/10.7740/kjcs.2013.58.4.432>.
- Lee, J.-S., Velasco-Pulzalan, M., Pacleb, M., Valdez, R., Kretzschmar, T., McNally, K.L., Ismail, A.M., Sta Cruz, P.C., Sackville Hamilton, N.R., Hay, F.R., 2019. Variation in seed longevity among diverse Indica rice varieties. *Ann. Bot.* 124, 447–460. <https://doi.org/10.1093/aob/mcz093>.
- Loskutov, I.G., 2020. Vavilov Institute (VIR): historical aspects of international cooperation for plant genetic resources. *Genet. Resour. Crop Evol.* 67, 2237–2253. <https://doi.org/10.1007/s10722-020-00979-4>.
- Merritt, D.J., Martyn, A.J., Ainsley, P., Young, R.E., Seed, L.U., Thorpe, M., Hay, F.R., Commander, L.E., Shackelford, N., Offord, C.A., Dixon, K.W., Probert, R.J., 2014. A continental-scale study of seed lifespan in experimental storage examining seed, plant, and environmental traits associated with longevity. *Biodivers. Conserv.* 23, 1081–1104. <https://doi.org/10.1007/s10531-014-0641-6>.
- Probert, R.J., Adams, J., Coneybeer, J., Crawford, A., Hay, F., 2007. Seed quality for conservation is critically affected by pre-storage factors. *Aust. J. Bot.* 55, 326–335. <https://doi.org/10.1071/BT06046>.
- Probert, R.J., Daws, M.I., Hay, F.R., 2009. Ecological correlates of *ex situ* seed longevity: a comparative study on 195 species. *Ann. Bot.* 104, 57–69. <https://doi.org/10.1093/aob/mcp082>.
- Sackville Hamilton, N.R., Chorlton, K.H., 1997. Regeneration of Accessions in Seed Collections: a Decision Guide. Handbook for Genebanks No. 5. International Plant Genetic Resources Institute, Rome, Italy.
- Salazar, D.E., Santos, L.G., Wenzl, P., Hay, F.R., 2020. Effect of dry heat on seed germination of *Desmodium* and *Stylosanthes* species. *Seed Sci. Technol.* 48, 419–437. <https://doi.org/10.15258/sst.2020.48.3.11>.
- Solberg, S.Ø., Yndgaard, F., Andreassen, C., von Bothmer, R., Loskutov, I.G., Asdal, Å., 2020. Long-term storage and longevity of orthodox seeds: a systematic review. *Front. Plant Sci.* 11, 1007. <https://doi.org/10.3389/fpls.2020.01007>.
- Van Treuren, R., de Groot, E.C., van Hintum, Th.J.L., 2013. Preservation of seed viability during 25 years of storage under standard genebank conditions. *Genet. Resour. Crop Evol.* 60, 1407–1421. <https://doi.org/10.1007/s10722-012-9929-0>.
- Van Treuren, R., Bas, N., Kodde, J., Groot, S.P.C., Kik, C., 2018. Rapid loss of seed viability in *ex situ* conserved wheat and barley at 4°C as compared to -20°C storage. *Cons. Physiol.* 6, coy033. <https://doi.org/10.1093/conphys/coy033>.
- Vavilov, N.I., 1997. Five Continents. International Plant Genetic Resources Institute, Rome, Italy.
- Walters, C., Wheeler, L.M., Grotenhuis, J.M., 2005. Longevity of seeds stored in a genebank: species characteristics. *Seed Sci. Res.* 15, 1–20. <https://doi.org/10.1079/SSR2004195>.
- Whitehouse, K.J., Hay, F.R., Ellis, R.H., 2018. Improvement in rice seed storage longevity from high-temperature drying is a consistent positive function of harvest moisture content above a critical value. *Seed Sci. Res.* 28, 332–339. <https://doi.org/10.1017/S0960258518000211>.
- Whitehouse, K.J., Hay, F.R., Lusty, C., 2020. Why seed physiology is important for genebanking. *Plants* 9, 584. <https://doi.org/10.3390/plants9050584>.
- Yamasaki, F., Domon, E., Tomooka, N., Baba-Kasai, A., Nemoto, H., Ebana, K., 2020. Thirty-year monitoring and statistical analysis of 50 species' germinability in genebank medium-term storage suggest specific characteristics in seed longevity. *Seed Sci. Technol.* 48, 269–287. <https://doi.org/10.15258/sst.2020.48.2.14>.