Variation for pea seed protein concentration in the USDA *Pisum* core collection

Coyne, C.J.[1], Grusak, M.A.[2], Razai, L.[1] and Baik, B.-K.[3]

[1]USDA-ARS, WRPIS, Pullman, WA, USA
[2]USDA-ARS Children's Nutrition Res. Center, Houston, TX, USA
[3]Dept. of Crop and Soil Sci., Washington State Univ., Pullman, WA, USA

Introduction

Food legumes are important sources of protein, complex carbohydrates, vitamins, and minerals in the diets of millions of people (12). While cereals supply nearly 50% of the protein in the human diet, an unfavorable balance in amino acids (poor in lysine) requires complementary protein sources (12). Legumes are good complements to cereals, as they are rich in lysine, but poor in sulfur containing amino acids (methionine and cysteine) (12). Pea seed proteins are composed of albumins and globins which separate into two major fractions: the 7S vicilin and convicilin fraction, and an 11S fraction that is predominantly composed of legumin (1). In this study, we characterized the total seed protein concentration of 480 accessions from the USDA *Pisum* core collection. The complete data set, which is summarized in this article, is available through the internet (http://www.ars-grin.gov/npgs/) or by contacting the curator (coynec@wsu.edu).

Materials and Methods

Plant material

The USDA *Pisum* core collection (504 accessions) was used as the source of germplasm (9). However, it should be noted that many of the accessions in the *Pisum* core collection are mixtures of diverse germplasm. Because we wished to avoid mixed samples for our quantitative analyses, randomly selected seeds of a single seed phenotype were chosen from each accession and documented for planting. For the most part, this resulted in plants with uniform characteristics within each planted accession. At harvest, if more than one plant phenotype was evident, seeds were selected from one plant phenotype only (usually the phenotype with the most plants, or the highest seed yield). Also, seeds harvested were compared to the original seeds to verify they were the same phenotype as planted. Phenotypic data were collected on harvested seeds of each accession, and these characteristics are noted in the GRIN descriptor dataset listed with the seed protein concentration (Seed Coat Color, Seed Coat Coloration Pattern, Smooth vs.Wrinkled Seeds, Cotyledon Color) (www.ars-grin.gov/cgi-bin/npgs).

Growth conditions

Six plants of each accession were grown in 5L black plastic pots filled with a synthetic soil mix composed of 2 parts Metro-Mix 360 (Scotts-Sierra Horticultural Products Co., Marysville, Ohio) and 1 part medium grade vermiculite (Strong-Lite Medium Vermiculite, Sun Gro Horticulture Co, Seneca Illinois). Plants were grown in a controlled environment greenhouse with a temperature regime of 22 ± 3° C/day and 20 ± 3° C/ night, with a relative humidity ranging from 45% to 65% throughout the day/night cycle. Sunlight was supplemented with metal halide lamps, set to a 15-h day, 9-h night cycle (lights on at 700 h). In order to maintain an adequate supply of all mineral nutrients, a complete fertilizer mixture was provided to each pot on a daily basis. Pots were irrigated with an automated drip irrigation system (one drip line to each pot); the system was regulated with a timer that delivered nutrient solution twice a day (younger plants) or three times a day (older plants) in sufficient quantity to saturate the soil mass at each irrigation. The nutrient solution contained the following concentrations of mineral salts: 1.0 mM KNO_3, 0.4 mM $Ca(NO_3)_2$, 0.1 mM $MgSO_4$, 0.15 mM KH_2PO_4 and 25 :M $CaCl_2$, 25 :M H_3BO_3, 2 :M $MnSO_4$, 2 :M $ZnSO_4$, 0.5 :M $CuSO_4$, 0.5 :M H_2MoO_4, 0.1 :M $NiSO_4$, 1 :M Fe(III)-*N*, *N*'-ethylenebis[2-(2-hydroxyphenyl)-glycine] (Sprint 138; Becker-Underwood, Inc., Ames, Iowa, USA). We thus attempted to maintain all essential minerals at sufficient, non-toxic levels in the soil.

Seed Samples

Plants were grown to maturity and all seeds were collected and combined from the six plants grown for each accession. Combined seeds were counted, dried to zero moisture in a 70 C oven, and weighed, in order to calculate 100 seed weights. Each combined seed sample was ground to a fine powder using a coffee grinder, prior to nitrogen analyses. Measurements were conducted on 480 of the 504 accessions in the USDA *Pisum* core collection.

Seed nitrogen analyses and protein calculations

Seed nitrogen concentrations were determined using a LECO FP-528 Nitrogen/Protein Determinator (Leco Corp., St. Joseph, MI, USA), according to the manufacturer's instruction manual. Weighed aliquots of EDTA (ethylenediamine tetraacetic acid) were used as nitrogen standards to calibrate the instrument. Two sub-samples (0.15 g each) of each accession were analyzed for nitrogen concentration; each sample was measured two times internally in the instrument with the average reported to the operator. The two sub-sample averages were then averaged to get a nitrogen concentration value for each accession. No sub-sample nitrogen values for any accession varied by more than 5%.

Protein concentrations were calculated using a conversion factor of 5.44 ([seed nitrogen concentration] x 5.44 = [seed protein concentration]), a multiplier determined by Mossé (6) as an average value for pea (based on 33 samples). This multiplier is specific for pea, as it takes into account the actual amino acid composition of pea seeds, and the nitrogen weight percentage of those amino acids. Comparison between seed protein concentration and 100 seed weight was calculated using Pearson's correlation (7).

Results

Once the nitrogen analyzer was calibrated, the readings for the two samples per accession were very similar. Using a tolerance level of a maximum 5% difference between samples, no accession required repeat analysis. The seed used in this study were grown under controlled conditions, which should significantly reduce the large effect environment can have on pea seed protein concentration (5). Protein concentration varied over two-fold in the accessions tested with the highest percentage of 30.93% and lowest of 12.38 % in the accessions tested. The results are summarized in a frequency histogram (Fig. 1). The mean seed protein concentration from round seed (426 accessions) was 20.62 and the mean of the wrinkled seed was higher at 23.76 (51 accessions). The accessions with the ten highest and ten lowest seed protein concentrations and their morphological characteristics are presented in Table 1, many with acceptable agronomic characteristics like white flower and clear seed coats. Table 2 summarizes the protein concentrations of the taxons of *Pisum* included in this study. While the sample size between taxon is large (from 1 accession to 437 accessions), the mean seed protein concentrations are very similar (Table 2). A comparison of the seed size (100 seed weight in g) and seed protein concentration was conducted, and no correlation was found for the USDA *Pisum* core (r=0.01; Figure 2). The complete data set and accessions are available through the internet (http://www.ars-grin.gov/npgs/) or by contacting the curator (coynec@wsu.edu).

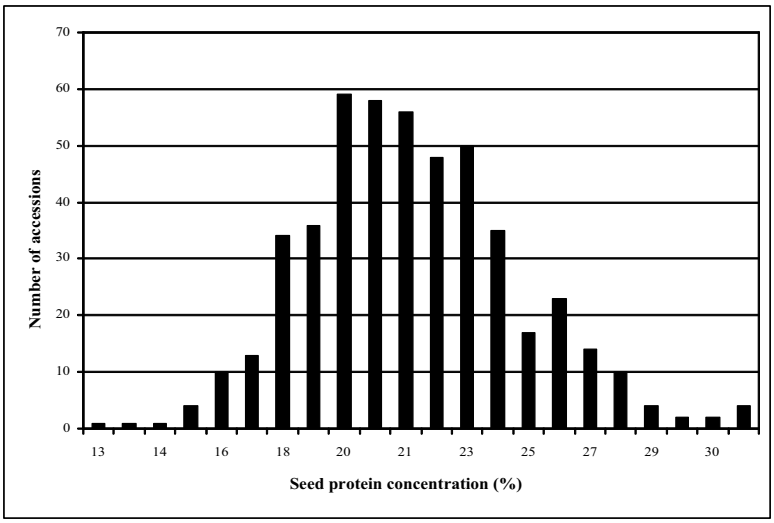

Fig. 1: Frequency histogram of total seed protein concentrations found in the USDA Pisum core collection (n=480).

Table 1. Summary of selected characteristics for those accessions in the USDA *Pisum* core collection that exhibited the ten highest and ten lowest seed protein concentrations

Accession	Taxon	Seed % Protein	Country	Flower	Cotyledon	Node to First Flower	Seed Surface
PI 357292	*Pisum sativum*	30.93	Yugoslavia	white	green	9-10	wrinkled
PI 343978	*Pisum sativum* subsp. *elatius*	30.77	Turkey	pigmented	yellow	12-15	round
PI 137118	*Pisum sativum*	30.51	Canada	pigmented	yellow	13-17	round
PI 288024	*Pisum sativum*	30.4	France	white	green	6	round
PI 102887	*Pisum sativum*	29.81	China	white	yellow	9-10	round
PI 165949	*Pisum sativum*	29.75	India	pigmented	yellow	10-13	round
PI 261671	*Pisum sativum*	29.08	Netherlands	white	yellow	10-13	round
PI 125840	*Pisum sativum*	28.51	Afghanistan	pigmented	yellow	13-16	round
PI 272207	*Pisum sativum*	28.08	Greece	pigmented	yellow	no data	round
PI 103709	*Pisum sativum*	27.95	India	white	green	10-11	round
PI 203944	*Pisum sativum*	15.48	Mexico	white	yellow	16-19	round
PI 358610	*Pisum sativum* subsp *abyssinicum*	15.28	Ethiopia	pigmented	yellow	10	round
PI 324706	*Pisum sativum*	15.21	Romania	pigmented	yellow	17-21	round
PI 204306	*Pisum sativum*	14.88	Australia	pigmented	yellow	20-22	round
PI 358623	*Pisum sativum*	14.67	Ethiopia	pigmented	yellow	15-18	round
PI 204307	*Pisum sativum*	14.62	Australia	white	yellow	11-14	round
PI 134271	*Pisum sativum*	14.37	Afghanistan	pigmented	yellow	12-14	round
PI 188698	*Pisum sativum*	13.87	Nigeria	pigmented	green	15-20	round
PI 356986	*Pisum sativum*	13.20	India	pigmented	yellow	12-13	round
PI 222071	*Pisum sativum*	12.38	Afghanistan	pigmented	yellow	12-19	round

Table 2. Taxon summary of the pea seed protein concentrations measured in the USDA *Pisum* core collection.

Taxon	Number of accessions	Protein minimum	Protein maximum	Protein mean (s.e.)
Pisum sativum	437	12.38	30.93	20.93 (0.15)
Pisum sativum ssp *elatius*	27	17.62	30.77	22.25 (0.56)
Pisum sativum ssp *abyssinicum*	713	15.28	23.11	19.39 (0.71)
Pisum sativum subsp *arvense*	2	18.17	23.59	20.88
Pisum sativum var *pumillio*	1	n.a.	n.a.	19.34

Discussion

Protein concentration in many biological tissues, such as seeds, is often measured indirectly as percent nitrogen. Protein concentration is then calculated from the nitrogen value using a nitrogen-to-protein conversion factor. For many foods, a factor of 6.25 is generally used, a value based on the average nitrogen percentage of a mix of common amino acids (11). However, because different amino acids vary in their nitrogen percentages, and different proteins contain varying mixtures of amino acids, it is more accurate to use a conversion factor that is based on the specific proteins contained in a given food (3, 6, 11). For this reason, a conversion factor of 5.44 was used in this study for all protein calculations, a value derived from actual amino acid analyses of several pea genotypes (6).

The seed protein concentration analysis of the USDA *Pisum* core collection revealed values ranging from 30.93% to 12.38% (Table 1). These values are comparable to the values of 34.1% to 14.5% for *Pisum sativum* seed obtained by Savage and Deo (8), especially when their values are recalculated with a 5.44 conversion factor instead of the 6.25 factor that was used (their recalculated range would be 29.7% to 12.6%). Similarly, our values for *Pisum sativum* subsp. *abyssinicum* accessions overlap with those of Yemane and Skjelvåg (12), who found

concentrations of 19.9% and 20.5% (recalculated with 5.44 conversion factor) for whole seed of two *abyssinicum* cultivars.

The 2.5-fold range reported here (30.93% to 12.38 %) is in contrast to a field study conducted with 1071 USDA accessions in 1975 (4), in which only a 1.4-fold variation in seed protein concentration was found (26.9% to 19.7%; recalculated with 5.44 conversion factor). Studies of seed protein concentration levels in a number of legume crops have shown extreme sensitivity to the environment (5). The controlled conditions of the greenhouse culture versus the field environment may be a possible explanation for the differences in the two studies using USDA pea germplasm. A slightly higher, three-fold variation was found in a field study of 255 accessions held in the John Innes pea germplasm collection (5). In that report (5), the protein values (presented as α-amino nitrogen) were negatively skewed (i.e., more to the lower concentrations; their Fig. 33.2), unlike the more symmetrical distribution found in the present study (Fig. 1), and this appears to account for their broader variation in protein values. Interestingly, the accessions analyzed in that study also represented a greater diversity of *Pisum* subspecies. No accessions, for instance, of *Pisum fulvum* or *Pisum sativum* subsp. *transcaucasicum* are included in the current seed protein concentration dataset (Table 2).

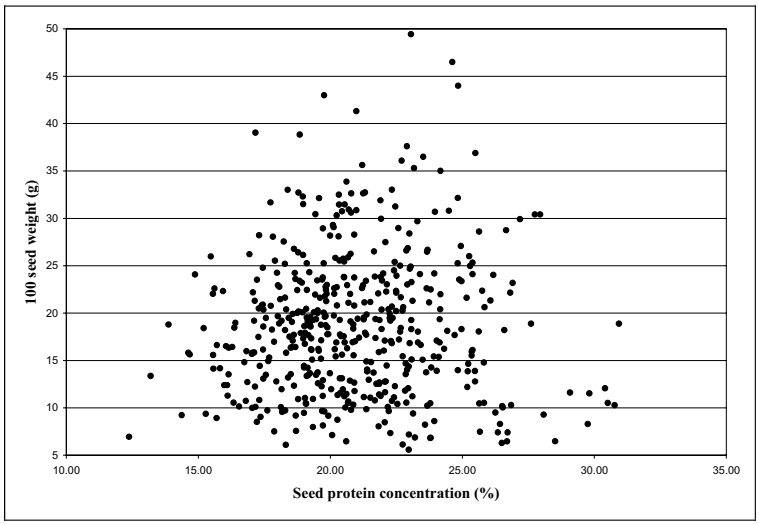

Figure 2. Correlation between seed protein concentration and seed size (g of 100 seed) in the USDA Pisum core collection.

Jermyn and Slinkard (4) presented field data demonstrating that as protein increased, yields decreased, when they assessed the USDA pea collection in the 1970's. Although yield was not measured in the current study, no correlation was found between the related trait, seed size, and seed protein concentration (Fig. 2). Additionally, a small replicated trial conducted in one location and one year (2004) of high yielding cultivars now in production in Washington State indicates most have seed protein concentrations in the higher range (~ 25 to 28%) (Coyne, unpublished). Thus, it appears that seed protein concentration can be enhanced independently of yield and/or seed size in pea.

The original USDA *Pisum* core collection was selected using only geography (country of origin) and flower color as trait variables (9). Seed protein concentration was not used in the selection process. Nonetheless, based on the seed protein concentrations reported in this investigation (Fig. 1, Tables 1, 2), relative to other studies (4, 5), it would appear that the core adequately represents seed protein variability found in pea. Therefore, these protein values have recently been included with other trait data to reduce the size of the original core to achieve a desired 10% representation of the entire USDA *Pisum* collection (2).

Several QTL studies on the heritability of pea seed protein concentration (9), along with recent studies using *Medicago truncatula* to study legume seed proteins (3), are increasing our understanding of the genetics of this important component of pea seeds. The accessions in Table 2 may aid in the future discovery of useful alleles for breeding enhanced seed protein levels in this crop, or could prove valuable for mapping novel regulatory genes associated with increased seed protein concentration in pea.

Acknowledgments: This work was supported by USDA-ARS Projects # 5348-21000-020-00D (Coyne) and # 6250-21520-042-00D (Grusak), and through USDA funds provided by the *Pisum* Crop Germplasm Committee (to Grusak).

1. Casey, R. and Domoney, C. 1985. In: Hebblethwaite, P.D., Heath, M.C. and Dawkins, T.C.K. (eds.) The Pea Crop. Butterworths, London, pp 359-368.
2. Coyne, C.J., Brown, A.F., Timmerman-Vaughan, G.M., McPhee, K.E. and Grusak, M.A. 2005. Pisum Genetics 37: 1-4.
3. Djemel, N., Guedon, D., Lechevalier, A., Salon, C., Miquel, M., Prosperi, J.-M., Rochat, C. and Boutin, J.-P. 2005. Plant Physiol. Biochem. 43: 557-566.
4. Jermyn, W.A. and Slinkard, A.E. 1977. Legume Res. 1: 33-37.
5. Mathews, P. and Arthur, E. 1985. In: Hebblethwaite, P.D., Heath, M.C. and Dawkins, T.C.K. (eds.) The Pea Crop. Butterworths, London, pp. 369-381.
6. Mossé, J. 1990. J. Agric. Food Chem. 38: 18-24.
7. SAS 9.1. 2002-2003. SAS Institute, Cary, NC, USA.
8. Savage, G.P. and Deo, S. 1989. Nutr. Abstr. Rev. Ser. A 59: 65-88.
9. Simon, C.J. and Hannan, R.M. 1995. HortSci. 30: 907.
10. Tar'an, B., Warkentin, T., Somers, D.J., Miranda, D., Vandenberg, A., Blade, S. and Bing, D. 2004. Euphytica 136: 297-306.
11. Tkachuk, R. 1977. In: Hulse, J.H., Rachie, K.O. and Billingsley, L.W. (eds.) Nutritional Standards and Methods of Evaluation for Food Legume Breeders. International Development Research Centre, Ottawa, Canada, pp. 78-82.
12. Yemane, A. and Skjelvåg, A.O. 2003. Plant Foods Human Nutr. 58: 275–283.