scientific reports

Check for updates

OPEN

Long-term preserved bean seeds exhibit high RNA integrity and high germination potential

Alisson Ferreira Dantas[1,3]✉, Tayara Colins Nunes[2,3], Priscila Grynberg[3], Ana Cristina Miranda Brasileiro[3], Antonieta Nassif Salomão[3], Solange Carvalho Barrios Roveri José[3], Rosana Pereira Vianello[4], Juliano Gomes Pádua[3], Marília de Castro Rodrigues Pappas[3] & Marcos Aparecido Gimenes[3]

Stored RNA plays a key role in seed germination, especially after seeds long-term storage. This study aimed to assess the correlation between germination potential (GP) and RNA integrity in common bean (*Phaseolus vulgaris* L.) seed accessions stored at -18 °C for 19.25 to 43.75 years and in seeds subjected to artificial aging at high temperature (42 °C) and relative humidity (100%). The GP of long-term preserved and artificially aged seed lots ranged from 3 to 100% and 38–87%, respectively. RNA integrity was evaluated using the RNA Integrity Number (RIN, Agilent Bioanalyzer software). RIN data suggested that: (1) Sample sizes for RIN analysis may need to vary based on the GP of the seeds; (2) RIN has the potential to predict physiological quality, especially in seeds with high GP; (3) RIN values by itself may not accurately reflect the rate of RNA degradation over time, necessitating comparison with a control to determine ΔRIN (the difference between the RIN values of the sample and the control); (4) ΔRIN has a significant positive correlation with GP. These findings highlight the potential of RNA integrity in seeds as a molecular marker for developing tests that complement germination tests.

Keywords Aged seeds, *Phaseolus vulgaris* L., RNA integrity number, Seed conservation

Gene banks safeguard the genetic variability of different plant species, which can be utilized in various fields of scientific research to meet both current and future needs of humanity. Typically, seeds are the germplasm samples preserved for long-term conservation in gene banks because they are the natural propagation units of many species, have high storability and allow the maintenance of wide intraspecific variation at low cost and in a safe manner[1]. Effective long-term seed conservation at -20 °C is possible only for species with orthodox seeds, which are seeds that can be dehydrated to low moisture content (≤ 10%)[2].

Germplasm utilization depends on maintaining high levels of seed viability during conservation. The viability of preserved germplasm accessions over time must be periodically checked by monitoring the physiological quality of individual accessions and regenerating them when necessary[3]. Seeds stored for the long-term are monitored using germination tests[4]. However, this method does not detect the initial stages of seed aging, and it may take several years for a decline in germination ability to become evident[5]. As a result, studies have been conducted to identify molecules stored in seeds that degrade in a linear relationship with seed aging, allowing for a more accurate and early estimation of their different aging levels, including those preceding sudden drops in germination capacity. These markers would be of great value in seed conservation, as the transition between life and death in stored seeds often occurs subtly, with the typical signs of life only becoming visible when water is supplied[6].

The loss of germination potential (GP) in orthodox seeds is the result of seed aging, which is associated with numerous forms of cellular and metabolic damage (membrane integrity, energy metabolism, and the impairment of DNA, RNA, and proteins) in which reactive oxygen species (ROS) play a prominent role[7]. DNA damage caused by ROS can be removed by specific repair functions. However, the same process does not occur with messenger RNA (mRNA), which is highly vulnerable to lesions caused by oxidation due to its single-stranded structure, and, in general, mRNA lacks a specific repair mechanism[8]. In our previous studies, we observed slight variation in DNA fragmentation between young and old rice (*Oryza sativa* L.) and common

[1]National Council for Scientific and Technological Development, Brasília, Brazil. [2]Federal District Research Support Foundation, Brasília, Brazil. [3]Present address: Embrapa Genetic Resources & Biotechnology, Brasília, Brazil. [4]Embrapa Rice & Beans, Santo Antônio de Goiás, Brazil. ✉email: alissonfdantas@yahoo.com.br

bean (*Phaseolus vulgaris* L.) seeds using the comet assay, and this variation differed depending on the degree of aging under storage[9] or artificial conditions[10]. Moreover, we observed evidence of variation in the efficiency of the DNA repair system during seed imbibition. Kranner et al.[11]. observed a gradual degradation of DNA and RNA artificially pea (*Pisum sativum* L.) aged seed lots as seed viability decreased, but only the DNA showed a significant decrease in fragmentation after imbibition in seed lots with GP of 85%. These data demonstrated that the loss of RNA integrity in artificially aged pea seeds can be used as a marker of early degradation in aged seed lots.

Assessment of RNA integrity has the potential to be an accurate tool for monitoring the quality of seeds stored in gene banks because RNA integrity declines over time, and RNA damage can be detected before the loss of viability and germinability in seeds[12]. Studies have demonstrated a strong correlation between the RNA integrity number (RIN) and germination percentage[13], with a linear reduction in the RIN occurring over time in the seeds of several long-term preserved species[6]. RIN is an RNA integrity index based on electropherogram profiles generated from high-performance automatic electrophoresis, and it accounts for the proportion of total RNA, the areas of 18 S and 25 S ribosomal RNA peaks and the proportion of the 'fast' area formed by small and fragmented RNAs[14]. RIN is a reliable and accurate RNA integrity measure that generates values ranging between 1 and 10, with 10 indicating fully intact RNA samples and 1 indicating fully degraded RNA in human blood samples[15].

The aim of this study was to evaluate RNA integrity in aged accessions. The results offer new insights into using RNA integrity as a tool for monitoring the viability of common bean seeds at intervals in germplasm collections.

Materials and methods

Long-term stored seeds

We analyzed 19 common bean accessions (Table 1) maintained at -18 °C with a water content of 5 to 7% for 19.25 to 43.75 years at the Embrapa Base Collection, located at Embrapa Genetic Resources and Biotechnology (Brasília, DF, Brazil). These accessions were selected according to the storage time and the quantity of seeds available for GP and RIN analysis. The seed samples were kindly provided by Dr. Juliano G. Pádua, Embrapa Gene Bank supervisor and Embrapa Base Collection curator. Accessions were identified with BRA and numerical sequences issued by the Alelo documentation system. Information on the 19 accessions analyzed is available at https://alelo.cenargen.embrapa.br/index_en.html. Out of the 19 accessions, 10 were introduced into the Base Collection on two separate occasions (BRA 11723-4, BRA 11853-9, BRA 11961-0, BRA 12025-3, BRA 13046-8, BRA 13303-3, BRA 16766-8, BRA 21296-9, BRA 24297-4, and BRA 114672-9), one was introduced three times (BRA 22058), and the remaining eight were introduced only once (BRA 11586-5, BRA 11682-2, BRA 11686-3, BRA 11962-8, BRA 12313-3, BRA 12327-3, BRA 14873-4, and BRA 18257-6). Thirteen samples (BRA 11723–4A, BRA 11723-4B, BRA 11853-9B, BRA 11961-0B, BRA 114672–9A, BRA 114672-9B, BRA 11586-5, BRA 11682-2, BRA 11686-3, BRA 12313-3, BRA 12327-3, BRA 14873-4 and BRA 18257-6) were selected for regeneration according to the current GP, i.e. samples with low GP (< 5%), intermediary GP (between 30 and 74%) and high GP (> 85%). Between eight and sixteen seeds from each sample were used for regeneration in greenhouses, yielding between 129 and 414 fresh seeds per sample. These fresh seeds were then used to compare RNA integrity with the corresponding long-term preserved seed accessions.

A total of 230 seeds from each long-term preserved sample were removed from the Embrapa Base Collection for evaluation. Of these, 150 seeds were divided equally into three aluminum envelopes (3 × 50 seeds) for RNA extraction, and 80 seeds were placed into a single aluminum envelope for the germination test. For the regenerated samples, 75 seeds were analyzed, with 50 seeds placed into one aluminum envelope for RNA extraction and 25 seeds into one aluminum envelope (Buckie and Moore, 70 × 150 cm) for the germination test. The envelopes for RNA extraction were stored at -20 °C (freezer) for two to six months, while the envelopes for germination test were stored at 4 °C (refrigerator) for two weeks to acclimatize the samples since this test is performed in room temperature.

Samples with the same number and different uppercase letters represent the same accession incorporated at least twice into the Embrapa Base Collection on different dates. GP = germination potential. VP (%) = viable plants during regeneration. N/A = not applicable. NR = not regenerated. ΔGP after regeneration = GP after regeneration – current GP. The probit was calculated transforming the proportion of current GP relative to initial GP into a probit value by applying the inverse of the cumulative standard normal distribution function. The storage time and initial GP data are represented by the mean, while the current GP and GP after regeneration data are represented by the mean ± standard deviation. Asterisk (*) represents no null difference from the Agresti–Caffo confidence interval with 95% confidence (CI 95%).

Freshly harvested common bean seeds from the cultivar 'Estilo' (87% of germination potential) were used to produce samples with different germination potentials through artificial aging treatment (see below). The freshly seeds were obtained from a farmer in Brasília who grows the cultivar for sale, then dried to 6% moisture content, sealed in aluminum bags, and stored at 20 °C until use.

High temperature and humidity treatment

Before carrying out the artificial aging at high temperature (42 °C) and relative humidity (100%) treatment, seeds of the cultivar 'Estilo' were placed in Kraft paper envelopes (Ipecol, 11.4 × 22.9 cm and 80 g/m²) and transferred to a Mangelsdorf germinator at 25 °C and 100% relative humidity (RH) for 24 h to increase their water content to facilitate aging. The initial and final water content were 8.17 ± 0.12 and 15.41 ± 0.88, respectively.

The treatment was performed according to the methods of Marcos Filho[16] with modifications. One hundred and eighty seeds per aging treatment were uniformly distributed on a screen suspended inside acrylic boxes (11 × 11 × 3.5 cm) containing 40 mL of water. The boxes were sealed and then incubated at 42 °C for 48, 96,

Accessions (BRA code)	Storage time (years)	Initial GP (%)	Current GP (%)	Probit (Current GP relative to initial GP)	VP (%)	GP after regeneration (%)	ΔGP after regeneration (%)
11586-5	28.83	93	3 ± 2.88*	-1,85	4	96 ± 8.94	93
11682-2	28.83	95	4 ± 2.50*	-1,73	12	100	96
11686-3	28.83	92	30 ± 10.80*	-0,45	25	100	70
11962-8	27.50	99	82 ± 2.88*	0,95	NR	NR	NR
12313-3	27.50	95	45 ± 4.08*	-0,07	37	100	55
12327-3	27.50	100	74 ± 7.50*	0,64	50	92 ± 17.85	18
14873-4	24.16	90	65 ± 8.16*	0,59	69	100	35
18257-6	27.50	53	35 ± 5.77*	0,41	69	96 ± 8.94	61
11723–4A	27.50	90	89 ± 4.78	2,29	62	96 ± 8.94	7
11723-4B	25.83	99	97 ± 2.88	2,05	62	96 ± 8.94	-1
11,853–9A	25.83	97	98 ± 2.88	N/A	NR	NR	NR
11853-9B	24.16	100	99 ± 2.50	2,33	87	96 ± 8.94	-3
11961–0A	43.75	66	94 ± 2.50*	N/A	NR	NR	NR
11961-0B	24.16	98	94 ± 6.29	1,74	62	92 ± 10.95	-2
12025–3A	35.50	85	62 ± 6.45*	0,61	NR	NR	NR
12025-3B	24.66	99	96 ± 2.50	1,88	NR	NR	NR
13046–8A	30.41	100	97 ± 5.00	1,88	NR	NR	NR
13046-8B	21.08	100	92 ± 6.45*	1,41	NR	NR	NR
13303–3A	43.75	97	96 ± 2.50	2,31	NR	NR	NR
13303-3B	24.16	100	100	N/A	NR	NR	NR
16766–8A	27.50	100	94 ± 2.50*	1,55	NR	NR	NR
16766-8B	27.50	100	92 ± 8.66	1,41	NR	NR	NR
21296–9A	24.66	98	91 ± 4.78	1,47	NR	NR	NR
21296-9B	24.16	100	99 ± 2.50	2,33	NR	NR	NR
24297–4A	24,16	96	98 ± 2.88	N/A	NR	NR	NR
24297-4B	24.16	98	97 ± 2.88	2,32	NR	NR	NR
114672–9A	28.91	98	96 ± 2.50	2,05	75	100	4
114672-9B	24.66	100	85 ± 7.07*	1,04	75	96 ± 8.94	11
22058–2A	39.33	95	90 ± 4.08	1,62	NR	NR	NR
22058-2B	28.75	100	97 ± 2.88	1,88	NR	NR	NR
22058–2C	19.25	94	92 ± 5.00	2,03	NR	NR	NR
Mean ± standard deviation	27.88 ± 5.57	94.42 ± 10.00	80.10 ± 27.29	-	-	96.31 ± 4.46	-

Table 1. Accessions analyzed in the present study and their corresponding information on germination potential (GP) before and after storage at the Embrapa Base Collection and regeneration.

120, 144 or 168 h. Usually, bean seed lots are artificially aged at 41 °C[17–19], however, we had to increase the temperature (42 °C) to obtain seed lots with different GP. After each period, the seeds were stored in Kraft paper envelopes and kept in a drying chamber (20 °C and 30% RH) until they reached a moisture content between 6.5 and 7%. The moisture contents were determined by the loss of fresh weight of 20 seeds in an oven at 105 °C.

After drying, the 160 seeds from each aging treatment were divided in two subsamples of 50 seeds and then stored in aluminized envelopes at -20 °C until total RNA extraction to assess RNA integrity. The remaining subsample of 60 seeds from each aging treatment was stored in the refrigerator (4 °C, and 30 to 50% RH) until the germination test was carried out to assess physiological quality.

Evaluation of germination potential

The germination potential was evaluated using 80 long-term preserved seeds (4 × 20) and 60 artificially aged seeds (3 × 20). The substrate used for germination was a Germitest® paper towel roll moistened with distilled water at a proportion of 2.5 times the weight of the dry paper. The Germitest® rolls moistened were put into unsealed plastic bags and placed in a BOD-type germinator at 25 °C. The percentages of normal seedlings, abnormal seedlings, and hard and nongerminating seeds were evaluated on the 9th day after sowing. Seeds considered not to have germinated were subjected to tetrazolium testing to determine their viability, according to Brasil[20]. The GP was defined as the sum of normal seedlings and hard seeds.

The differences between the GP values for each sample recorded at the time the seeds had been introduced at the *Phaseolus* Base Collection (Table 1) and the GP at the time of the present study experiments, as well as the differences among the GP values of artificially aged seed lots of the cultivar 'Estilo', were evaluated by the Agresti–Caffo method with a 95% confidence interval (CI 95%) using the R software[21].

The relationship between the current and the initial GP for each accession was calculated transforming the proportion of current GP relative to initial GP into a probit value by applying the inverse of the cumulative standard normal distribution function (Φ^{-1}), as described below:

$$\text{Probit} = \Phi^{-1}(\text{Initial GP}/\text{Current GP})$$

where, Φ^{-1} is inverse of the cumulative standard normal distribution function, initial GP is initial germination potential of the accession, and current GP is current germination of the accession.

Total RNA extraction

Dry seeds that had been kept in aluminized envelopes at -20 °C were pulverized in liquid nitrogen using a grain grinder, followed by manual maceration in a RNAse-free mortar. Approximately 100 mg of powder from each sample was collected and stored in microtubes at -20 °C until RNA extraction.

Total RNA extraction was performed using PureLink™ Plant RNA Reagent (Thermo Fisher Scientific) and an Aurum™ Total RNA Fatty and Fibrous Tissue Kit (Bio-Rad). PureLink™ Plant RNA Reagent was added (1 mL) to each 2 mL microtube containing 100 mg of pulverized seeds. After complete resuspension of the samples using a vortex, the microtubes were incubated horizontally at room temperature for 5 min to maximize the contact area between the sample and the reagent. The subsequent extraction protocol followed the manufacturer's instructions. Total RNA samples were eluted in elution solution from Bio-Rad kit, then they were quantified, and their purity was assessed by spectrophotometry via a Nanodrop (Thermo Fisher), after which the samples were stored at -20 °C.

High-performance electrophoresis

The total RNA from each sample was diluted in RNase free water (Ludwig, Brazil) to approximately 100 ng/μL. Then, the RNA integrity number (RIN) of the diluted RNA was determined by high-performance electrophoresis on an Agilent 2100 Bioanalyzer using an Agilent RNA 6000 Nano kit (Waldbronn, Germany) and a Plant RNA Nano assay according to the manufacturer's protocols. Agilent 2100 software (software version B.02.10, SI764, Agilent Technologies) was used to calculate the RIN. The mean RIN value of freshly harvested seed samples from the regeneration treatment was used as a standard (control) to estimate the decrease in the RIN of the long-term preserved seed samples during the period of preservation at the Embrapa Base Collection (i.e. ΔRIN = mean RIN of the sample – RIN of the control). A comparison of the RIN of the control with the RIN of the long-term preserved accessions and the comparison of the RIN of the artificially aged samples with that of the nonaged samples (0 h) was carried out via the Dunnett test ($p < 0.05$) using SPSS software.

Two sample sizes were performed to evaluate the impact of sampling on the RIN. All long-term preserved accessions had their RIN values obtained from one biological replicate with 50 seeds (1×50) that had three technical. Thirteen long-term preserved accessions were chosen to obtain RIN values from three biological replicas with 17 seeds (3×17).

The long-term preserved samples were grouped into five clusters by *k*-means clustering using SPSS software. The R^2 values obtained with *k*-means groups are often higher than those from individual data due to reduced noise and variability, emphasizing general trends. *K*-means minimizes internal variability, centralizing values around averages and increasing correlations[22]. The mean ΔRIN was used as the criterion for clustering these samples because the RIN value by itself does not allow observing the degradation rate over the storage time. The correlation between the ΔRIN and GP was evaluated using linear regression ($p < 0.05$) with GraphPad Prism software.

Results and discussion

Physiological characterization of long-term preserved samples

The initial and GP recorded in 2022 (Current GP) are presented in Table 1. The 31 samples (i.e. introductions) of 19 accessions had been long-term preserved for an average of 27.88 years, varying from 19.25 to 43.75 years. All of these accessions had high initial GP at the time they were incorporated in the Base Collection (mean ± standard deviation: 94.42 ± 10.00). The accessions were introduced one, two or three times during their conservation period for different reasons, including loss of viability or germinability and complementation of the number of seeds (Table 1). Three introductions exhibit increase in GP (BRA 24297–4 A: 96% of initial GP to 98% of current GP and 24.16 years of storage; BRA 11853–9 A: 97% of initial GP to 98% of current GP and 25.83 years of storage; and BRA 11961–0 A: 68% of initial GP to 94% of current GP and 43.75 years of storage) and one current GP equal to the GP initial after storage (BRA 13303-3B: 100% of initial GP to 100% of current GP and 24.16 years of storage) (Table 1). However, only the introduction BRA 11961–0A had the initial and current GP different statistically using CI 95% and therefore this introduction was excluded from the analysis of GP and RNA integrity. This unexpected increase in germination ability in this introduction during storage could be due to the breaking of the initial physical dormancy after a period of storage, which is associated with a rapid water uptake and respiration rate[23]. For example, BRA 18257-6 regenerated seeds had an increase in GP after seeds scarification (from 48 to 96%) demonstrating the occurrence of a physical dormancy in a common bean.

In general, the mean current GP of the introductions was 94.42 ± 10.00 (mean ± standard deviation) after 27.88 ± 5.57 years of storage (Table 1). The current GP ranged from 3% (BRA 11586-5) to 100% (BRA 13303-3B). The probit analysis provided a standardized measure to evaluate germination retention across accessions, highlighting significant variability in seed viability. Accessions such as BRA 11586-5 (93% of initial GP to 3% of current GP) and BRA 11682-2 (95 to 4%) exhibit very low germination retention, resulting in highly negative probit values (-1.85 and − 1.73, respectively). These low values suggest pronounced viability loss, potentially linked to aging or inherent genetic factors. Conversely, accessions like BRA 11723–4A (90 to 89%) and BRA 11853-9B (100 to 99%) demonstrate high retention, with positive probit values of 2.29 and 2.33, respectively, reflecting strong stability and longevity. These results align with studies indicating that seed longevity is affected

by chemical composition, maturity, seed processing methods[24] and genetic components since several genes or chromosome regions have been associated with greater seed longevity[25–27]. Accessions with high initial GP, such as 13303-3B (100% from 100%) and 24,297–4 A (98% from 96%), consistently display high probit values (2.33 and 2.05), confirming that seed aging progresses more slowly in high-quality seeds under optimal conditions[28].

The 13 regenerated samples had a mean GP of 96.92 ± 3.01% (Table 1). All accessions introduced into the gene bank had a GP ≥ 92%. The percentage of viable greenhouse-grown seedlings was lower for nine out of the 13 regenerated introductions than the viability percentage in the GP (Table 1). This likely occurred because, upon germination, the seedling may grow normally, terminate at various stages of development, or exhibit an abnormal morphology resulting in a failure to survive; however, seedling establishment ultimately determines seedling vigor and crop yield[29]. It is also worth mentioning that the germination testing was carried out under controlled conditions to determine the maximum germination potential of a seed lot, and this does not occur when seedlings are grown in a greenhouse due to the different environmental conditions[20].

GP of high temperature and humidity treatment samples

Artificial aging at high temperature (42 °C) and relative humidity (100%) treatment was used to produce extra samples with reduced GP but without genotype variation for analysis of RNA degradation using RIN. The GP of treated seeds decreased by as much as 38% after 168 h of aging, and differences in the GP values of the aged seeds were observed from 120 h of aging onward (Fig. 1). Three groups could be observed regarding the aging level: (1) samples aged for 48 and 96 h with a GP of 83,00 ± 5,77% and 80,00 ± 10,00%, respectively; (2) samples aged for 120 and 144 h with a GP of 62,00 ± 5,77% and 62,00 ± 2,88%, respectively; and (3) samples aged for 168 h with a GP of 38,00 ± 2,88% (Fig. 1).

The GPs in the artificially aged seeds gradually decreased by a mean of 4% between the initial aging periods (48 to 96 h) until they decreased more intensely at 120 h (25%) and 168 h (49%) in comparison to nontreated seeds (0 h), which had a GP of 87% (Fig. 1). According to Bernal-Lugo and Leopold[30], seed aging is characterized by a sigmoid relationship between viability and time, in which a phase of no apparent change in seed vigor is followed by a phase of rapid onset of mortality. Our results do not allow us to state that the relationship between artificial aging in common beans and GP is characterized by a sigmoid curve, as it would require more aging periods to verify this; however, we observed a downward trend in GP as aging time increases.

RNA integrity evaluation

The number of seeds used to monitor long-term preserved seed germplasm is an important issue since the quantity must be large enough to properly represent the original sample GP and small enough to reduce the risk of allele loss since the GP test is a destructive method. The RIN used to monitor seed collections must also be evaluated considering the number of seeds to be used, but there is little information on the impact of sample size on the RIN. Tetrault et al.[31] observed that a decline in RNA integrity number (RIN) can be detected in 'Williams 82' soybean seeds stored at 5 °C within 10 years using an RNA monitoring test that consumes approximately 30 seeds.

In the present study, we evaluated two sample sizes: one replicate of 50 seeds and three replicates of 17 seeds. The RIN values are presented in Table 2. We expected more homogeneity when sampling 50 seeds of long-term preserved accessions that had a current GP < 5% and > 91%, since in theory, those could be more adequately sampled with smaller samples (34 and 49 seeds, respectively, with 95% confidence level) than those with GP between 30% and 90%, which required between 65 and 51 seeds, respectively[32]. We observed greater RIN variation between 17- and 50-seed samples of introductions that had a GP below 74% (mean ± standard deviation: 0.52 ± 0.17; data not shown) than those that had a GP above 74% (mean ± standard deviation: 0.37 ± 0.15; data not shown). A smaller RIN variation (0.1) was found between samples of BRA 12327-3 [RIN (1 × 50): 6.50 ± 0.10; RIN (3 × 17): 6.40 ± 0.26], which had a current GP of 74% (Fig. 2). These data suggested that the sample size should be accounted when using the RIN as a marker, and sample sizes might have to vary according to the germination potential of the sample, as accessions with a GP ≥ 74% could be properly sampled from 17 seeds given that less variation in the RIN was observed. The RIN for BRA 11586-5 [RIN (1 × 50): 5.76 ± 0.33 and 3% GP] was unexpected since the value was similar to that for accessions with a GP between 30% and 45% [BRA 12313-3 RIN (1 × 50): 6.73 ± 0.15, BRA 12327-3 RIN (1 × 50): 6.50 ± 0.10, and BRA 18257-6 RIN (1 × 50): 6.20 ± 0.00] and was dissimilar to that for other accessions with a low GP [BRA 11682-2 RIN (1 × 50): 4.30 ± 0.42; GP = 4%]. This may be because whole-seed RNA was analyzed and seeds may maintain some level of RNA integrity even when they have already lost their GP[6].

Control = regenerated samples. N/A = not applicable. RIN (1 × 50) = one biological replicate with 50 seeds that had three technical. RIN (3 × 17) = three biological replicates with 17 seeds. The storage time (years), ΔRIN and ΔRIN.t−1 are represented by the mean, while the other data are represented by the mean ± standard deviation. Accessions with samples incorporated more than one time are represented with the same number and different uppercase letters. ΔRIN = difference between the mean RIN value of the samples and the mean of the control in RIN (1 × 50). $\Delta RIN.t^{-1}$ = difference between the mean RIN value of the samples and the mean RIN value of the control by the storage time in years ($RIN.year^{-1}$) in RIN (1 × 50). Uppercase letters in the same row represent significant differences ($p < 0.05$) between the RIN values of the RIN (1 × 50) and RIN (3 × 17), according to the unpaired t test. Lowercase letters in the same row represent significant differences ($p < 0.05$) between the RIN values before [RIN (3 × 17)] and after regeneration according to unpaired t tests. Asterisks (*) represent significant differences ($p < 0.05$) between the control and Embrapa Base Collection samples.

The GP and RIN of 43 samples (long-term preserved accessions and regenerated samples) were evaluated in this study (Table 2). The long-term preserved accessions with a GP < 85% had a mean RIN of 6.18 ± 1.26, while the long-term preserved accessions with a GP ≥ 85% had a mean RIN of 7.68 ± 0.61 (Fig. 3), suggesting that samples with a high GP also have a high RIN. Thus, the RIN seems to have the potential to predict the

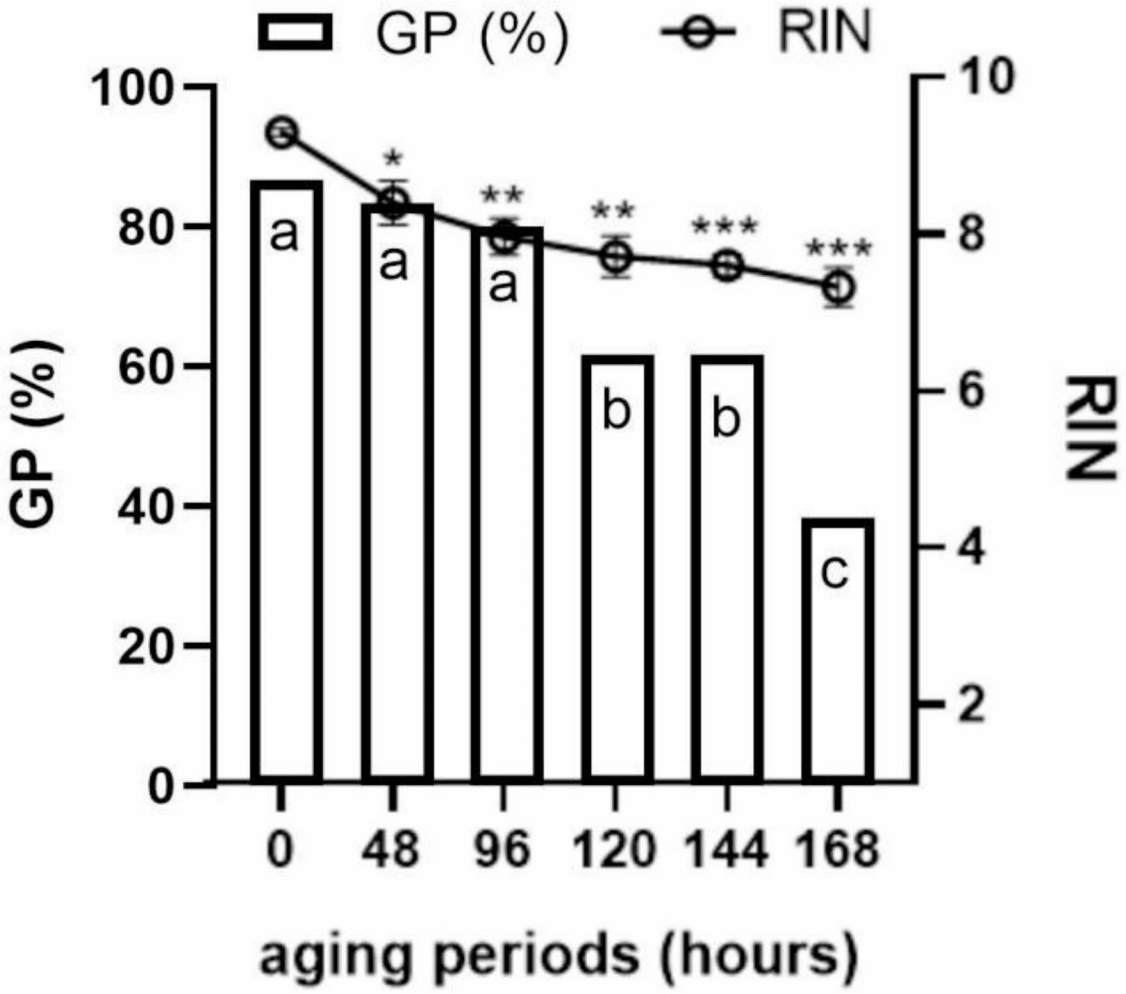

Fig. 1. Germination Potentials (GPs) and RNA integrity number (RIN) from artificially aged seeds of *Phaseolus vulgaris* L at 42 °C for 48, 96, 120, 144 or 168 h. GPs and RIN are represented by bars and circles, respectively. Different lowercase letters represent no null differences in the Agresti–Caffo confidence interval with 95% confidence between GPs. Asterisks represent significant differences ($^{*}p < 0.05$, $^{**}p < 0.01$ and $^{***}p < 0.001$) between RIN values of nonaged samples (0 h) and RIN values of artificially aged samples according to one-way ANOVA followed by Dunnett's test.

physiological quality of bean samples with high GP because RIN > 7 is considered indicative of high RNA integrity[6]. For example, regenerated BRA 18257-6 (RIN = 8.5) showed a GP of 48 and 96% before and after the scarification treatment, respectively, suggesting that the RIN reflects the physiological quality of introduction without the requirement for dormancy breaking during the scarification treatment. This approach could be used in initial trials, and in cases where the RIN was lower, a traditional GP test would be carried out to determine the GP of the sample.

We compared the RIN of three biological replicates of fresh seeds from the cultivar 'Estilo' to the mean RIN of three biological replicates of 13 regenerated introductions to establish a RIN control for long-term preserved bean seeds (Fig. 3). The RIN of the cultivar 'Estilo' was 9.47 ± 0.29, ranging from 9.2 to 9.9, while the mean RIN of the regenerated introductions was 8.98 ± 0.56, ranging from 7.8 to 9.8. No significant difference ($p > 0.05$) was observed between these groups. Therefore, it was assumed that a RIN value of 8.98 would correspond to the initial value of a common bean seed sample properly processed to be incorporated into a gene bank [i.e., > 85%[4]], and it could be used as a control value.

Samples (BRA code)	Storage time (years)	RIN (1 × 50 seeds)	RIN (3 × 17 seeds)	RIN variation between samples size	RIN after regeneration	ΔRIN	ΔRIN.t^{-1}
Control	––	8.98 ± 0.56	N/A	N/A	N/A	––	––
11586-5	28.83	5.76 ± 0.33*	5.27 ± 0.72^{a}	0.49	8.76 ± 0.75^{b}	-3.22	-0.111
11682-2	28.83	4.30 ± 0.42*	3.60 ± 0.00^{a}	0.7	8.90 ± 0.10^{b}	-4.68	-0.162
11686-3	28.83	3.93 ± 0.05*	3.53 ± 0.41^{a}	0.4	9.16 ± 0.20^{b}	-5.05	-0.175
11962-8	27.5	6.13 ± 0.05*	N/A	N/A	N/A	-2.85	-0.103
12313-3	27.5	6.73 ± 0.15*.^{A}	6.10 ± 0.17^{a}.^{B}	0.63	9.73 ± 0.05^{b}	-2.25	-0.081
12327-3	27.5	6.50 ± 0.10*	6.40 ± 0.26^{a}	0.1	9.50 ± 0.10^{b}	-2.48	-0.090
14873-4	24.16	8.06 ± 0.05*.^{A}	7.40 ± 0.10^{a}.^{B}	0.66	8.93 ± 0.11^{b}	-0.92	-0.038
18257-6	27.5	6.20 ± 0.00*.^{A}	6.46 ± 0.11^{a}.^{B}	0.26	8.50 ± 0.40^{b}	-2.78	-0.101
11723–4A	27.5	7.26 ± 0.05*	7.53 ± 0.63^{a}	0.27	9.46 ± 0.05^{b}	-1.72	-0.062
11723-4B	25.83	7.16 ± 0.05*.^{A}	7.37 ± 0.05^{a}.^{B}	0.21	9.66 ± 0.23^{b}	-1.82	-0.070
11853–9A	25.83	7.06 ± 0.05*	N/A	N/A	N/A	-1.92	-0.074
11853-9B	24.16	8.05 ± 0.49*	8.46 ± 0.05	0.41	8.56 ± 0.45	-0.93	-0.004
11961-0B	24.16	7.88 ± 0.05*	N/A	N/A	8.80 ± 0.10	-1.10	-0.045
12025–3A	35.5	7.26 ± 0.05*	7.60 ± 0.78	0.28	N/A	-1.72	-0.048
12025-3B	24.66	7.86 ± 0.05*	N/A	N/A	N/A	-1.12	-0.045
13046–8A	30.41	7.96 ± 0.05*	N/A	N/A	N/A	-1.02	-0.033
13046-8B	21.08	7.31 ± 0.44*	N/A	N/A	N/A	-1.67	-0.079
13303–3A	43.75	7.84 ± 0.25*	N/A	N/A	N/A	-1.14	-0.026
13303-3B	24.16	8.08 ± 0.08*	N/A	N/A	N/A	-0.90	-0.037
16766–8A	27.5	6.80 ± 0.17*	N/A	N/A	N/A	-2.18	-0.079
16766-8B	27.5	8.06 ± 0.65*	N/A	N/A	N/A	-0.92	-0.033
21296–9A	24.66	7.76 ± 0.11*	N/A	N/A	N/A	-1.22	-0.049
21296-9B	24.16	7.53 ± 0.05*	N/A	N/A	N/A	-1.45	-0.06
24297–4A	24.16	7.63 ± 0.05*	N/A	N/A	N/A	-2.48	-0.063
24297-4B	24.16	8.73 ± 0.05	N/A	N/A	N/A	-2.05	-0.071
114672–9A	28.91	8.66 ± 0.05^{A}	8.03 ± 0.23*.^{B}	0.63	7.96 ± 0.15	-1.12	-0.058
114672-9B	24.66	8.40 ± 0.17	8.00 ± 0.26^{a}	0.4	8.76 ± 0.30^{b}	-1.35	-0.056
22058–2A	39.33	6.50 ± 0.00*	N/A	N/A	N/A	-0.25	-0.010
22058-2B	28.75	6.93 ± 0.05*	N/A	N/A	N/A	-0.32	-0.011
22058–2C	19.25	7.86 ± 0.05*	N/A	N/A	N/A	-0.58	-0.023
Mean ± standard deviation	27.65 ± 5.10	7.21 ± 1.12	6.60 ± 1.54	0.42 ± 0.19	8.97 ± 0.49	-1.77 ± 1.12	-0.06 ± 0.04

Table 2. The RNA integrity number (RIN) values of *Phaseolus vulgaris* L. long-term preserved accessions from the Embrapa Base Collection.

The RINs of the 13 regenerated introductions were determined both before and after regeneration. The mean RINs of these samples were 6.60 ± 1.54 before regeneration and 8.98 ± 0.56 after regeneration (Table 2). For most samples, the RIN significantly increased ($p < 0.05$) after regeneration compared to their respective values before regeneration. However, three samples (BRA 11853-9B, BRA 11961-0B, and BRA 114672–9 A) showed no significant increase, primarily because they already exhibited high RNA quality before regeneration, i.e. RIN > 7[6].

Afterward, the RIN was evaluated for 30 samples preserved at the Embrapa Base Collection (Table 1) and for two biological replicates of artificially aged samples, using pools of 50 seeds for each replicate. The mean RIN values of long-term preserved samples ranged from 3.93 (BRA 11686-3) to 8.73 (BRA 24297-4B), with a mean of 7.21 ± 1.12 (Table 2). These values indicate a reduction in the RIN during the storage time since the RIN values of the long-term preserved samples differed significantly ($p < 0.05$) from the value of the control (RIN = 8.98; Table 2), except the introductions BRA 24297-4B, 114672-9B and 114672-9B ($p > 0.05$). Conversely, the mean RIN values of the artificially aged samples were greater, ranging from 8.40 (48 h) to 7.32 (168 h). When comparing these RIN values with the mean RIN value determined for the nonaged seeds (0 h), there was a significant reduction ($p < 0.05$) [Fig. 1]. A decrease in RNA integrity, as measured by the RIN has been observed in seeds of different species aged under storage and artificial conditions. Fleming et al.[6]. reported a significant decrease in RNA integrity when analyzing the RIN values of *Trifolium incarnatum* L., *Sorghum bicolor* (L.) Moench, *Carthamus tinctorius* L. and *Sesamum indicum* L. seeds stored for 59 years at -12 °C. Saighani et al.[33]. reported a decrease in the RIN with a reduction in the germinability of artificially aged *O. sativa* seeds through controlled deterioration control (CDT). They also demonstrated that the RNA degradation process of rice seeds was induced for two weeks under CDT conditions at 36 °C and 80% RH.

The RIN value by itself may not show the RNA degradation rate of degradation over storage time, requiring a comparison with a control. The use of ΔRIN emphasizes deviations from a reference state, capturing meaningful biological variations while reducing baseline noise and seed lots effects. This improves the signal-to-noise

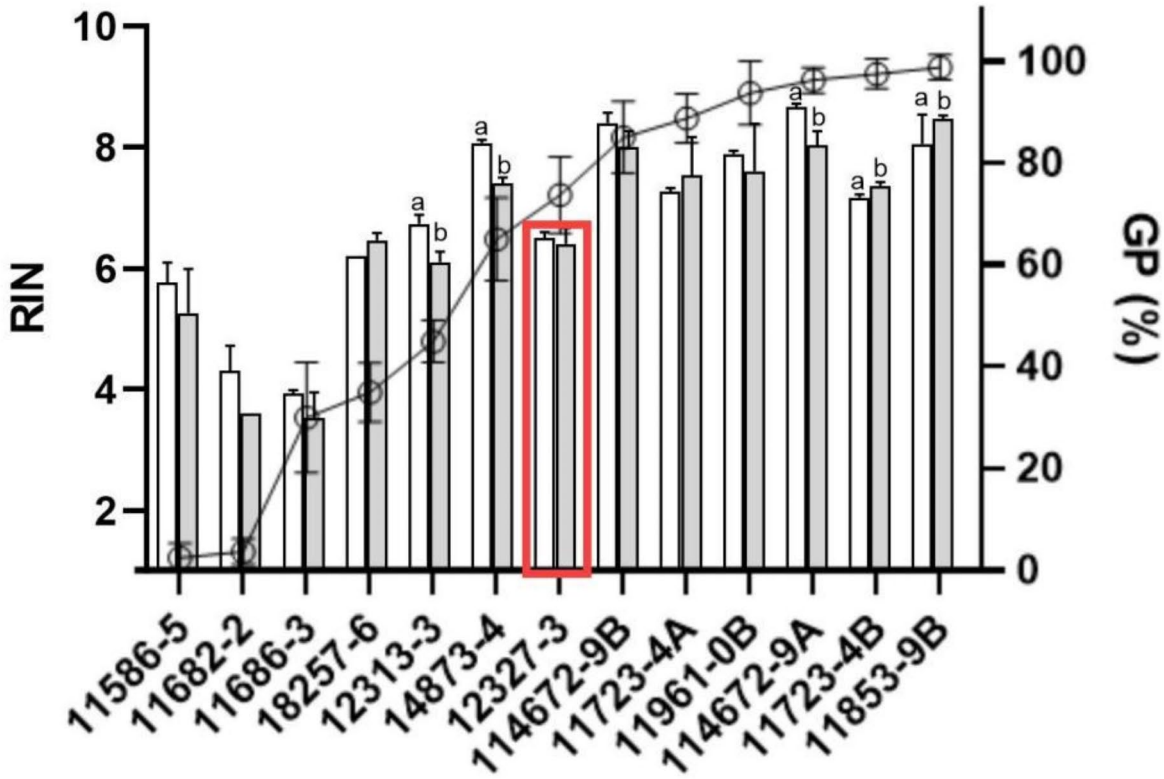

Fig. 2. RIN variation between two sample sizes, (1 × 50 and 3 × 17 seeds) as germination potential (GP) increases. White bars represent the mean RIN for one replicate of 50 seeds. The gray bars represent the mean RIN for three replicates of 17 seeds. Circles represent GP. Red rectangle represents smaller RIN variation between sample sizes. Lowercase letters represent significant differences ($p < 0.05$) between the replicates according to unpaired t tests.

ratio and highlights RNA integrity dynamics, such as degradation or recovery, which are more biologically relevant than static measurements[34]. ΔRIN better reflects the processes influencing GP and enhances statistical relationships, leading to stronger correlations and improved model fits[35]. Thus, ΔRIN is a sensitive metric for studies prioritizing relative trends over absolute measures.

The linear regressions between the GP and ΔRIN values of the long-term preserved and artificially aged samples are shown in Fig. 4. Linear regression revealed a positive and significant correlation between GP and ΔRIN for both groups ($R^2 = 0.8774$ and $p = 0.0189$ for long-term preserved accessions, and $R^2 = 0.8859$ and $p = 0.0233$ for the artificially aged cultivar 'Estilo'). Previous studies have demonstrated a significant correlation by linear regression between germination and RIN values in long-term preserved or artificially aged seeds from other species. Fleming et al.[13]. studied the relationship between the RIN and the germination potential of soybean seeds (the 'Williams 82' cultivar) stored at -5 °C for 1 to 27 years. Linear regression revealed a stronger correlation between the total germination percentage and the mean RIN of cotyledons ($R^2 = 0.91$, $p < 0.0001$) than between apparently healthy germination and the RIN of embryonic axes ($R^2 = 0.66$, $p = 0.0045$). Saighani et al.[33]. demonstrated that the germination rates of three rice cultivars artificially aged by CDT or stored for 8 to 11 years without CDT had a positive and significant correlation with the RIN ($R^2 = 0.75$, $p = 0.00053$). The Fig. 4 demonstrate a clear positive correlation between ΔRIN and GP, emphasizing the importance of RNA integrity in maintaining GP, while the difference in slopes suggests that the aging under storage conditions represented in Fig. 4B may be more responsive to molecular aging as measured by ΔRIN, and the slightly stronger statistical fit in Fig. 4A (higher R^2) may indicate a more consistent or controlled dataset compared to Fig. 4B.

The comparisons between the current and initial GP of the same sample using the Agresti–Caffo 95% CI demonstrated that 18 of the 30 samples maintained a high GP, while 12 samples had a reduced GP. Both groups had similar mean storage times (27 years), but significantly different ($p < 0.01$) mean RIN (Fig. 5). This shows that the storage time acted differently in RNA degradation of these distinct groups. Fleming et al.[13]. suggested that time affects germination potential and RNA integrity in similar, but not identical, ways. We highlighted that

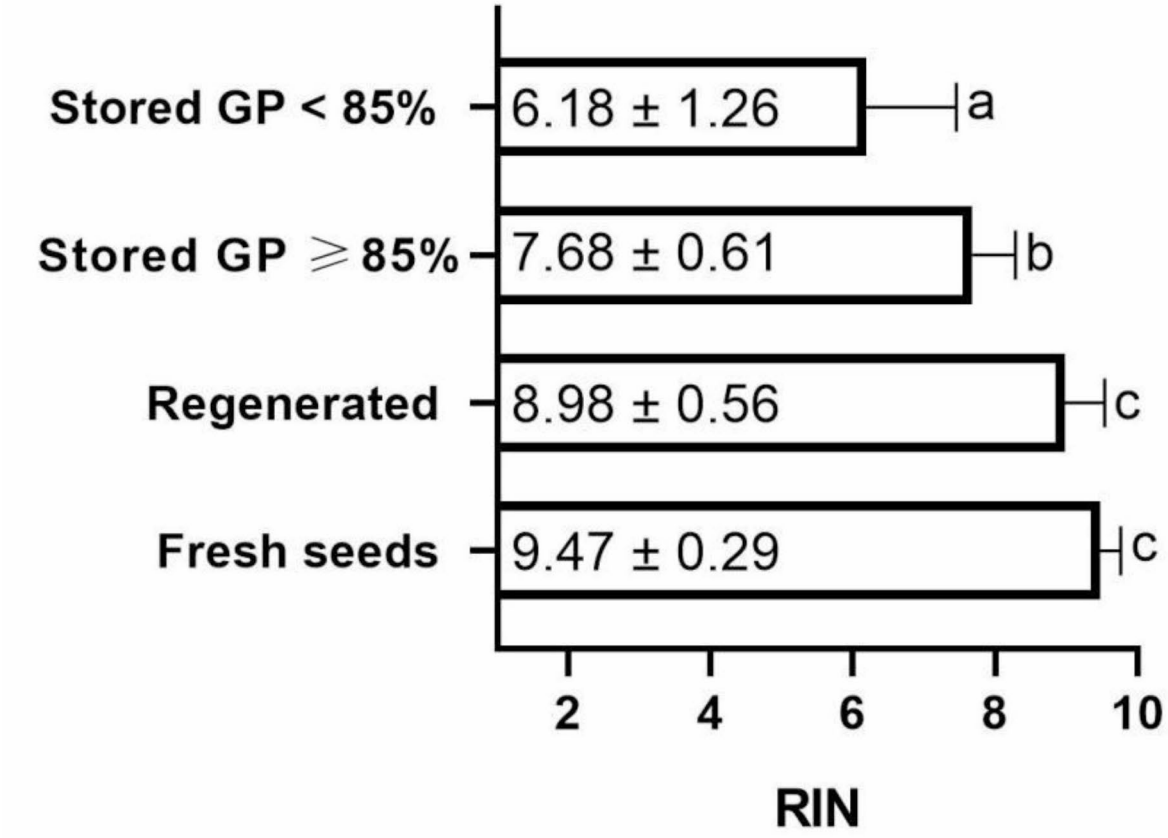

Fig. 3. Mean RIN values in long-term preserved samples that had a GP < 85%, long-term preserved samples that had a GP ≧ 85%, 13 regenerated samples and the fresh seeds ('Estilo' cultivar). In general, samples with a GP ≧ 85% had a RIN above 7, suggesting that the RIN might be a marker of high germination potential.

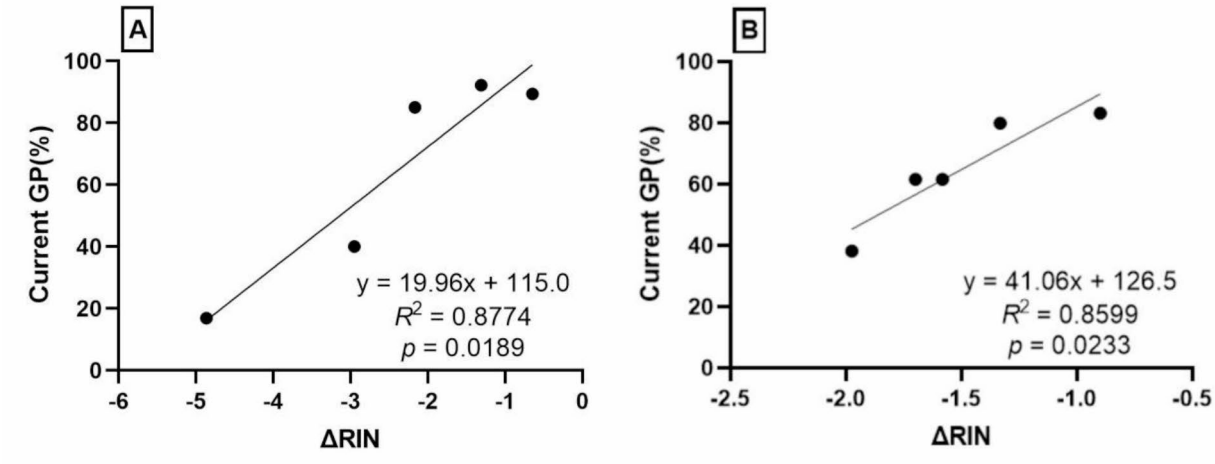

Fig. 4. Linear regressions between the current germination potential (current GP) and ΔRIN of data from long-term preserved accessions clustered by *k*-means (**A**) and from the artificially aged cultivar 'Estilo' (**B**).

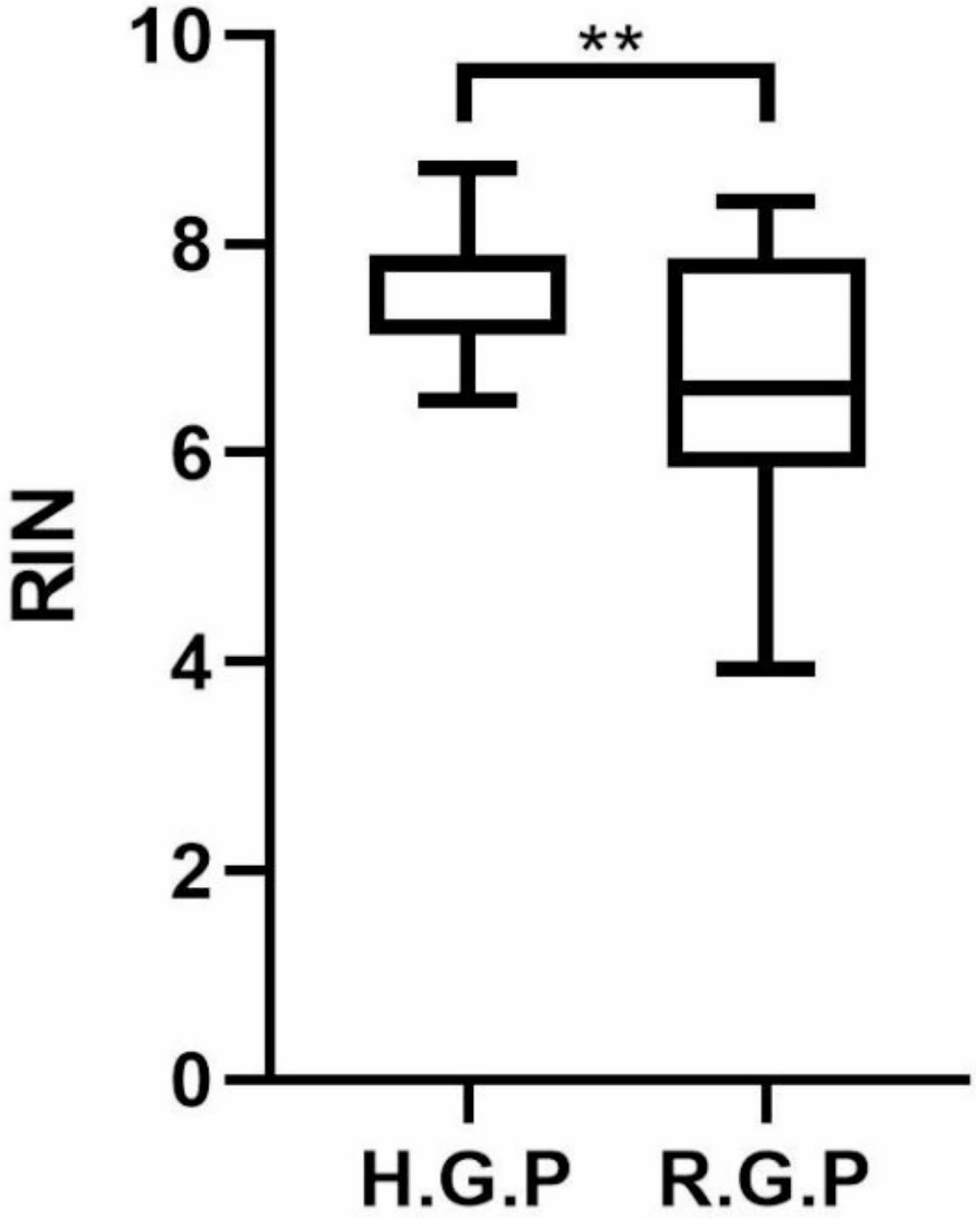

Fig. 5. RIN comparison between accessions with high germination potential maintained (H.G.P) and accessions with reduced germination potential (R.G.P). Asterisks represent significant differences (**$p < 0.01$) between H.G.P and R.G.P using unpaired t tests.

Zhao et al.[36]. showed a mRNA degradation at a constant rate in stored seeds under the same conditions during aging in *Arabidopsis*.

We estimated the decrease in RNA integrity of the long-term preserved samples during storage at the Embrapa Base Collection using the RNA degradation rate ($\Delta RIN.t^{-1}$). In general, the mean $\Delta RIN.t^{-1}$ was -0.06 ± 0.04 $RIN.year^{-1}$ of storage (Table 2). Using this index, it was observed that the oldest sample [BRA 13303–3 A, 43.75 years, GP (%) = 96.00 ± 2.50 and RIN = 7.84 ± 0.25] had a $\Delta RIN.t^{-1}$ mean of -0.026 $RIN.year^{-1}$, and the most recent sample [BRA 22058–2 C, 19.25 years, GP (%) = 92.00 ± 5.00 and RIN = 7.86 ± 0.05] had a mean $\Delta RIN.t^{-1}$ of -0.058 $RIN.year^{-1}$, suggesting that RNA from the more recent sample was degraded more quickly than that

from the older sample. To predict the longevity limit, Tetreault et al.[31]. proposed the use of RIN decline values at the early stages of aging in long-term preserved seeds to optimize viability monitoring and regeneration. This approach enables the loss of valuable samples due to excessive testing or a lack of longevity limits to be avoided. Our data suggested that $\Delta RIN.t^{-1}$ has potential to serve as a parameter for differentiating samples with similar GP and RIN values.

The RIN data suggest that RNA integrity could serve as a reliable predictive marker for seed viability in long-term preserved gene banks. By tracking RNA integrity, gene banks can better optimize seed regeneration strategies, thereby reducing the risk of sample loss due to degradation. This approach offers an alternative to germination testing, reducing the number of seeds analyzed. However, further research is required to investigate the relationship between RNA integrity and seed viability across various species, ensuring the broader applicability of this method. Additionally, understanding the molecular mechanisms that govern RNA degradation in stored seeds could provide valuable insights for enhancing storage conditions and extending seed longevity in gene banks. Previous research by Fleming et al.[37] and Zhao et al.[36,38] suggested that qPCR using two pairs of primers for specific transcripts may offer a more reproducible and accessible means of estimating RNA integrity. This other method could provide consistency across different laboratories and facilitate long-term monitoring of seed collections, offering a practical solution for gene banks worldwide.

Data availability

All data are disclosed in the manuscript. Further information on data can be directed to the corresponding author.

Received: 1 May 2024; Accepted: 30 December 2024
Published online: 11 April 2025

References

1. Rao, N. K. et al. Biodiversity International. Manual of Seed Handling in Genebanks. (2006).
2. Roberts, E. H. Predicting the storage life of seeds. *Seed Sci. Technol.* **1**, 499–514 (1973).
3. Hay, F. R. & Probert, R. J. Advances in seed conservation of wild plant species: A review of recent research. *Conserv. Physiol.* **1**, 1–11 (2013).
4. Fao *Genebank Standards for Plant Genetic Resources for Food and Agriculture* (Food and Agriculture Organization of the United Nations, 2014).
5. Walters, C., Wheeler, L. & Stanwood, P. C. Longevity of seeds stored in a genebank: Species characteristics. *Seed Sci. Res.* **15**, 1–20 (2004).
6. Fleming, M. B., Hill, L. M. & Walters, C. The kinetics of ageing in dry-stored seeds: A comparison of viability loss and RNA degradation in unique legacy seed collections. *Ann. Bot.* **123**, 1133–1146 (2019).
7. Corbineau, F. The effects of Storage conditions on seed deterioration and ageing: How to improve seed longevity. *Seeds* **3**, 56–75 (2024).
8. Kong, Q. & Lin, C. L. G. Oxidative damage to RNA: Mechanisms, consequences, and diseases. *Cell. Mol. Life Sci.* **67**, 1817–1829 (2010).
9. Dantas, A. F. et al. Analysis of the genetic integrity of rice (*Oryza sativa* L.) and bean (*Phaseolus vulgaris* L.) accessions stored in gene banks. *Genet. Resour. Crop Evol.* **67**, 1999–2007 (2020).
10. Dantas, A. F. et al. Loss of genetic integrity in artificially aged seed lots of rice (*Oryza sativa* L.) and common bean (*Phaseolus vulgaris* L). *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **846**, 403080 (2019).
11. Kranner, I., Chen, H., Pritchard, H. W., Pearce, S. R. & Birtic, S. Inter-nucleosomal DNA fragmentation and loss of RNA integrity during seed ageing. *Plant. Growth Regul.* **63**, 63–72 (2011).
12. Walters, N. J., Pritchard, C., Ballesteros, H. W. & Colville, D. L. Seed Longevity – The Evolution of Knowledge and a Conceptual Framework. *Plants*. 12, 471 (2023).
13. Fleming, M. B., Richards, C. M. & Walters, C. Decline in RNA integrity of dry-stored soybean seeds correlates with loss of germination potential. *J. Exp. Bot.* **44**, 34–49 (2017).
14. Schroeder, A. et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol. Biol.* **7**, 3 (2006).
15. Farrel, R. E. Jr *RNA Methodologies: Laboratory Guide for Isolation and Characterization* (Academic, 2017).
16. Marcos Filho, J. Abrates,. Testes de vigor: importância e utilização. In *Vigor de sementes: Conceitos e testes* (ed. Krzyzanowski, F. C., Vieira, R. D.) 3–24 (1999).
17. Machado Neto, N. B., Custódio, C. C. & Takaki, M. Evaluation of naturally and artificially aged seeds of *Phaseolus vulgaris* L. *Seed Sci. Technol.* **29**, 137–149 (2001).
18. Santos, C. M. R., Menezes, N. L. & de., Villela, F. A. Alterações fisiológicas e bioquímicas em sementes de feijão envelhecidas artificialmente. *Rev. Bras. Sementes* **26**, 10–119 (2004).
19. da Binotti, F. F. Efeito do período de envelhecimento acelerado no teste de condutividade elétrica e na qualidade fisiológica de sementes de feijão. *Acta Sci. Agron.* **30**, 247–254 (2008).
20. Brasil. Regras para análise de sementes (Mapa/ACS. (2009).
21. Bilder, C. R. & Loughin, T. M. *Analysis of Categorical Data with R* (CRC, 2015).
22. Hartigan, J. A., Wong, M. A. & Algorithm AS 136: A K-means clustering algorithm. *J. R Stat. Soc. C: Appl. Stat.* **28**, 100–108 (1979).
23. Soltani, A. et al. The genetics and physiology of seed dormancy, a crucial trait in common bean domestication. *BMC Plant Biol.* **21**, 58 (2021).
24. Bonner, F. T. Storage of Seeds. In The woody plant seed manual (ed. Bonner, F. T., Karrafalt, R. P.) 85–95Agriculture Handbook, (2008).
25. Raquid, R. et al. Genetic factors enhancing seed longevity in tropical japonica rice. *Curr. Plant. Biol.* **26**, 100196 (2021).
26. Niñoles, R. et al. Transcription factor DOF4.1 regulates seed longevity in *Arabidopsis* via seed permeability and modulation of seed storage protein Accumulation. *Front. Plant. Sci.* **13**, 915184 (2022).
27. Arif, M. A. R. et al. Genetic analyses of seed longevity in *Capsicum annuum* L. *Cold Storage Conditions Plants*. **12**, 1321 (2023).
28. Bewley, J. D., Bradford, K. J., Hilhorst, H. W. M., Nonogaki, H. & Seeds *Physiology of Development, Germination and Dormancy* (Springer Science & Business Media, 2013).
29. Rehmani, M. S. et al. Seedling establishment: The neglected trait in the seed longevity field. *Plant. Physiol. Biochem.* **200**, 107765 (2023).
30. Bernal-Lugo, I. & Leopold, A. C. The dynamics of seed mortality. *J. Exp. Bot.* **49**, 1455–1461 (1998).

31. Tetreault, H. et al. A power analysis for detecting aging of dry-stored soybean seeds: Germination versus RNA integrity assessments. *Crop Sci.* **63**, 1481–1493 (2023).
32. Dhand, N. K., Khatkar, M. S. & Statulator An online statistical calculator. Sample size calculator for estimating a single proportion. Preprint at (2014). http://statulator.com/SampleSize/ss1P.html
33. Saighani, K. et al. Correlation between seed longevity and RNA integrity in the embryos of rice seeds. *Plant. Biotechnol.* **38**, 277–283 (2021).
34. Everitt, B. S., Landau, S. & Leese, M. *Stahl, D. Cluster Analysis* (Wiley, 2011).
35. Jain, A. K., Murty, M. N. & Flynn, P. J. Data clustering: A review. *ACM Comput. Surv.* **31**, 264–323 (1999).
36. Zhao, L., Wang, S., Fu, Y. B. & Wang, H. Arabidopsis seed stored mRNAs are degraded constantly over aging time, as revealed by new quantification methods. *Front. Plant. Sci.* **10**, 1764 (2020).
37. Fleming, M. B. et al. Exploring the fate of mRNA in aging seeds: Protection, destruction, or slow decay? *J. Exp. Bot.* **69**, 4309–4321 (2018).
38. Zhao, L., Wang, H. & Fu, Y. B. Analysis of stored mRNA degradation in acceleratedly aged seeds of wheat and canola in comparison to arabidopsis. *Plants* **9**, 1707 (2020).

Acknowledgements

We are very grateful for the support of Paula Torga and Rogério da Costa Vieira in obtaining multiple replicates from the accessions studied, and Luis Alberto Martins Palhares de Melo for the statistical analysis of the germination potential of the bean accessions.

Author contributions

A.F.D. performed the germination test and high temperature and humidity treatment, helped with RNA extraction and high-performance electrophoresis, and wrote the paper; T.C.N. performed RNA extraction and high-performance electrophoresis; R.P.V., P.G., and A.C.M.B. supervised the writing of the paper; A.N.S. supervised the germination potential test; S.C.B.R.J. supervised the high temperature and humidity treatment; J.G.P. helped with the selection of accessions preserved in the Embrapa Base Collection; M.deC.R.P. supervised DNA and RNA extraction and high-performance electrophoresis; M.A.G. supervised all steps and wrote the paper. All the authors have read and contributed to writing the paper.

Funding

This study was supported by Embrapa Genetic Resources and Biotechnology, the National Council for Scientific and Technological Development (CNPq) [grant number 152343/2022-9] and the Foundation for Research Support of the Federal District (FAPDF).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.F.D.