B. Pipan, L. Sinkovič, M. Neji, V. Meglič, K. Susek,
M. Kroc, E. Bellucci, E. Bitocchi and R. Papa (2025).
Seed Science and Technology, **53**, 1, 51-72.
https://doi.org/10.15258/sst.2025.53.1.06

Application of non-destructive MARVIN and NIR seed screening of pre-breeding material: A case study on common bean (*Phaseolus vulgaris*), white lupin (*Lupinus albus*) and Andean lupin (*L. mutabilis*)

Barbara Pipan[1]*, Lovro Sinkovič[1], Mohamed Neji[1], Vladimir Meglič[1], Karolina Susek[2], Magdalena Kroc[2], Elisa Bellucci[3], Elena Bitocchi[3] and Roberto Papa[3]

[1] Crop Science Department, Agricultural Institute of Slovenia, Hacquetova ulica 17, SI-1000 Ljubljana, Slovenia (E-mail: barbara.pipan@kis.si; lovro.sinkovic@kis.si; Mohamed.Neji@kis.si; Vladimir.Meglic@kis.si)
[2] Legume Genomics Team, Institute of Plant Genetics, Polish Academy of Sciences, Strzeszynska 34, Poznan 60-479, Poland (E-mail: ksus@igr.poznan.pl; mkro@igr.poznan.pl)
[3] Department of Agricultural, Food and Environmental Sciences, Polytechnic University of Marche, via Brecce Bianche, Ancona 60131, Italy (E-mail: e.bellucci@univpm.it; e.bitocchi@univpm.it; r.papa@staff.univpm.it)
* Author for correspondence (E-mail: barbara.pipan@kis.si)

(Submitted October 2024; Accepted February 2025; Published online February 2025)

Abstract

Grain legumes are important for sustainable agriculture due to their agronomic and nutritional properties. In this study, a rapid, high-throughput, non-destructive seed screening was performed on 835 lines of *Phaseolus vulgaris, Lupinus albus* and *L. mutabilis* grown under a sub-alpine climate. The lines were analysed for seed size characteristics (thousand seed weight – TSW, width and length) using a MARVIN analyser and for protein and moisture content using near-infrared transmittance spectroscopy. The results showed a strong correlation between TSW and seed size in all species, while protein content varied inversely with moisture content, especially in the *Lupinus* species. Regression analysis identified moisture content as a key factor for protein content, which explained up to 86.2% of the variance in *L. mutabilis*. Principal component analysis showed that trait variation was more strongly influenced by genetic and other environmental factors than by geographical origin. Lines selected based on normalised rank (NR) and weighted trait index (WTI) methods were characterised by balanced trait profiles, such as high TSW and protein content with optimal moisture contents. These results underline the importance of integrating non-destructive screening methods into pre-breeding programs. In this way, breeders can efficiently select legume seeds with superior agronomic and nutritional traits, such as optimal seed size, high protein content and adequate moisture. This approach not only improves the selection process, but also ensures the development of legume varieties that meet both agricultural and nutritional standards.

Keywords: legumes, NIR spectroscopy, plant genetic resources, pre-breeding collection, protein content, seed traits

---

Introduction

Grain legumes play an indispensable role in global food systems, serving as a primary source of protein for human consumption and as livestock feed (Bulut *et al*., 2023). Among these legumes, common bean (*Phaseolus vulgaris* L.) stands out for its significant economic and nutritional value. The evolutionary development of this species has been crucial for the structure and extent of its current and available genetic diversity (Cortinovis *et al*., 2021, 2024). Similarly, two lupin species have gained importance in agriculture and are cultivated as food crops: the European *Lupinus albus* L. (white lupin) and the American *L. mutabilis* Sweet (tarwi or Andean lupin) (Kroc *et al*., 2021). The characterisation, conservation and use of genetic resources of food legumes such as common bean and lupin are fundamental aspects for the development of sustainable agriculture and a healthy food system (Belluci *et al*., 2021).

The measurement of seed characteristics is a fundamental aspect of legume breeding and basic research into the genetic control of seed development. Seed size plays an important role in determining grain yield and also influences end-use quality, such as protein content, and consumer preference (Gupta *et al*., 2006; Whan *et al*., 2014). *P. vulgaris* genotypes show a wide range of thousand seed weights (TSW), with values ranging from about 200 to 1000 g, reflecting the diversity of its genetic pool and its adaptability (Arteaga *et al*., 2019; Sinkovič *et al*., 2019). The TSW of *L. albus* is generally higher than that of *L. mutabilis*, with a range of 250–500 g and ~100–300 g, respectively (Annicchiarico *et al*., 2014; Galek *et al*., 2016; Guilengue *et al*., 2019; Gresta *et al*., 2023). The protein content in the seeds of *P. vulgaris* germplasm is typically between 17 and 28% (Celmeli *et al*., 2018; Los *et al*., 2018). *L. mutabilis* has the highest seed quality of all cultivated lupins with a protein content of up to 44% of dry weight and a fat content of up to 18% of dry weight. In contrast, *L. albus* has a lower protein content (34–42%) and fat content (5–11%) (Gulisano *et al*., 2021).

The conservation of well-characterised genetic resources is fundamental to provide the different genotypes needed for pre-breeding programmes to produce new varieties (Cortinovis *et al*., 2021; Kroc *et al*., 2021). Characterisation of available genetic resources based on seed traits is central to advancing breeding programmes aimed at improving legume varieties. Effective breeding requires precise and efficient evaluation methods to accurately assess these traits and provide information for selection strategies. Traditional approaches for characterisation, while useful, often involve time-consuming and destructive procedures that limit the number of samples that can be analysed and damage the seeds (Murube *et al*., 2021). To overcome these challenges, there is a growing need for innovative, high-throughput, non-destructive screening methods. Such techniques not only preserve the integrity of seed samples, but also enable large-scale assessments with greater efficiency. The use of advanced technologies enables breeders to obtain comprehensive and reliable data on seed traits, facilitating the development of improved legume varieties with higher agronomic performance and nutritional quality.

This study addresses this need by utilising advanced non-destructive screening technologies, specifically MARVIN for seed size characteristics and near-infrared (NIR) transmittance spectroscopy for protein and moisture content analysis. The aim of this study was to apply these techniques to the diverse germplasm of common beans and lupins to characterise the variation patterns in seed traits of these legumes and to identify important traits and lines that can be used in breeding. We anticipate that the results of this study will improve our understanding of legume trait variation and provide valuable insights for the development of new varieties that meet both agronomic and nutritional requirements, thereby contributing to more sustainable agriculture. The data on seed traits of interest would represent the first stage of the screening process, including the evaluation, description and classification of the extensive germplasm collections from the perspective of the capabilities of the seed testing laboratories.

Materials and methods

Plant materials and seed characteristics measurement

A set of 835 lines of common beans and lupins (463 lines of *P. vulgaris*, 258 lines of *L. albus* and 114 lines of *L. mutabilis*), derived from the INCREASE project collections (Bellucci *et al*., 2021), provided a limited number of seeds. These lines are part of the intelligent collection developed within the framework of the INCREASE project, which aims to enhance the genetic diversity and adaptability of legume crops. Our study is a follow-on investigation that utilised this germplasm to further explore and validate the potential of these lines. The selected lines represent a subset of the broader INCREASE project collection, chosen based on their multiplication ability in Slovenia under central European growing conditions. While they do not encompass the entire germplasm from the INCREASE project, they are representative of the genetic diversity within three species. The number of lines selected for our study is intended to provide a comprehensive overview of the genetic variation and potential traits of interest within these species. This subset is representative of the broader genus, allowing us to draw meaningful conclusions and insights that can be applied to the larger collection. Common bean seed multiplication was conducted under field conditions at the Jablje Infrastructure Centre of the Agricultural Institute of Slovenia (46°8'42" N, 14°33'29" E; 320 m a.s.l., sub-alpine climate) during the 2021 and 2022 growing seasons using established cultivation technology. The lupins were multiplied during the growing season 2021 at the same location under greenhouse conditions with an average relative humidity and temperature of 77.9% and 18.6°C, respectively, and a photoperiod of 16 hours/8 hours (light/dark). The qualified air-dried seeds of each line and species were stored under dark, dry (40–50% relative humidity) and cold (4°C) conditions before being analysed for various seed characteristics. The thousand seed weight (TSW; g), mean width (mm) and mean length (mm) of each representative sample ($\leq$ 20 mL) were determined non-destructively using the MARVIN ProLine seed analyser (MarviTech GmbH, Wittenburg, Germany) and MARViN 6 software, following the instructions provided by the manufacturer. The moisture (%; as is) and protein (%; dry matter, 6.25 factor) content of each representative seed sample was determined non-destructively with an Infratec NOVA seed analyser (FOSS Analytics, Denmark) using a sample transport module (STM10; 20 mL) for small samples and a global FOSS calibration curve for beans and pulses.

Data analysis

To investigate the variation in seed traits between the legume species, a comprehensive analysis of five key parameters was carried out. Data from the 835 samples were first analysed using descriptive statistics, including minimum, maximum, range, median, mean, standard error of the mean, standard deviation and coefficient of variation (CV). To gain insight into the relationships between seed size, protein content and moisture content, Pearson's correlation coefficients were calculated for each pair of traits within each species. In addition, stepwise multiple regression analyses were performed to examine the factors influencing protein content. In these models, protein content was treated as the dependent variable, with TSW, seed width, seed length and moisture content as predictors, allowing us to quantify the contribution of each predictor to the variance in protein content between species. Principal component analysis (PCA) was then used to identify patterns of trait variation and differentiation between lines, highlighting the most important traits determining variation within and between species. Finally, two approaches were used to select the best-performing lines based on their agronomic traits. The first approach was to rank each line for individual traits and then normalise these ranks by the total number of lines. This normalisation provides a standardised measure of relative performance and facilitates comparison between traits and lines. The second approach was based on calculating a weighted trait index (WTI) for each accession, a method commonly used in multi-trait selection (Mohammadi and Amri, 2013). In this method, each trait was first normalised to a 0-1 scale using min-max normalisation, where the value of each trait was adjusted relative to the minimum and maximum values observed across all lines. The normalised value of a trait Ti for a line was calculated using the following formula:

$$\text{Normalised value} = \frac{\text{Ti} - \min(\text{T})}{\max(\text{T}) - \min(\text{T})}$$

where Ti is the trait value for the line, min (T) is the minimum value of the trait across all lines and max (T) is the maximum value. A weighted sum of these normalised values was then calculated to determine the WTI for each line. The WTI is defined as:

$$\text{WTI} = \sum_{i=1}^{n} W_i \times \text{Normalised value}$$

where W_i is the weight assigned to the i^{th} trait based on its relative importance in the selection criteria, and n is the total number of traits. The WTI provides a single composite score for each line, allowing a holistic assessment of performance across multiple traits. The lines were then ranked based on their WTI scores and the best-performing lines were selected for further analysis. All statistical analyses were performed using RStudio version 3.4.4 (R Core Team, 2021). The following R packages were used: "stats" for descriptive statistics, correlation and regression analyses, “FactoMineR” for PCA (Lê *et al.*, 2008), and custom scripts for the ranking methods.

Results

Patterns of variation in seed characteristics

The mean TSW was 362.3 ± 112.3 g for *P. vulgaris*, 272.6 ± 58.6 g for *L. albus* and 177.7 ± 36.6 g for *L. mutabilis* (table 1). For the width and length traits, the means were 8.09 and 12.56 mm, respectively for *P. vulgaris*, 9.69 and 10.34 mm for *L. albus* and 8.52 and 10.35 ± 0.83 mm for *L. mutabilis*. For protein and moisture content, the means were 23.61 and 13.09%, respectively for *P. vulgaris*, 24.18 and 13.40% for *L. albus* and 23.48 and 14.99%, respectively for *L. mutabilis*. Of the traits measured, the highest CV was calculated for TSW (31.0% for *P. vulgaris*, 21.5% for *L. albus* and 20.6% for *L. mutabilis*). In contrast, protein content showed the lowest CV: 3.4% for *P. vulgaris,* 2.7% for *L. albus* and 2.2% for *L. mutabilis*.

Table 1. Summary statistics for the seed traits of interest in three legume species studied.

Species	Statistic	TSW (g)	Width (mm)	Length (mm)	Protein (%)	Moisture (% f.wt.)
Phaseolus vulgaris (n = 463)	Minimum	92.79	5.41	8.22	20.70	8.80
	Maximum	740.91	10.71	19.31	27.50	15.25
	Range	648.12	5.30	11.10	6.80	6.45
	Median	356.13	8.11	12.46	23.59	13.47
	Mean	362.25	8.09	12.56	23.61	13.09
	SE.mean	5.22	0.04	0.09	0.04	0.06
	Std.dev	112.33	0.89	1.85	0.80	1.26
	CV (%)	31.00	11.00	14.70	3.40	9.60
Lupinus albus (n = 258)	Minimum	111.61	7.47	7.99	22.81	11.26
	Maximum	527.16	12.31	13.56	26.19	15.39
	Range	415.55	4.84	5.57	3.38	4.13
	Median	267.61	9.64	10.31	24.18	13.48
	Mean	272.60	9.69	10.34	24.18	13.40
	SE.mean	3.72	0.05	0.05	0.04	0.05
	Std.dev	58.58	0.72	0.77	0.65	0.74
	CV (%)	21.50	7.40	7.40	2.70	5.50
Lupinus mutabilis (n = 114)	Minimum	75.62	6.79	8.26	22.66	13.84
	Maximum	260.38	10.54	12.76	24.96	15.59
	Range	184.75	3.75	4.50	2.30	1.75
	Median	177.19	8.53	10.38	23.43	15.03
	Mean	177.65	8.52	10.35	23.48	14.99
	SE.mean	3.42	0.07	0.08	0.05	0.04
	Std.dev	36.55	0.78	0.83	0.53	0.40
	CV (%)	20.60	9.20	8.00	2.20	2.70

TSW = thousand seed weight; SE.mean = standard error of the mean; Std.dev = standard deviation; CV = coefficient of variation.

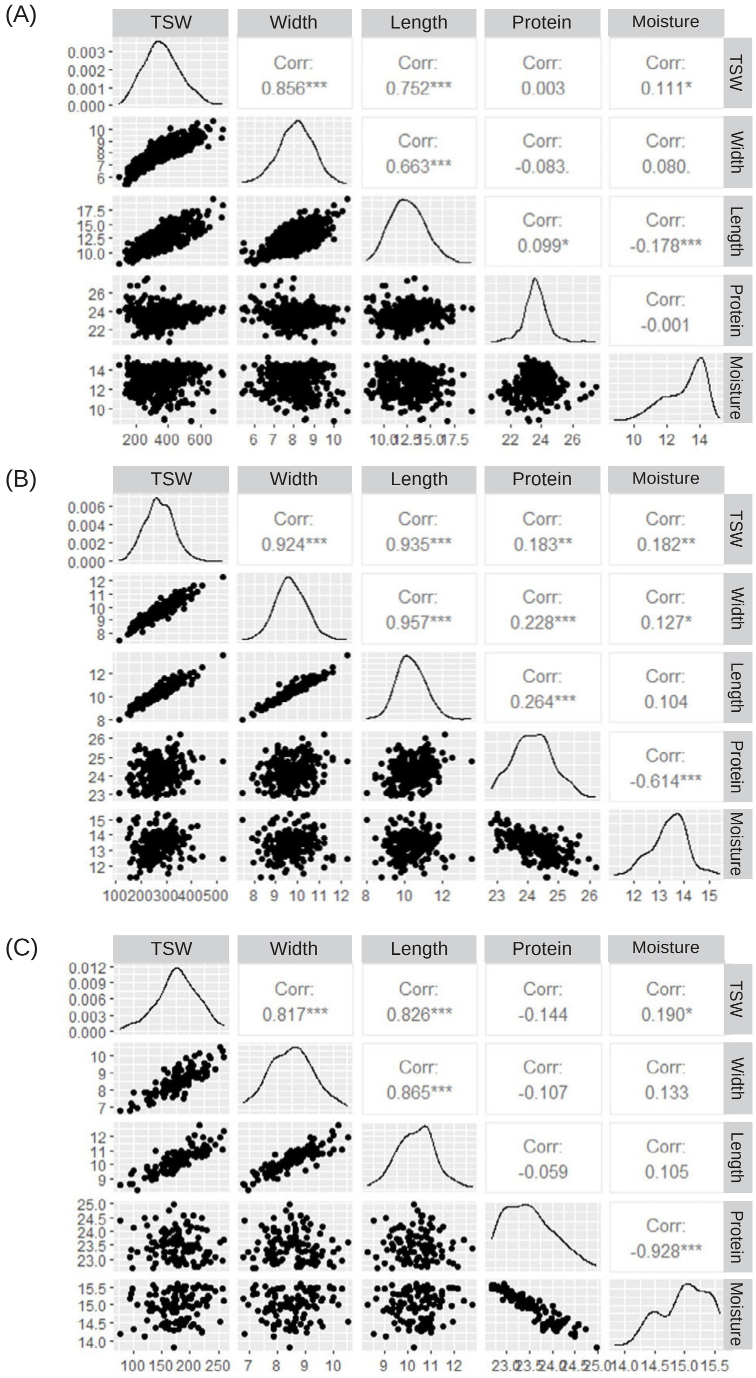

Figure 1. Correlation of seed characteristics of **(A)** *Phaseolus vulgaris,* **(B)** *Lupinus albus* and **(C)** *L. mutabilis.*

Correlations among measured traits

For *P. vulgaris*, TSW showed strong positive correlations with both seed width ($r = 0.86$) and seed length ($r = 0.75$) (figure 1). TSW showed no correlation with protein content ($r = 0.003$) and only a weak positive correlation with moisture content ($r = 0.11$). Seed width and seed length were also positively correlated ($r = 0.66$). Protein content, on the other hand, showed a weak correlation with other traits and a slight negative correlation with seed width ($r = –0.08$) and a weak positive correlation with seed length ($r = 0.10$). Similarly, TSW in *L. albus* showed a strong positive correlation with both seed width ($r = 0.92$) and seed length ($r = 0.94$). However, TSW showed only a weak positive correlation with protein and moisture content ($r = 0.18$). Notably, protein content showed a low positive correlation with seed width ($r = 0.23$) and length ($r = 0.26$), but a pronounced negative correlation with moisture content ($r = –0.61$), where the higher moisture content is often associated with lower protein content. In *L. mutabilis*, TSW again correlated strongly with seed width ($r = 0.82$) and seed length ($r = 0.83$). In contrast to the other species, however, TSW showed a weak negative correlation with protein content ($r = –0.14$) and a weak positive correlation with moisture content ($r = 0.19$). Width and length of the seeds were strongly correlated ($r = 0.87$). Protein content showed a significant negative correlation with moisture content ($r = –0.93$), reflecting a significant inverse relationship between these two traits. Overall, seed size traits (TSW, width and length) were positively associated for all species, meaning that high weight seeds tend to be both wider and longer. However, the relationship between protein content and other traits varied by species, with a particularly strong inverse correlation between protein content and moisture content observed in *L. mutabilis*.

Regression analysis of protein content determinants

A stepwise regression analysis was performed to evaluate the influence of TSW, seed width, seed length and moisture content on the protein content of the three legume species (table 2). For *P. vulgaris*, the regression model showed a statistically significant negative relationship between seed width and protein content ($\beta = –0.305$, SE = 0.079, $t = –3.860$, $p < 0.001$), with a partial R^2 of 0.040, meaning that width accounted for 4% of the variance in protein content. Seed length showed a positive correlation with protein content ($\beta = 0.119$, SE = 0.032, $t = 3.680$, $p < 0.001$), although it contributed only 0.8% to the variance (partial $R^2 = 0.008$). Neither TSW nor moisture content had significant effects on protein content in *P. vulgaris* ($p > 0.05$). In *L. albus*, moisture content proved to be the most significant predictor of protein content and showed a strong negative correlation ($\beta = –0.560$, SE = 0.041, $t = –13.480$, $p <$ 2e-16), which explained 37.4% of the variance (partial $R^2 = 0.374$). Seed length also had a significant positive impact ($\beta = 0.381$, SE = 0.151, $t = 2.507$, $p = 0.013$), accounting for 11.3% of the variance. TSW and width, however, did not significantly affect protein content in *L. albus* ($p > 0.05$). In *L. mutabilis*, moisture content was the most important determinant of protein content, with a strong negative correlation ($\beta = –1.227$, SE = 0.047, $t = –25.721$, $p <$ 2e-16), accounting for a substantial 86.2% of the variance (partial $R^2 = 0.862$). No significant associations were found between protein content and the other traits (TSW, seed width or length) in this species ($p > 0.05$). The R^2 values of the overall model showed that the regression models were most effective in explaining the variance in protein content in *L. mutabilis* ($R^2 = 0.862$), followed by *L. albus* ($R^2 = 0.487$) and *P. vulgaris* ($R^2 = 0.048$). The results emphasise the significant influence of moisture content on protein content, especially in *Lupinus* species.

Table 2. Stepwise regression analyses of seed traits in the three legume species studied.

Species	Parameter	Estimate	Std.Error	*t*-value	*p*-value	Partial R^2	Model R^2
Phaseolus vulgaris	(Intercept)	23.823	0.719	33.110	**< 2e-16**		0.048
	TSW	0.001	0.001	0.759	0.448		
	Width	–0.305	0.079	–3.860	**0.0001**	0.040	
	Length	0.119	0.032	3.680	**0.0002**	0.008	
	Moisture	0.042	0.032	1.330	0.184		
Lupinus albus	(Intercept)	28.544	1.057	27.003	**< 2e-16**		0.487
	TSW	0	0.001	–0.422	0.674		
	Width	–0.063	0.149	–0.425	0.671		
	Length	0.381	0.151	2.507	**0.013**	0.113	
	Moisture	–0.560	0.041	–13.480	**< 2e-16**	0.374	
Lupinus mutabilis	(Intercept)	41.716	0.784	53.207	**< 2e-16**		0.862
	TSW	0	0.001	0.395	0.693		
	Width	–0.051	0.051	–1.011	0.314		
	Length	0.052	0.049	1.058	0.292		
	Moisture	–1.227	0.047	–25.721	**< 2e-16**	0.862	

TSW, thousand seed weight; values in bold are significant at $p > 0.05$.

Patterns of trait differentiation among accessions: PCA-based analysis

The PCA effectively summarised the variance within each species, but did not show distinct geographic clustering, so trait variation is determined by a complex interplay of factors rather than geographic origin alone (figure 2). For *P. vulgaris*, the first two principal components (PCs) explained a substantial proportion of the total variance, with PC1 accounting for 50.3% and PC2 for 21.7% (figure 2A). PC1 was primarily influenced by TSW, seed width and seed length, all of which were strongly positively correlated. Conversely, PC2 was primarily driven by protein content and moisture content, though these traits had weaker correlations with seed size. The PCA biplot for *P. vulgaris* displayed a dispersed pattern of trait variation, indicating that the observed diversity is likely influenced by a combination of environmental conditions, genetic diversity, and local agricultural practices rather than specific geographic regions. In *L. albus*, PC1 and PC2 explained 55.6 and 36.5% of the variance, respectively (figure 2B).

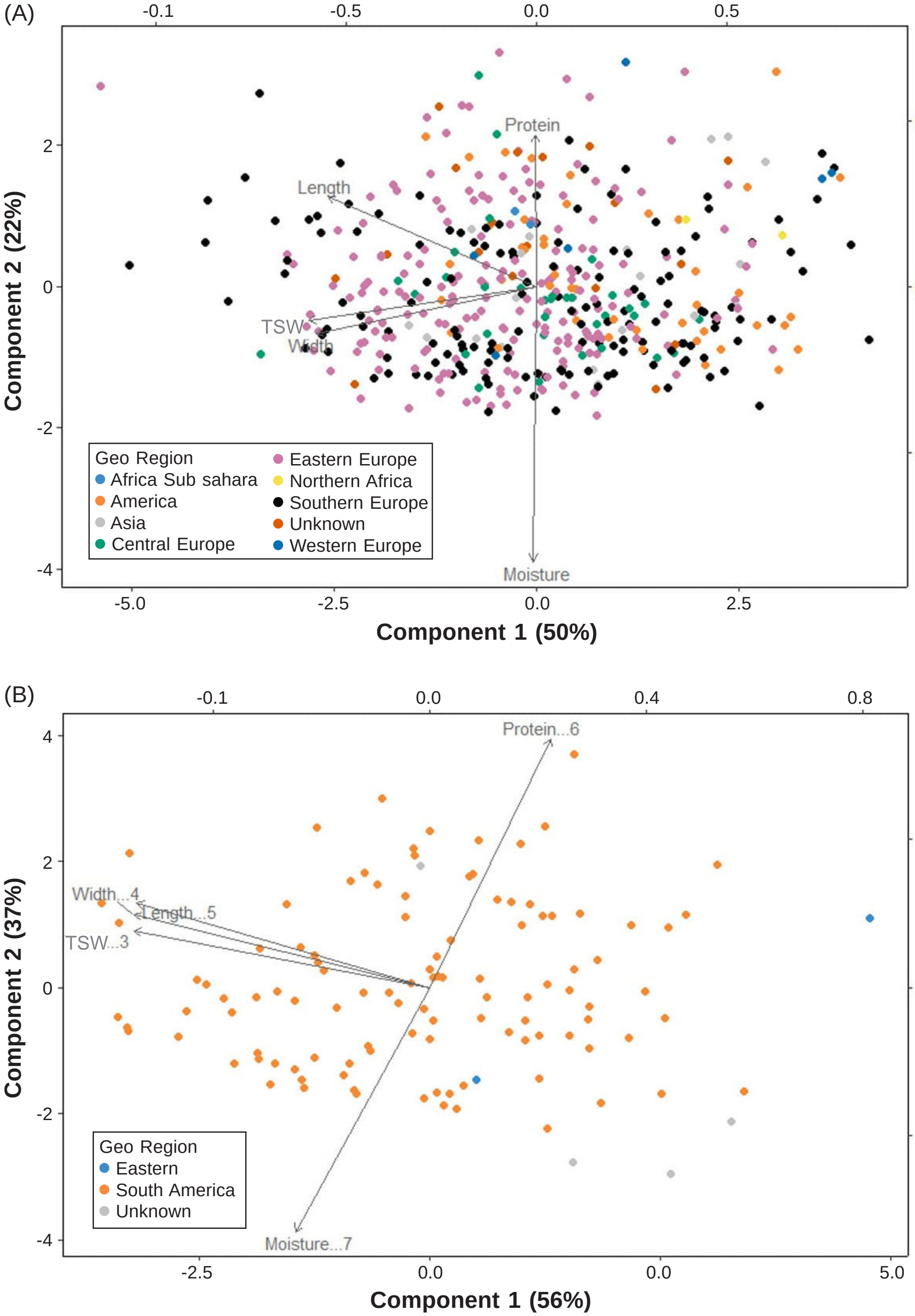

Figure 2. PCA biplot according to the geographical origin of the grain legume lines of **(A)** *Phsaelous vulgaris*, **(B)** *Lupinus mutabilis* and **(C)** *L. albus*.

Figure 2. *Con't*

Figure 2. *Continued.*

(C)

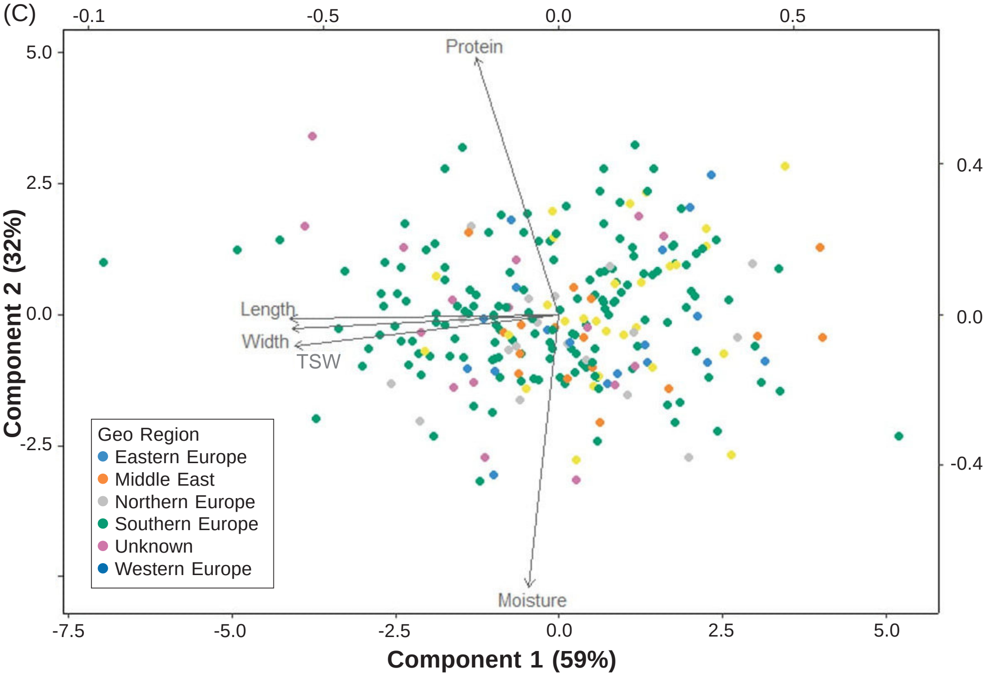

Figure 2. PCA biplot according to the geographical origin of the grain legume lines of **(A)** *Phsaelous vulgaris*, **(B)** *Lupinus mutabilis* and **(C)** *L. albus*.

Similar to *P. vulgaris*, seed size traits such as TSW, seed width and seed length were the main contributors to PC1, while PC2 was driven by moisture content and protein content. The lack of geographic clustering in the PCA biplot for *L. albus* suggests that variation in these traits is shaped by multiple factors beyond geographic origin, underscoring the influence of environmental and genetic interactions on trait expression. For *L. mutabilis*, PC1 and PC2 accounted for 59.3 and 32.2% of the variance, respectively (figure 2C). Seed size traits were closely associated with PC1, while protein content was primarily associated with PC2, showing a negative correlation with moisture content. The PCA biplot for *L. mutabilis* also revealed a dispersed pattern of trait variation, indicating that trait diversity in this species is influenced by a number of factors rather than being tied to specific geographic locations.

At the inter-species level, the first two PCs explained 63.7% of the total variance (39.8% by PC1 and 23.9% by PC2) (figure 3). The PCA biplot highlighted distinct clustering patterns between species, reflecting their unique trait profiles. *P. vulgaris* formed a distinct cluster along the positive side of PC1, strongly associated with TSW and seed length, suggesting that these traits are the dominant factors contributing to the observed variance in this species. *L. albus*, on the other hand, clustered on the positive side of PC2,

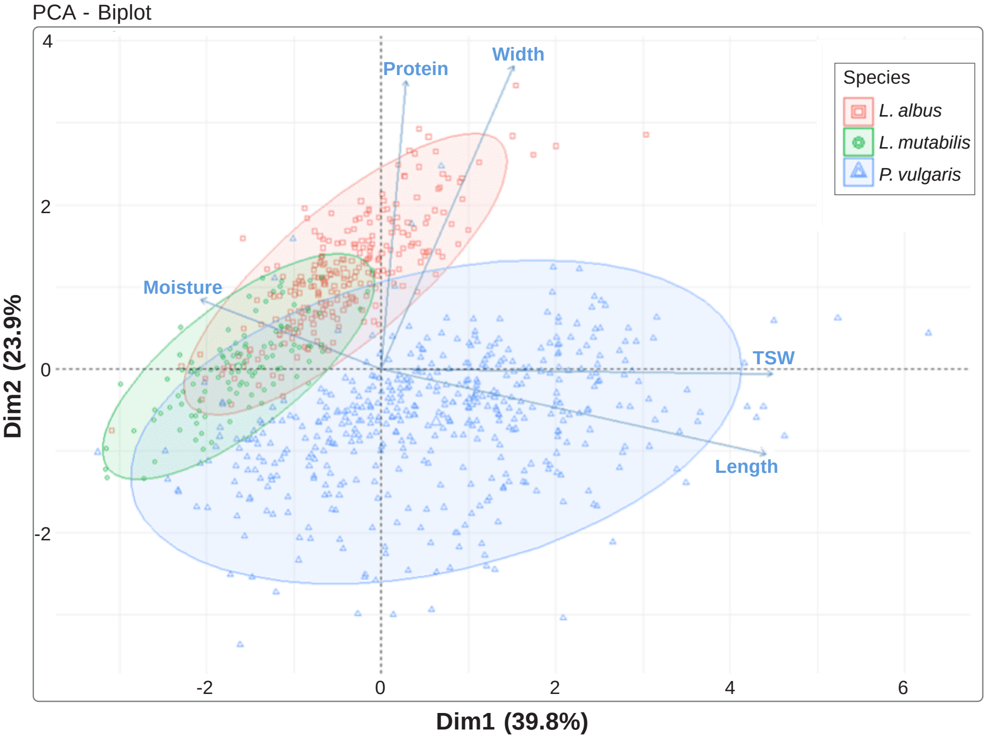

Figure 3. PCA biplot according to the geographical origin of the grain legume lines of *Phaselous vulgaris*, *Lupinus albus* and *L. mutabilis*.

indicating a strong association with seed width and protein content, which are key traits distinguishing this species. *L. mutabilis* was primarily associated with moisture content, clustering towards the negative side of PC1 and the positive side of PC2, indicating that moisture content is a critical factor differentiating this species from the others. The biplot also reveals an inverse relationship between moisture content and protein content across species, particularly in *L. mutabilis*, where the negative correlation is most pronounced.

Best-performing lines based on normalised rank (NR) and weighted trait index (WTI)

The best 20 lines of each species, selected by NR (figure 4, supplementary table 1) and WTI (figure 5, supplementary table 2), showed significant variability in seed traits. For *P. vulgaris*, TSW varied greatly, with IGP_23 (Italy) having the highest TSW and INCBN_00052 (Ecuador) the lowest (figure 4A). Protein content was consistent across lines, but moisture content varied, with INCBN_00054 (Mexico) having higher moisture content than INCBN_06704 (Poland). *L. albus* also showed significant TSW variation, with INLUP_00443 (unknown origin) having the highest TSW and INLUP_00739 (Portugal) the lowest (figure 4B). INLUP_00443 excelled in seed width, while INLUP_00368 (unknown origin) excelled in seed length. Protein content was high

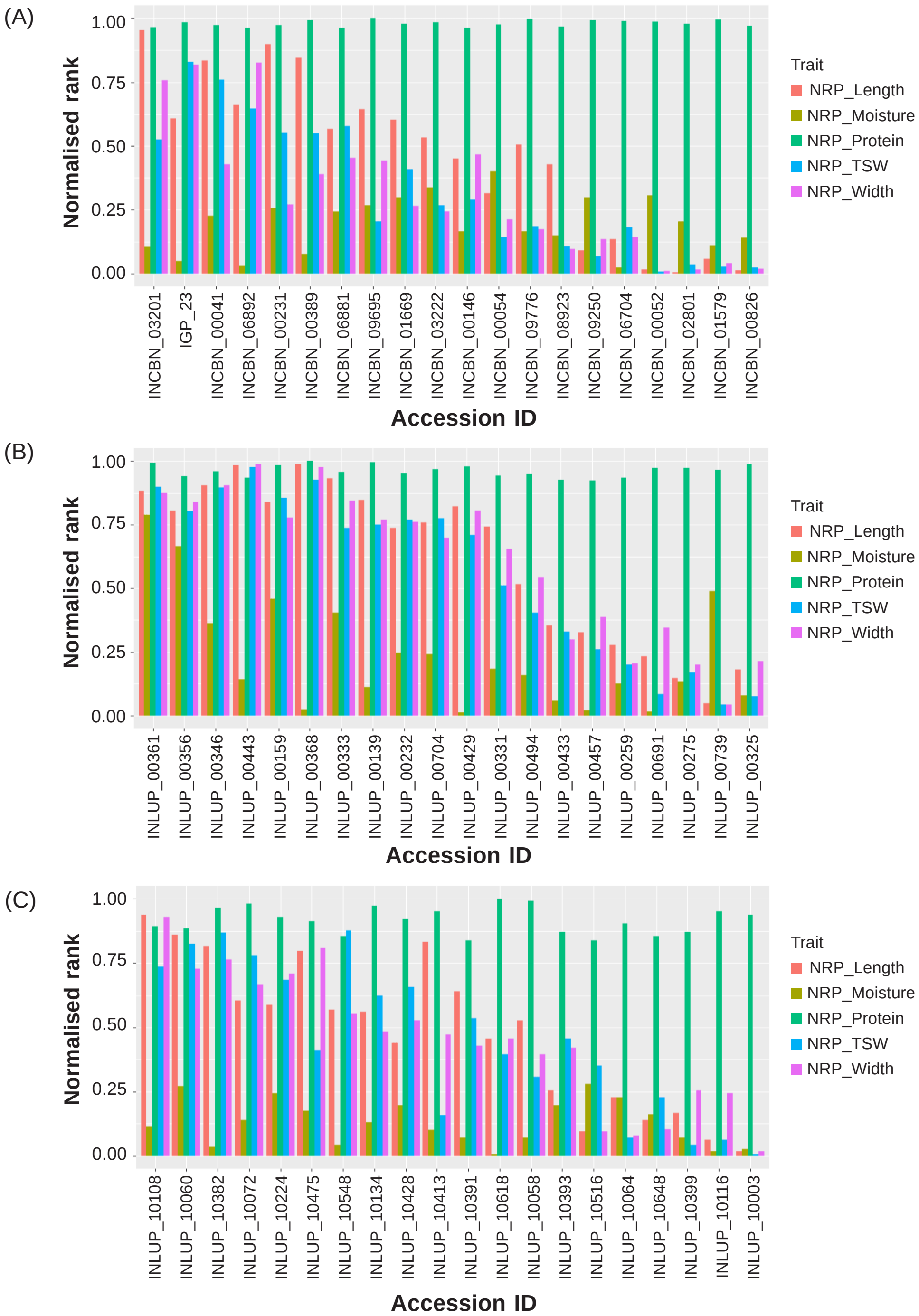

Figure 4. Normalised rank (NR) for the 20 best-performing lines in **(A)** *Phsaelous vulgaris*, **(B)** *Lupinus albus* and **(C)** *L. mutabilis*.

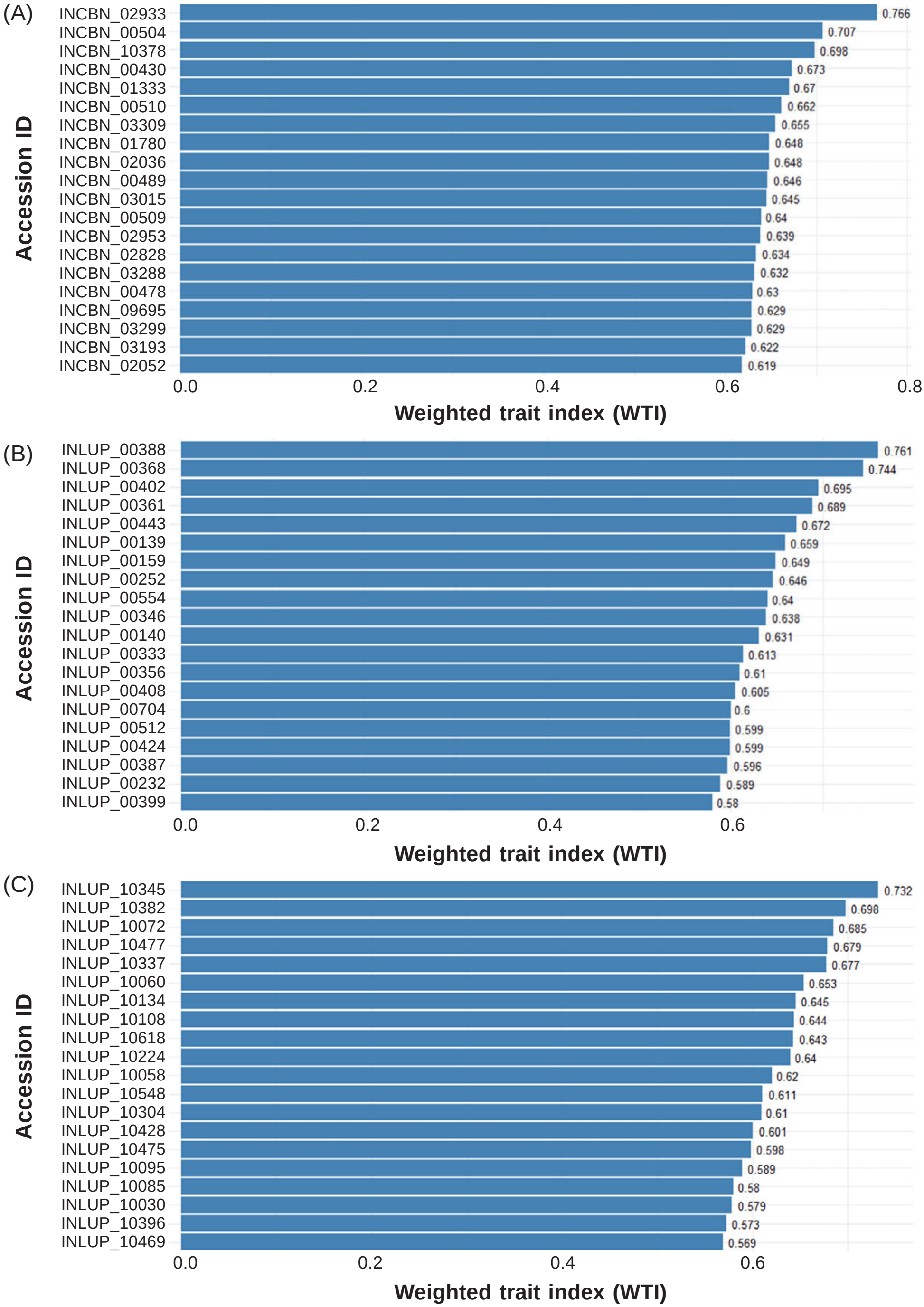

Figure 5. Weighted trait index (WTI) of the 20 best-performing lines in **(A)** *Phaseolus vulgaris*, **(B)** *Lupinus albus* and **(C)** *L. mutabilis* based on the five seed traits analysed.

in INLUP_00368 and INLUP_00139 (Spain), with moisture content varying less. For *L. mutabilis*, TSW ranged from INLUP_10548 (Peru) with the highest to INLUP_10003 (Czech Republic) with the lowest (figure 4C). INLUP_10108 (Peru) had high seed width and length, with all lines showing high protein content and varying moisture content. The WTI method provided a comprehensive evaluation by weighting multiple traits. For *P. vulgaris*, INCBN_02933 (Italy) ranked highest due to high TSW and protein content, followed by INCBN_00504 (Spain) (figure 5A). For *L. albus*, INLUP_00388 (Spain) led due to high TSW, seed width and length, despite INLUP_00368 having high protein content (figure 5B). In *L. mutabilis*, INLUP_10345 (Peru) was the best, characterised by high TSW and seed length (figure 5C). While NR emphasised TSW, WTI offered a more comprehensive assessment by integrating multiple traits.

Discussion

The application of non-destructive MARVIN and NIR seed screening techniques in this study has provided valuable insights into the seed characteristics of the three species, namely *P. vulgaris*, *L. albus* and *L. mutabilis*. Understanding the variation in seed size, protein content and moisture content in these legume species is critical for optimising pre-breeding programmes, as it highlights the importance of focusing on genetic and environmental interactions to improve trait expression and enhance breeding outcomes in different legume species. These non-destructive screening techniques can be further used on genebank collections, aiding in the identification of valuable genetic resources for breeding programmes. Additionally, NIR can be adapted to assess other important seed traits such as fat, starch or gluten content in soybeans, corn and wheat and MARVIN can be applied to various species. However, there are limitations in terms of sample size, seed size and colour for different species.

The variation in TSW between the three legume species is consistent with previous studies showing diversity in seed size within legumes. For example, Beebe *et al*. (2013) found similar TSW patterns in *P. vulgaris*, highlighting the importance of this trait for yield potential and market preferences, where larger seeds are preferred. The high coefficient of variation (CV) observed in *P. vulgaris* in the current study reflects the high genetic diversity within the species. Pérez-Vega *et al*. (2010) also found a wide range of seed sizes in *P. vulgaris*, highlighting the potential for breeding to improve specific traits such as seed size, which is closely linked to overall yield. Generally, higher-yielding crops often produce a larger number of smaller seeds. This trade-off is due to the plant's allocation of resources towards producing more seeds rather than increasing the size of each individual seed. However, the trade-off between seed size and yield is influenced by genetic factors, environmental conditions and agricultural practices (Gupta *et al*., 2006). In contrast, the lower CVs for protein content observed in all legume species suggest that this trait is genetically more stable and less influenced by environmental factors. This result aligns with Subedi *et al*. (2021), who reported that protein content tends to be a more stable trait in legumes due to its strong genetic control. The stability of protein content is beneficial for breeding programmes aiming to improve or maintain high protein content in legumes as it allows for more predictable outcomes when selecting for this trait. The relationship between TSW and protein content adds an interesting dimension to our understanding of these traits. While TSW shows strong positive correlations with seed width and length, suggesting that high weight seeds are typically wider and longer, the lack of a significant correlation between TSW and protein content in *P. vulgaris* ($r = 0.00$) suggests that these traits may be controlled by different genetic mechanisms or respond differently to environmental conditions. This result mirrors the work of Duc *et al*. (2015), who reported that both seed size and protein content are crucial for breeding programmes but are often not strongly correlated, so a balanced approach to breeding strategies is needed to optimise both traits simultaneously.

The variation in TSW observed in *L. albus* and *L. mutabilis* and its relationship to other traits mirrors patterns observed in other legumes. For example, the robust correlation between TSW and seed size traits in *L. albus* ($r = 0.92$ for width and $r = 0.93$ for length) and *L. mutabilis* ($r = 0.82$ for width and $r = 0.83$ for length) is consistent with the results of Kosev and Georgieva (2022), who observed similar relationships in their study on lupins. This agreement between different studies and species suggests that the relationship between TSW and seed size is a robust trait in legumes that can be used in breeding programmes to improve yield traits. The pronounced negative correlation between protein and moisture content, especially in *L. mutabilis* ($r = –0.93$), which explained 86.2% of the variance, reflects previous findings that higher moisture content tends to dilute protein concentration in legume seeds (Erbas *et al*., 2005; Kohajdová *et al*., 2011; Shevkani, 2023). This relationship is crucial for breeding strategies aimed at increasing protein content, as it emphasises the importance of moisture management in both the growth and post-harvest stages. Such a trade-off was discussed by Czubinski *et al*. (2021), who noted that breeding for higher protein content often requires careful moisture management to avoid compromising seed quality.

The strong positive correlations observed in this study between TSW, seed width and seed length across all species are consistent with the findings of previous research conducted by Beyene (2020), who also identified similar relationships in *Lupinus* species. These correlations indicate that seed size traits are closely linked and potentially regulated by analogous genetic pathways, with implications for breeding programmes aimed at increasing seed size. It is noteworthy that the weak or non-existent correlations between TSW and protein content observed in our study corroborate the findings of Muranaka *et al*. (2016), who similarly reported the independence of seed size and protein content in legumes. This independence is advantageous for breeding programmes as it allows for the simultaneous enhancement of yield-related traits and nutritional quality without compromising the other. The results of the regression analysis indicate that the moisture content is a primary determinant of protein content, particularly in the case of *Lupinus* species. The reduced predictive capacity of the regression models for *P. vulgaris* indicates that protein content in this species may be influenced by a more intricate set of factors, as previously discussed by Koinange *et al*. (1996), who emphasised the multifaceted nature of protein content regulation in common beans.

Overall, the PCA results emphasise the distinct trait profiles of each species, with TSW and seed length being pivotal for *P. vulgaris*, seed width and protein content for *L. albus*, and moisture content for *L. mutabilis*. The lack of geographic clustering across all three species suggests that environmental factors, genetic diversity, and agricultural practices play a more significant role in trait variation than geographic origin. The PCA results provide a clear picture of trait differentiation between species, with distinct trait profiles identified for each species. The lack of geographical clustering in the PCA is consistent with the findings of Rodiño *et al.* (2003), who also reported that trait variation in legumes is influenced by genetic diversity and environmental factors rather than geographical origin. This indicates that breeding programmes should prioritise the use of genetic diversity and environmental adaptability over regional selection. The distinct clustering patterns observed in PCA are in accordance with the trait differentiation reported by Atnaf *et al.* (2017) in their study on Ethiopian lupin species. In particular, *P. vulgaris* is strongly associated with TSW and seed length, *L. albus* with seed width and protein content and *L. mutabilis* with moisture content.

Conclusions

It is of particular importance for seed testing laboratories to process data on prospective rapid, non-destructive, high throughput applications such as those presented in this study for three different legume species. This study illustrates the value of rapid and non-destructive seed screening technologies, such as MARVIN and NIR, in the assessment of important seed traits in legumes, providing crucial data for pre-breeding interventions. The results emphasise the necessity of trait-specific breeding strategies, particularly the management of moisture content to optimise protein content in *Lupinus* species. The distinct trait profiles identified in the three species (*P. vulgaris, L. albus* and *L. mutabilis*) provide valuable insights for the development of targeted breeding programmes that enhance both agronomic performance and nutritional quality. The results contribute to the growing knowledge about the variability of legume seed traits and their significance for breeding programmes.

Acknowledgements

The authors gratefully acknowledge financial support from the research programme Agrobiodiversity (P4-0072; Slovenian Research and Innovation Agency), the INCREASE project as funded by the European Union Horizon 2020 Research and Innovation Programme under grant agreement No 862862, and the Slovenian Ministry of Agriculture, Forestry and Food under the Public Service for Vegetables and Herbs. The authors would also like to thank Žiga Doljak, Živa Jurič, Danijela Cvijin and Boštjan Ogorevc for their assistance during vegetation and collecting seed materials and Umesh K. Tanwar for his contribution that improved the clarity and consistency of this research work.

References

Annicchiarico, P., Manunza, P., Arnoldi, A. and Boschin, G. (2014). Quality of *Lupinus albus* L. (white lupin) seed: extent of genotypic and environmental effects. *Journal of Agricultural and Food Chemistry*, **62**, 6539-6545. <https://doi.org/10.1021/jf405615k>

Arteaga, S., Yabor, L., Torres, J., Solbes, E., Muñoz, E., Díez, M.J., Vicente, O. and Boscaiu, M. (2019). Morphological and agronomic characterization of Spanish landraces of *Phaseolus vulgaris* L. *Agriculture*, **9**, article 149. <https://doi.org/10.3390/agriculture9070149>

Atnaf, M., Tesfaye, K., Dagne, K. and Wegary, D. (2017). Genotype by trait biplot analysis to study associations and profiles of Ethiopian white lupin (*Lupinus albus* L.) landraces. *Australian Journal of Crop Science*, **11**, 55-62. <https://doi.org/10.21475/ajcs.2017.11.01.pne226>

Beebe, S.E., Rao, I.M., Blair, M.W. and Acosta-Gallegos, J.A. (2013). Phenotyping common beans for adaptation to drought. *Frontiers in Physiology*, **4**, article 35. <https://doi.org/10.3389/fphys.2013.00035>

Bellucci, E., Aguilar, O.M., Alseekh, S., Bett, K., Brezeanu, C., Cook, D., De la Rosa, L., Delledonne, M., Dostatny, D.F., Ferreira, J.J., Geffroy, V., Ghitarrini, S., Kroc, M., Agrawal, S.K., Logozzo, G., Marino, M., Mary-Huard, T., McClean, P., Meglič, V., Messer, T., Muel, F., Nanni, L., Neumann, K., Servalli, F., Strajeru, S., Varshney, R.K., Vasconcelos, M.W., Zaccardelli, M., Zavarzin, A., Bitocchi, E., Frontoni, E., Fernie, A.R., Gioia, T., Graner, A., Guasch, L., Prochnow, L., Oppermann, M., Susek, K., Tenaillon, M. and Papa, R. (2021). The INCREASE project: Intelligent Collections of food-legume genetic resources for European agrofood systems. *The Plant Journal*, **108**, 646-660. <https://doi.org/10.1111/tpj.15472>

Beyene, C. (2020). Genetic variation among white lupin (*Lupinus albus* L.) landraces from Northwestern and Southern Ethiopia for agronomic traits and nutrient contents of grain. *Journal of Plant Breeding and Crop Science*, **12**, 156-169. <https://doi.org/10.5897/JPBCS2019.0869>

Bulut, M., Wendenburg, R., Bitocchi, E., Bellucci, E., Kroc, M., Gioia, T., Susek, K., Papa, R., Fernie, A.R. and Alseekh, S. (2023). A comprehensive metabolomics and lipidomics atlas for the legumes common bean, chickpea, lentil and lupin. *The Plant Journal*, **116**, 1152-1171. <https://doi.org/10.1111/tpj.16329>

Celmeli, T., Sari, H., Canci, H., Sari, D., Adak, A., Eker, T. and Toker, C. (2018). The nutritional content of common bean (*Phaseolus vulgaris* L.) landraces in comparison to modern varieties. *Agronomy*, **8**, article 166. <https://doi.org/10.3390/agronomy8090166>

Cortinovis, G., Vincenzi, L., Anderson, R., Marturano, G., Marsh, J.I., Bayer, P.E., Rocchetti, L., Frascarelli, G., Lanzavecchia, G., Pieri, A. and Benazzo, A. (2024). Adaptive gene loss in the common bean pan-genome during range expansion and domestication. *Nature Communications*, **15**, article 6698. <https://doi.org/10.1038/s41467-024-51032-2>

Cortinovis, G., Oppermann, M., Neumann, K., Graner, A., Gioia, T., Marsella, M., Alseekh, S., Fernie, A.R., Papa, R., Bellucci, E. and Bitocchi, E. (2021). Towards the development, maintenance, and standardized phenotypic characterization of single-seed-descent genetic resources for common bean. *Current Protocols*, **1**, article e133. <https://doi.org/10.1002/cpz1.133>

Czubinski, J., Grygier, A. and Siger, A. (2021). *Lupinus mutabilis* seed composition and its comparison with other lupin species. *Journal of Food Composition and Analysis*, **99**, article 103875. <https://doi.org/10.1016/j.jfca.2021.103875>

Duc, G., Agrama, H., Bao, S., Berger, J., Bourion, V., De Ron, A.M., Gowda, C.L., Mikic, A., Millot, D., Singh, K.B. and Tullu, A. (2015). Breeding annual grain legumes for sustainable agriculture: new methods to approach complex traits and target new cultivar ideotypes. *Critical Reviews in Plant Sciences*, **34**, 381-411. <https://doi.org/10.1080/07352689.2014.898469>

Erbas, M., Certel, M. and Uslu, M.K. (2005). Some chemical properties of white lupin seeds (*Lupinus albus* L.). *Food Chemistry*, **89**, 341-345. <https://doi.org/10.1016/j.foodchem.2004.02.040>

Galek, R.A., Kozak, B., Bıela, A., Zalewskı, D., Sıenkıewıcz, E.S., Spychała, K. and Stawińskı, S. (2016). Seed coat thickness differentiation and genetic polymorphism for *Lupinus mutabilis* sweet breeding. *Turkish Journal of Field Crops*, **21**, 305-312. <https://doi.org/10.17557/tjfc.99967>

Gresta, F., Oteri, M., Scordia, D., Costale, A., Armone, R., Meineri, G. and Chiofalo, B. (2023). White lupin (*Lupinus albus* L.), an alternative legume for animal feeding in the Mediterranean area. *Agriculture*, **13**, article 434. <https://doi.org/10.3390/agriculture13020434>

Guilengue, N., Alves, S., Talhinhas, P. and Neves-Martins, J. (2019). Genetic and genomic diversity in a tarwi (*Lupinus mutabilis* Sweet) germplasm collection and adaptability to Mediterranean climate conditions. *Agronomy*, **10**, article 21. <https://doi.org/10.3390/agronomy10010021>

Gulisano, A., Alves, S., Rodriguez, D., Murillo, A., Van Dinter, B.J., Torres, A.F., Gordillo-Romero, M., Torres, M.D.L., Neves-Martins, J., Paulo, M.J. and Trindade, L.M. (2022). Diversity and agronomic performance of *Lupinus mutabilis* germplasm in European and Andean environments. *Frontiers in Plant Science*, **13**, article 903661. <https://doi.org/10.3389/fpls.2022.903661>

Gupta, P.K., Rustgi, S. and Kumar, N. (2006). Genetic and molecular basis of grain size and grain number and its relevance to grain productivity in higher plants. *Genome*, **49**, 565-571. <https://doi.org/10.1139/g06-063>

Kohajdová, Z., Karovičová, J. and Schmidt, Š. (2011). Lupin composition and possible use in bakery – a review. *Czech Journal of Food Sciences*, **29**, 203-211. <https://doi.org/10.17221/252/2009-CJFS>

Koinange, E.M., Singh, S.P. and Gepts, P. (1996). Genetic control of the domestication syndrome in common bean. *Crop Science*, **36**, 1037-1045. <https://doi.org/10.2135/cropsci1996.0011183X003600040037x>

Kosev, V. and Georgieva, N. (2022). Dependencies between quantitative traits, determining the seed productivity in white lupine accessions. *Bulgarian Journal of Agricultural Science*, **28**, 401-407. <https://www.agrojournal.org/28/03-05.pdf>

Kroc, M., Tomaszewska, M., Czepiel, K., Bitocchi, E., Oppermann, M., Neumann, K., Guasch, L., Bellucci, E., Alseekh, S., Graner, A. and Fernie, A.R. (2021). Towards development, maintenance, and standardized phenotypic characterization of single-seed-descent genetic resources for lupins. *Current Protocols*, **1**, article 191. <https://doi.org/10.1002/cpz1.191>

Lê, S., Josse, J. and Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, **25**, 1-18. <https://doi.org/10.18637/jss.v025.i01>

Los, F.G.B., Zielinski, A.A.F., Wojeicchowski, J.P., Nogueira, A. and Demiate, I.M. (2018). Beans (*Phaseolus vulgaris* L.): whole seeds with complex chemical composition. *Current Opinion in Food Science*, **19**, 63-71. <https://doi.org/10.1016/j.cofs.2018.01.010>

Mohammadi, R. and Amri, A. (2013). Phenotypic diversity and relationships among a worldwide durum wheat (*Triticum turgidum* L. var. *durum*) germplasm collection under rainfed conditions of Iran. *Crop and Pasture Science*, **64**, 87-99. <https://doi.org/10.1071/CP12403>

Muranaka, S., Shono, M., Myoda, T., Takeuchi, J., Franco, J., Nakazawa, Y., Boukar, O. and Takagi, H. (2016). Genetic diversity of physical, nutritional and functional properties of cowpea grain and relationships among the traits. *Plant Genetic Resources*, **14**, 67-76. <https://doi.org/10.1017/S147926211500009X>

Murube, E., Beleggia, R., Pacetti, D., Nartea, A., Frascarelli, G., Lanzavecchia, G., Bellucci, E., Nanni, L., Gioia, T., Marciello, U., Esposito, S., Foresi, G., Logozzo, G., Frega, G.N., Bitocchi., E. and Papa, R. (2021). Characterization of nutritional quality traits of a common bean germplasm collection. *Foods*, **10**, article 1572. <https://doi.org/10.3390/foods10071572>

Pérez-Vega, E., Pañeda, A., Rodríguez-Suárez, C., Campa, A., Giraldez, R. and Ferreira, J.J. (2010). Mapping of QTLs for morpho-agronomic and seed quality traits in a RIL population of common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics*, **120**, 1367-1380. <https://doi.org/10.1007/s00122-010-1261-5>

R Core Team (2021). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Rodiño, A.P., Santalla, M., De Ron, A.M. and Singh, S.P. (2003). A core collection of common bean from the Iberian peninsula. *Euphytica*, **131**, 165-175. <https://doi.org/10.1023/A:1023973309788>

Shevkani, K. (2023). Protein from land – legumes and pulses (chapter 3). In *Future Proteins, Sources, Processing, Applications and the Bioeconomy* (eds. B.K. Tiwari and L.E. Healy), pp. 35–68, Academic Press, Elsevier. <https://doi.org/10.1016/B978-0-323-91739-1.00003-9>

Sinkovič, L., Pipan, B., Sinkovič, E. and Meglič, V. (2019). Morphological seed characterization of common (*Phaseolus vulgaris* L.) and runner (*Phaseolus coccineus* L.) bean germplasm: A Slovenian gene bank example. *BioMed Research International*, article 6376948. <https://doi.org/10.1155/2019/6376948>

Subedi, M., Khazaei, H., Arganosa, G., Etukudo, E. and Vandenberg, A. (2021). Genetic stability and genotype × environment interaction analysis for seed protein content and protein yield of lentil. *Crop Science*, **61**, 342-356. <https://doi.org/10.1002/csc2.20282>

Whan, A.P., Smith, A.B., Cavanagh, C.R., Ral, J.P.F., Shaw, L.M., Howitt, C.A. and Bischof, L. (2014). GrainScan: a low cost, fast method for grain size and colour measurements. *Plant Methods*, **10**, 1-10. <https://doi.org/10.1186/1746-4811-10-23>

Supplementary table 1.

Normalised rank (NRP) of 20 best-performing lines in beans (*Phaseolus vulgaris*) and lupins **(***Lupinus albus* and *L. mutabilis***)** based on five seed traits analysed.

					P. vulgaris					
ID	TSW	Width	Length	Protein	Moisture	NRP _TGW	NRP _Width	NRP _Length	NRP _Protein	NRP _Moisture
INCBN_09695	268.87	7.98	13.20	27.50	12.30	0.20	0.44	0.64	1.00	0.27
INCBN_09776	255.15	7.29	12.49	27.20	11.70	0.19	0.17	0.51	1.00	0.17
INCBN_01579	162.91	6.42	9.85	26.80	11.30	0.03	0.04	0.06	1.00	0.11
INCBN_09250	206.69	7.12	10.19	26.70	12.50	0.07	0.14	0.09	0.99	0.30
INCBN_00389	368.68	7.90	14.43	26.50	11.00	0.55	0.39	0.84	0.99	0.08
INCBN_06704	253.74	7.16	10.46	25.70	10.20	0.18	0.14	0.14	0.99	0.02
INCBN_00052	142.24	5.80	8.90	25.63	12.54	0.01	0.01	0.02	0.99	0.31
IGP_23	467.73	8.90	12.98	25.60	10.70	0.83	0.82	0.61	0.98	0.05
INCBN_03222	290.38	7.52	12.63	25.45	12.66	0.27	0.24	0.53	0.98	0.34
INCBN_01669	327.19	7.56	12.95	25.40	12.50	0.41	0.27	0.60	0.98	0.30
INCBN_02801	176.55	5.94	8.26	25.40	11.90	0.04	0.02	0.00	0.98	0.21
INCBN_00054	235.55	7.45	11.57	25.25	13.04	0.14	0.21	0.32	0.98	0.40
INCBN_00041	435.80	7.95	14.35	25.20	12.00	0.76	0.43	0.83	0.97	0.23
INCBN_00231	368.79	7.56	14.89	25.20	12.20	0.55	0.27	0.90	0.97	0.26
INCBN_00826	157.72	6.05	8.86	25.16	11.51	0.02	0.02	0.01	0.97	0.14
INCBN_08923	222.06	6.91	12.10	25.10	11.60	0.11	0.10	0.43	0.97	0.15
INCBN_03201	362.57	8.73	15.76	25.02	11.23	0.52	0.76	0.95	0.97	0.11
INCBN_06892	398.15	8.92	13.26	25.00	10.30	0.65	0.83	0.66	0.96	0.03
INCBN_06881	376.38	8.00	12.79	25.00	12.10	0.58	0.45	0.57	0.96	0.24
INCBN_00146	294.96	8.04	12.22	25.00	11.70	0.29	0.47	0.45	0.96	0.17
					L. albus					
INLUP_00368	357.48	11.20	12.16	26.19	11.86	0.93	0.98	0.99	1.00	0.02
INLUP_00139	309.61	10.22	11.12	25.98	12.41	0.75	0.77	0.85	1.00	0.11
INLUP_00361	341.11	10.51	11.24	25.66	13.93	0.90	0.88	0.88	0.99	0.79
INLUP_00325	191.36	9.12	9.65	25.62	12.26	0.08	0.21	0.18	0.99	0.08
INLUP_00159	330.38	10.24	11.08	25.57	13.40	0.85	0.78	0.84	0.98	0.46
INLUP_00429	304.56	10.30	11.02	25.52	11.48	0.71	0.81	0.82	0.98	0.01
INLUP_00275	214.93	9.10	9.56	25.44	12.51	0.17	0.20	0.15	0.97	0.14
INLUP_00691	196.52	9.42	9.80	25.44	11.58	0.08	0.35	0.23	0.97	0.02
INLUP_00704	311.72	10.05	10.87	25.43	12.94	0.77	0.70	0.76	0.97	0.24
INLUP_00739	174.58	8.51	9.17	25.39	13.46	0.04	0.04	0.05	0.96	0.49

Supplementary table 1. *Con't*

Supplementary table 1. *Continued.*

L. albus										
ID	TSW	Width	Length	Protein	Moisture	NRP _TGW	NRP _Width	NRP _Length	NRP _Protein	NRP _Moisture
INLUP_00346	338.03	10.63	11.30	25.37	13.21	0.90	0.90	0.90	0.96	0.36
INLUP_00333	307.53	10.38	11.49	25.34	13.30	0.74	0.84	0.93	0.96	0.40
INLUP_00232	311.39	10.19	10.80	25.32	12.95	0.77	0.76	0.74	0.95	0.25
INLUP_00494	257.59	9.73	10.32	25.30	12.60	0.40	0.54	0.52	0.95	0.16
INLUP_00331	268.53	9.93	10.81	25.29	12.73	0.51	0.65	0.74	0.94	0.18
INLUP_00356	317.30	10.38	10.97	25.24	13.75	0.80	0.84	0.81	0.94	0.67
INLUP_00443	392.21	11.37	11.98	25.23	12.52	0.98	0.99	0.98	0.93	0.14
INLUP_00259	220.01	9.10	9.90	25.23	12.48	0.20	0.21	0.28	0.93	0.13
INLUP_00433	247.22	9.32	10.00	25.21	12.17	0.33	0.30	0.35	0.93	0.06
INLUP_00457	236.40	9.47	9.97	25.19	11.71	0.26	0.39	0.33	0.92	0.02
L. mutabilis										
INLUP_10618	171.19	8.43	10.24	24.96	13.84	0.39	0.46	0.46	1.00	0.01
INLUP_10058	161.83	8.32	10.43	24.74	14.36	0.31	0.39	0.53	0.99	0.07
INLUP_10072	204.67	8.83	10.60	24.57	14.48	0.78	0.67	0.61	0.98	0.14
INLUP_10134	188.87	8.47	10.54	24.56	14.46	0.62	0.48	0.56	0.97	0.13
INLUP_10382	219.56	9.05	10.97	24.50	14.25	0.87	0.76	0.82	0.96	0.04
INLUP_10413	138.46	8.46	11.02	24.49	14.39	0.16	0.47	0.83	0.95	0.10
INLUP_10116	119.87	7.87	9.05	24.49	14.13	0.06	0.25	0.06	0.95	0.02
INLUP_10003	75.62	6.80	8.53	24.40	14.21	0.01	0.02	0.02	0.94	0.03
INLUP_10224	195.55	8.91	10.58	24.34	14.63	0.68	0.71	0.59	0.93	0.25
INLUP_10428	191.64	8.61	10.20	24.27	14.55	0.66	0.53	0.44	0.92	0.20
INLUP_10475	172.90	9.20	10.94	24.26	14.51	0.41	0.81	0.80	0.91	0.18
INLUP_10064	123.38	7.43	9.66	24.22	14.57	0.07	0.08	0.23	0.90	0.23
INLUP_10108	200.33	9.63	11.52	24.16	14.40	0.74	0.93	0.94	0.89	0.11
INLUP_10060	211.80	8.98	11.06	24.15	14.76	0.82	0.73	0.86	0.89	0.27
INLUP_10393	174.86	8.35	9.76	24.12	14.55	0.46	0.42	0.25	0.87	0.20
INLUP_10399	99.22	7.88	9.57	24.12	14.36	0.04	0.25	0.17	0.87	0.07
INLUP_10548	220.89	8.65	10.55	24.09	14.29	0.88	0.55	0.57	0.86	0.04
INLUP_10648	154.15	7.59	9.39	24.09	14.50	0.23	0.11	0.14	0.86	0.16
INLUP_10391	179.07	8.37	10.72	24.03	14.36	0.54	0.43	0.64	0.84	0.07
INLUP_10516	168.44	7.55	9.32	24.03	14.81	0.35	0.10	0.10	0.84	0.28

Supplementary table 2.
Weighted trait index (WTI) of 20 best-performing lines in beans and lupins based on five seed traits analysed.

ID	TSW	Protein	Width	Length	Moisture	WTI
P. vulgaris						
INCBN_02933	740.91	24.34	10.01	18.39	13.56	0.77
INCBN_00504	727.93	23.97	9.26	16.19	13.72	0.71
INCBN_10378	678.18	24.30	10.71	19.31	9.80	0.70
INCBN_00430	638.54	24.23	10.18	16.69	11.80	0.67
INCBN_01333	657.70	23.78	9.99	14.49	14.25	0.67
INCBN_00510	648.94	24.16	8.55	17.25	12.69	0.66
INCBN_03309	590.50	24.13	9.69	14.13	14.54	0.66
INCBN_01780	595.92	24.35	9.14	14.51	13.55	0.65
INCBN_02036	559.57	24.87	8.77	13.96	13.40	0.65
INCBN_00489	611.39	23.94	9.38	14.79	14.06	0.65
INCBN_03015	566.55	24.24	9.69	14.07	14.21	0.64
INCBN_00509	545.16	24.80	8.72	15.68	12.63	0.64
INCBN_02953	572.62	23.94	9.93	14.32	14.34	0.64
INCBN_02828	583.12	23.91	9.48	14.17	14.46	0.63
INCBN_03288	527.18	24.33	9.50	13.76	14.63	0.63
INCBN_00478	626.42	23.74	9.38	13.88	13.84	0.63
INCBN_09695	268.87	27.50	7.98	13.20	12.30	0.63
INCBN_03299	507.93	24.49	8.27	17.08	13.95	0.63
INCBN_03193	605.09	23.90	8.97	13.36	14.15	0.62
INCBN_02052	592.81	23.65	9.79	13.19	14.35	0.62
L. albus						
INLUP_00388	527.16	24.77	12.31	13.56	12.46	0.76
INLUP_00368	357.48	26.19	11.20	12.16	11.86	0.74
INLUP_00402	442.02	24.98	11.60	12.55	12.62	0.70
INLUP_00361	341.11	25.66	10.51	11.24	13.93	0.69
INLUP_00443	392.21	25.23	11.37	11.98	12.52	0.67
INLUP_00139	309.61	25.98	10.22	11.12	12.41	0.66
INLUP_00159	330.38	25.57	10.24	11.08	13.40	0.65
INLUP_00252	402.56	24.97	11.01	11.43	13.14	0.65
INLUP_00554	408.75	24.83	11.62	12.17	12.33	0.64
INLUP_00346	338.03	25.37	10.63	11.30	13.21	0.64
INLUP_00140	335.00	25.05	10.82	11.29	14.41	0.63
INLUP_00333	307.53	25.34	10.38	11.49	13.30	0.61
INLUP_00356	317.30	25.24	10.38	10.97	13.75	0.61

Supplementary table 2. *Con't*

Supplementary table 2. *Continued.*

L. albus

ID	TSW	Protein	Width	Length	Moisture	WTI
INLUP_00408	362.53	24.88	10.52	11.36	13.55	0.61
INLUP_00704	311.72	25.43	10.05	10.87	12.94	0.60
INLUP_00512	364.78	24.78	10.88	11.34	13.42	0.60
INLUP_00424	360.28	25.08	10.59	11.22	12.42	0.60
INLUP_00387	368.73	24.52	11.23	11.84	13.80	0.60
INLUP_00232	311.39	25.32	10.19	10.80	12.95	0.59
INLUP_00399	355.97	24.78	10.66	11.69	12.84	0.58
L. mutabilis						
INLUP_10345	260.38	23.71	9.97	12.39	15.12	0.73
INLUP_10382	219.56	24.50	9.05	10.97	14.25	0.70
INLUP_10072	204.67	24.57	8.83	10.60	14.48	0.68
INLUP_10477	228.38	24.02	9.07	11.34	14.98	0.68
INLUP_10337	253.88	23.61	10.54	11.93	14.56	0.68
INLUP_10060	211.80	24.15	8.98	11.06	14.76	0.65
INLUP_10134	188.87	24.56	8.47	10.54	14.46	0.65
INLUP_10108	200.33	24.16	9.63	11.52	14.40	0.64
INLUP_10618	171.19	24.96	8.43	10.24	13.84	0.64
INLUP_10224	195.55	24.34	8.91	10.58	14.63	0.64
INLUP_10058	161.83	24.74	8.32	10.43	14.36	0.62
INLUP_10548	220.89	24.09	8.65	10.55	14.29	0.61
INLUP_10304	216.29	23.37	10.16	12.76	15.03	0.61
INLUP_10428	191.64	24.27	8.61	10.20	14.55	0.60
INLUP_10475	172.90	24.26	9.20	10.94	14.51	0.60
INLUP_10095	256.25	22.94	10.31	11.04	15.44	0.59
INLUP_10085	195.85	23.99	8.66	10.85	14.65	0.58
INLUP_10030	232.47	23.23	9.99	11.03	15.20	0.58
INLUP_10396	243.71	23.09	9.26	11.46	15.38	0.57
INLUP_10469	214.31	23.79	9.03	10.82	14.37	0.57