

**Genetic variation and relationships of total seed protein content with some agronomic traits in pigeonpea (*Cajanus cajan* (L.) Millsp.)**

Jimmy Obala<sup>1,2</sup>, Rachit K. Saxena<sup>1,\*</sup>, Vikas K. Singh<sup>1</sup>, Suryanarayana Vechalapu<sup>1</sup>, Roma Das<sup>1</sup>, Abhishek Rathore<sup>1</sup>, Chanda V Sameer-Kumar<sup>1</sup>, Kulbhushan Saxena<sup>1</sup>, Pangirayi Tongoona<sup>2</sup>, Julia Sibiya<sup>2</sup>, Rajeev K. Varshney<sup>1</sup>

<sup>1</sup>International Crops Research Institute for the Semi-Arid Tropics, Patancheru 502 324, Hyderabad, India

<sup>2</sup>University of KwaZulu-Natal, Pietermaritzburg, South Africa

\*Corresponding author: [r.saxena@cgiar.org](mailto:r.saxena@cgiar.org)

**Abstract**

Seed protein content (SPC) is an important grain quality trait, which impacts the nutritional importance of pigeonpea seed in the diet of over a billion people globally. The present study was carried out to determine variation in SPC and its relationships with some agronomic traits among 23 parental lines of different types of pigeonpea mapping populations. The parental lines were evaluated under field conditions during 2014-2015 growing season. A randomised complete block design in two replications was used. Data were recorded on SPC, days to first flower (DTF), plant height at maturity (PltH), number of pods per plant (NPP), number of seeds per pod (NSP), hundred-seed weight (SW) and seed yield per plant (SY). There were significant differences among genotypes for all traits. Broad-sense heritability was 0.693 for SPC but ranged from 0.519 (NPP) to 0.999 (DTF) while genetic advance was 2.4% for SPC but ranged from 1.2 % (NSP) to 141.2 % (SY), and genetic gain ranged from 11.0 % (SPC) to 230.0 % (SY). Simple correlation showed that SPC is only significantly but negatively correlated with SW ( $r = -0.30$ ,  $P < 0.05$ ), while path analyses revealed that SPC is negatively associated SW and NPP but positively with DTF, PltH, NSP and SY. It is concluded that genetic variation for SPC and agronomic traits exist among pigeonpea genotypes studied. The variation is accompanied by both favourable and unfavourable relationships of SPC with the agronomic traits.

**Keywords:** *Cajanus cajan*, dietary protein, genetic advance, correlation, path analysis, heritability.

**Abbreviations:** DTF\_days to first flower, GA\_genetic advance, GCV\_genotypic coefficient of variation, ICRISAT\_International Crops Research Institute for the Semi-Arid Tropics, IL\_introgression line, MAGIC\_multiparent advanced generation intercross, NAM\_nested association mapping, NPP\_number of pods per plant, NSP\_number of seeds per pod, PCV\_phenotypic coefficient of variation, PltH\_plant height, RIL\_recombinant inbred line, SPC\_seed protein content, SW\_100-seed weight, SY\_seed yield per plant.

**Introduction**

Among food plants, grain legumes are a major source of dietary protein in the developing world (Iqbal et al., 2006; Akibode and Maredia, 2011). For sustained supply of dietary protein, there is need not only to improve the agronomic practices but also to use crop cultivars which give reliable yields even under severe conditions (Foley et al., 2011). In the scenario mentioned above, pigeonpea seems to be a promising crop as it is tolerant to heat and drought, and can give relatively better yields in marginal soils than any other food legume (Rao et al., 2010). Pigeonpea is a sub-tropical and tropical grain legume, which has diverse uses including source of food, feed, fodder, building material and fuel wood. Pigeonpea also contributes to biological nitrogen fixation (Rao et al., 2010). It is a cash crop that supports the livelihoods of millions of resources-poor farmers in Asia and Africa (Mula and Saxena, 2010). Global annual production stands at ~4.5 million ton from ~5.4 million ha (FAOSTAT, 2017). Despite the importance of pigeonpea as a source of dietary protein to more than one billion people globally, breeding objectives in the crop have for a long time almost

entirely focused on increasing yield and crop adaptability (Odeny, 2007; Mligo and Craufurd, 2005; Upadhyaya et al., 2007). Very little or no attention has been targeted at genetic enhancement of the nutritional quality such as seed protein content (SPC) of pigeonpea. A critical step to genetic improvement of any trait is the understanding of the level of genetic variation in the trait and the extent to which the trait relates with other traits of agronomic importance within a target set of genetic materials. A few studies have investigated variation of SPC in pigeonpea (Saxena et al., 2002; Upadhyaya et al., 2007; Sawargoankar, 2010). Upadhyaya et al. (2007) reported mean SPC ranging from 19.7 to 20.3% among 310 germplasm collections while Sawargoankar (2010) reported a range of 17.4 to 23.0% among 37 parental lines. Environmental effects on SPC have been reported to be generally large but genotype  $\times$  environment interactions (GEI) are often small, with relative differences between genotypes being similar in several environments (Baudoin and Maquet, 1999; Saxena et al., 2002). There are also few reports on the relationships of

SPC with agronomic traits in pigeonpea (Saxena et al., 1987; Rekha et al., 2013). Saxena et al. (1987) reported highly significant negative or positive correlations or no correlations between SPC and SW among 1,974 single F<sub>7</sub> plants from intergeneric crosses of pigeonpea. Similarly, Rekha et al. (2013) in an evaluation of 40 pigeonpea genotypes found a significant though small negative correlations between SPC and NSP, while correlation of SPC with SW, NPP and PltH were small, positive and non-significant. In other legume crops, Kulwal and Mhase (2017) observed strong positive correlation between SPC and SW in chickpea, while Lawn and Rebetzke (2006) reported significant negative correlation between SPC and SW in mungbean. Significant correlation between SPC and PltH also has been reported in pea (Burstin et al. 2007). The different reports on variation in SPC and its relationships with agronomic traits in pigeonpea and other legume crops show that the trait's variates including variance, heritability, and its relationships with other traits depend upon set of genotypes evaluated and environment in which they are tested. This therefore warrants continuous assessment of germplasms for variation in SPC and its relationship with agronomic traits of importance before applying specific materials in genetic improvement programs. To study trait variations, measures such as phenotypic and genotypic coefficients of variation (PCV and GCV, respectively) are often used in addition to estimates of heritability and genetic advance (GA). On the other hand simple correlation and path coefficient analyses are used for studying interrelationships among traits. Simple correlation indicates how change in the variance of one trait affects the change in the variance of the other trait regardless of cause and effect relationship. Unlike simple correlation, path coefficient analysis helps to measure the direct effect of one trait on another by separating correlation coefficient into direct and indirect components, which enables detection of the most influential traits. At ICRISAT, Patancheru, India, a number of different types of mapping populations including introgression lines (IL), nested association mapping (NAM), multiparent advanced generation intercross (MAGIC) and recombinant inbred lines (RIL) are being developed in pigeonpea for the identification of quantitative trait loci (QTLs) and molecular markers for various traits. These populations can also be used for dissecting genetic control of SPC, including marker-SPC associations, or even to directly select for lines with improved levels of the trait. However, the variation of SPC and its relationships with important agronomic traits among the parental lines are not known. The objectives of this study were to (i) characterize variability for SPC among 23 pigeonpea parental genotypes, and (ii) estimate correlation and path coefficients of SPC with seed yield and yield-related characters including seed weight, pod characters, plant height and days to flower.

## Results

### **Performance of genotypes for seed protein content and agronomic traits**

Mean square for each of the seven studied traits is presented in Table 1. Highly significant ( $p \leq 0.01$ ) differences existed among the 23 genotypes for all seven traits, and

therefore genotype means were compared to determine differences. Mean, range and coefficient of variation (CV) values are presented in Table 2. Mean SPC in the present study ranged from 19.3% (ICPL 87119) to 25.5 % (HPL 31) with an overall mean value of 22.1%. Of the 23 genotypes, HPL 24, ICP 14486, ICP 5529, HPL 28 and HPL 31 recorded relatively high SPC while genotypes ICPL 87, ICPL 20097, ICPL 85063, ICP 99050 and ICPL 87119 recorded low SPC in that order. For the agronomic traits, days to first flower (DTF) ranged from 48.0 days (MN 1) to 156.0 days (ICPL 332) with a mean of 100.0 days (Table 2). Plant height (PltH) ranged from 67.5 cm (MN 1) to 230.0 cm (ICPL 20097) with an average of 179.7 cm (Table 2). Number of pods per plant (NPP) ranged from 31.7 (MN 1) to 582.3 (HPL 24) with a mean of 229.1 while number of seeds per pod (NSP) ranged from 2.9 to 4.6 with a mean of 3.5 (Table 2). Hundred seed weight (SW) ranged from 6.2 g/plant (ICP 7426) to 20.8 g/plant (ICP 7035) with a mean of 10.1 g/plant while seed yield per plant (SY) varied from 7.9 g/plant (MN 1) to 333.4 g/plant (ICP 7035) with a mean of 61.2 g/plant. The highest CV was shown by NPP followed by NSP, SY, SW, PltH, SPC and DTF. The relatively low CV values across traits (Table 2) is expected because the genotypes used in the study are highly inbred landraces or breeding lines.

### **Heritability, genotypic and phenotypic coefficients of variation and genetic gain**

Whereas the mean, range and CV suggest the extent to which improvement can be made for a given trait they, however, depict nothing about effect of genotype on trait variation. Hence, in the present study, parameters such as genotypic ( $\sigma^2G$ ), environmental ( $\sigma^2E$ ) and phenotypic ( $\sigma^2P$ ) variances, GCV and PCV, broad-sense heritability ( $H^2$ ) and genetic gain were estimated (Table 3).

In general,  $\sigma^2G$  and GCV were always close to  $\sigma^2P$  and PCV, respectively, with  $\sigma^2G$  always larger than  $\sigma^2E$  for all traits. This was also consistent with the generally high  $H^2$  ranging from 0.519 for NPP to 0.999 for DTF (Table 3). There were small differences between PCV and GCV values for SPC and most of the other traits except NPP (Table 3). Although SPC showed high  $H^2$  estimate ( $>0.60$ ), the GCV and GA were low resulting in a relatively low genetic gain estimate for the trait. High  $H^2$  with high GCV and high or moderate GA estimates for DTF, PH, SW, SY, NPP and NSP resulted in  $>50$  % genetic gain (Table 3).

### **Genotypic correlations and path analysis between SPC and agronomic characters**

Results of simple genotypic correlations between SPC and agronomic traits is presented in Table 4. Generally, SPC had negative correlations with all traits being significant only with SW (Table 4). On the basis of path coefficient analysis results (Table 5), all values of direct effects were below one, showing that increments resulting from multicollinearity were marginal. The values of direct path coefficient were relatively large and negative between SPC and NPP (-0.73) and SPC and SW (-0.68). It was positive and large between SPC and SY (0.63) but small between SPC and DTF (0.08), SPC and PH, and SPC and NSP. Indirect effects of agronomic

**Table 1.** Mean squares for seed protein content and six agronomic traits in 23 pigeonpea genotypes.

Trait	Mean square	
	Genotype (DF = 22)	Error (DF = 22)
Seed protein content (%)	4.9 ***	0.9
Days to first flower	1754.6 ***	0.1
Plant height (cm)	5151.2 ***	10.4
Number of pods per plant	13782.0 ***	4359.0
Number of seeds per pod	1.1 ***	0.2
100-seed weight (g)	18.8 ***	0.6
Seed yield (g)	9534.2 ***	59.2

DF, Degrees of freedom, \*\*\* significant at P = 0.001.

**Table 2.** Mean, range and coefficient of variation for seed protein content and six agronomic characters studied in 23 pigeonpea genotypes.

Genotype	SPC (%)	DTF	PltH (cm)	NPP	NSP	SW (g)	SY (g)
HPL 31	25.5	100.0	188.3	167.8	2.95	9.9	38.6
HPL 28	25.2	101.0	191.7	310.5	3.51	10.0	41.2
ICP 5529	24.6	103.7	210.0	152.8	4.37	8.6	23.3
ICP 14486	24.1	86.0	133.3	31.7	4.10	8.6	10.2
ICP 14209	23.1	138.0	208.0	212.2	3.57	8.7	17.3
HPL 24	23.0	111.8	208.3	582.3	2.93	8.1	152.7
ICP 8863	22.3	90.2	210.0	124.8	3.47	9.9	24.5
ICPL 85010	22.2	50.3	82.5	76.5	3.41	9.0	18.0
ICPL 88039	22.2	60.5	149.2	178.8	3.62	11.5	54.3
MN 1	22.2	48.0	67.5	37.7	3.02	7.2	7.9
ICP 7426	22.1	120.0	205.0	363.3	3.48	6.2	70.2
ICPL 20096	22.1	131.3	215.8	162.0	3.22	13.1	48.2
UQ 50	21.9	106.8	204.2	352.5	3.57	13.6	127.4
ICP 28	21.6	79.8	128.3	132.5	2.91	8.6	26.2
ICP 11605	21.5	66.0	93.3	53.2	3.38	12.2	22.5
ICP 7035	21.3	129.0	226.7	517.7	4.60	20.8	333.4
ICPL 332	21.3	156.0	228.3	124.2	3.03	15.6	36.9
ICPB 2049	20.8	102.0	206.7	131.8	3.37	9.9	37.4
ICPL 87	20.8	68.7	116.7	219.8	3.77	11.1	69.2
ICPL 20097	20.7	151.3	230.0	251.3	3.35	11.8	78.0
ICPL 85063	20.4	92.5	228.3	375.0	3.38	9.6	67.3
ICPL 99050	20.2	104.8	210.0	264.2	3.35	10.3	64.1
ICPL 87119	19.3	103.2	191.7	445.7	3.75	11.5	38.5
Mean	22.1	100.0	179.7	229.1	3.50	10.7	61.2
S.e.m	0.5	0.5	3.1	26.9	0.20	0.4	8.8
Range	19.3-25.5	48.0-156.0	67.5-230.0	31.7-582.3	2.9-4.6	6.2-20.8	7.9-333.4
CV (%)	4.3	1.1	4.8	28.8	13.9	7.8	12.6
LSD 5%	1.4	1.3	8.5	75.5	0.55	1.1	24.6

CV, Coefficient of variation; S.e.m, Standard error of mean; SPC, Seed protein content; DTF, Days to first flower; PltH, Plant height; NPP, Number of pods per plant; NSP, Number of seeds per pod; SW, Hundred-seed weight; SY, Seed yield.

**Table 3.** Estimates of broad-sense heritability, genotypic and phenotypic coefficients of variation, and genetic gain for seven traits in 23 pigeonpea genotypes.

Trait	$\sigma^2G$	$\sigma^2E$	$\sigma^2P$	$H^2$	GCV (%)	PCV (%)	GA	GG (%)
DTF	877.25	0.07	877.32	0.999	29.6	29.6	61.0	60.9
PltH	2570.40	10.44	2580.84	0.996	28.2	28.3	104.2	58.2
SW	9.10	0.61	9.71	0.937	28.3	29.2	6.0	56.4
SY	4737.52	59.20	4796.72	0.988	112.5	113.2	141.0	230.4
SPC	2.01	0.89	2.90	0.693	6.4	7.7	2.4	11.0
NPP	4711.50	4359.00	9070.50	0.519	30.0	41.6	101.8	713.4
NSP	0.50	0.20	0.70	0.712	19.2	23.6	1.2	65.5

DTF, Days to first flower; PltH, Plant height; NPP, Number of pods per plant; SW, Hundred-seed weight; SY, Seed yield; SPC, Seed protein content;  $\sigma^2G$ , Genetic variance;  $\sigma^2E$ , Environmental variance;  $\sigma^2P$ , Phenotypic variance;  $H^2$ , Broad-sense heritability; GCV, Genotypic coefficient of variability; PCV, Phenotypic coefficient of variability; GA, Genetic advance; GG, Genetic gain.

**Table 4.** Genotypic correlation coefficients for pair-wise association of seed protein content with agronomic traits.

	PltH	NPP	NSP	SW	SY	SPC
DTF	0.85***	0.41**	0.06 <sup>NS</sup>	0.39**	0.33*	-0.07 <sup>NS</sup>
PltH		0.56***	0.15 <sup>NS</sup>	0.32*	0.38**	-0.07 <sup>NS</sup>
NPP			0.20 <sup>NS</sup>	0.27 <sup>NS</sup>	0.73***	-0.27 <sup>NS</sup>
NSP				0.43***	0.46**	0.00 <sup>NS</sup>
SW					0.68***	-0.30*
SY						-0.20 <sup>NS</sup>

NS, not significant at 0.1 probability level; \*, \*\*, and \*\*\* significant at the 0.05, 0.01, 0.001 probability levels, respectively; DTF, Days to first flower; PltH, Plant height; NPP, Number of pods per plant; SW, Hundred-seed weight; SY, Seed yield; SPC, Seed protein content.

**Table 5.** Direct (boldfaced main diagonals) and indirect path (off-diagonals) coefficient values of seed protein content against agronomic traits of pigeonpea.

	Trait						
	DTF	PltH	NPP	NSP	SW	SY	
DTF	<b>0.08</b>	0.19	-0.30	0.01	-0.27	0.21	
PltH	0.07	<b>0.23</b>	-0.41	0.02	-0.22	0.24	
NPP	0.03	0.13	<b>-0.73</b>	0.02	-0.18	0.46	
NSP	0.00	0.04	-0.15	<b>0.11</b>	-0.29	0.29	
HSW	0.03	0.07	-0.20	0.05	<b>-0.68</b>	0.43	
SY	0.03	0.09	-0.53	0.05	-0.47	<b>0.63</b>	

DTF, Days to first flower; PltH, Plant height; NPP, Number of pods per plant; NSP, Number of seeds per pod; SW, Hundred-seed weight; SY, Seed yield.

**Table 6.** Features of pigeonpea genotypes evaluated for seed protein content and some agronomic traits.

Accession	Features	Source population
HPL 28	High seed protein content breeding line	-†
HPL 31	High seed protein content breeding line	-
ICPL 87119 (Asha)	Genome sequence available, leading variety, resistant to Fusarium wilt (FW) and sterility mosaic disease (SMD)	IL, NAM
ICP 7426	High pod numbers, medium duration	MAGIC
HPL 24	High protein content, medium duration, compact, susceptible to FW and resistant to SMD, inter-specific derivative	MAGIC, NAM
ICP 11605	Early flowering, germplasm line	MAGIC
ICP 14209	High number of pods, germplasm line	MAGIC
ICP 14486	Early flowering, germplasm line	MAGIC
ICP 5529	Medium duration, obcordate leaves, compact plant, poor yielding, modified flower	MAGIC
ICP 7035	Medium duration, SMD resistant to both Patancheru and Bangalore races, large purple seed, high sugar	MAGIC, NAM
ICP 8863	Erect, mid-late, highly resistant to FW and susceptible to SMD, red seeded genotype	MAGIC, NAM, RIL
ICPL 87	Early duration, determinate, short, high combiner	NAM
ICPL 88039	Extra early maturity, indeterminate, good yield	NAM
ICP 85063 (Lakshmi)	Medium duration, indeterminate, good yield, more branching	NAM
MN-1	Super early, small seeded, determinate	NAM
ICP 28	Early maturity, local varieties	NAM
ICP 85010 (Sarita)	Early maturity, local varieties	NAM
UQ 50	Determinate, long-podded, white seeded	NAM
ICPL 20096	Resistant to FW and SMD	RIL
ICPL 20097	Resistant to both SMD and FW	RIL
ICPL 332	Tolerant to pod borer, high yielding.	NAM
ICPB 2049	Susceptible to FW	RIL
ICPL 99050	Resistant to FW	NAM

† Not a parent in any population; IL, Introgression line; NAM, Nested association mapping population; MAGIC, Multiple parent advanced generation intercross; RIL, Pigeonpea recombinant inbred line population.

traits on SPC were large and negative for NPP via SY, but positive for SW also via SY.

## Discussion

The knowledge of genetic variation for a trait and trait correlations are important components of any breeding

objective. SPC in pigeonpea is an important grain quality trait, and it impacts the nutritional importance of pigeonpea in the human diet. The range of SPC values obtained in the present study is within 12.0 to 30.0 % reported earlier among 1,974 germplasm accessions at ICRISAT (Remanandan et al., 1988). It is also close to 15.9 to 24.1% reported recently among 300 germplasm collection from

different altitudes of Kenya (Upadhyaya et al., 2007). In general, the range in SPC values of 19.3 to 25.5% among the 23 genotypes may also be a reflection of the low genetic diversity within the cultivated pigeonpea gene pool that has been repeatedly reported (Saxena et al., 2002; Saxena et al., 2014). Among the genotypes tested in this study, interspecific derivatives (HPL 24, HPL 28, HPL 31) from the cross between wild (*C. scarabaoiedes*) and cultivated (*C. cajan*) pigeonpea (Saxena et al., 2002) showed the highest SPC. This suggests that the wild accessions or their interspecific progenies could provide the needed source of high SPC genes for trait improvement. Landrace cultivars that showed comparable level of SPC included ICP 5529 (24.6%) and ICP 14486 (24.1%), and they are equally potential sources of desirable genes for improving SPC.

The significant differences among pigeonpea genotypes in the present study indicate presence of variability for all traits measured. This is supported by the generally  $H^2$  indicating influence of genetic factors on phenotype. Whereas heritability estimates can be used to predict the reliability of the phenotypic value as a guide to breeding value (Falconer and Mackay, 1996), heritability alone does not reveal the extent of response to selection.  $H^2$  along with GCV and GA provide reliable estimates of the amount of genetic gain to be expected through phenotypic selection (Burton, 1952). The combination of high  $H^2$ , GCV, GA and genetic gain (%) for DTF, PltH, SY and NPP indicates that the variation in these traits is largely due to genetic factors, and selection would be effective for these traits. However, SPC as a core trait in this study had high  $H^2$  but low GCV and low genetic gain estimates, depicting a low response to selection. Similarly, SW and NSP with high and moderate  $H^2$ , respectively, had low genetic gain values indicative of a poor response to selection. Given the poor predicted response to selection based on SPC alone, determining the relationships of SPC with agronomic traits could provide an indication of which of the agronomic traits could be used to indirectly select for improved SPC. It could also pinpoint which of the agronomic traits affect SPC either positively or negatively, which in turn helps in deciding on appropriate selection or breeding strategy.

Few studies have been conducted on the relationship of SPC with agronomic traits in pigeonpea. Results of simple genotypic correlations in the present study indicated that SW was the major trait that negatively influenced SPC in the set of genotypes tested. This observation is in agreement with that of earlier studies in pigeonpea (Saxena et al., 1987), soybean (Filho et al., 2001), mungbean (Afzal et al., 2003), and cowpea (Asante et al., 2004) who reported significant negative correlations between SPC and SW but contrasts with results of other similar studies which reported either positive or no correlation between the two traits (Saxena et al., 1987; Filho et al., 2001; Rekha et al., 2013).

If only simple genotypic correlations were considered in the present study, SW would be the only agronomic trait that influences SPC but negatively in the set of pigeonpea genotypes tested. However, path analysis allocated the strongest negative direct effects on SPC to NPP and SW indicating that selection for increased NPP or SW would lead to reduced SPC. On the other hand the strong positive direct effect due to SY indicates that simultaneous selection for

high SPC and high SY is possible, and is in agreement with conclusions from previous studies that selection for high SPC does not always lead to SY reduction in the grain legumes (Leleji et al., 1972; Brim and Burton, 1978; Wilcox and Cavins, 1995). Similarly, through path coefficient analysis, a large negative indirect effect of NPP on SPC via SY was detected indicating that simultaneous selection for high NPP and SY would lead to reduced SPC. In a similar manner, SW had a large positive indirect effect on SPC also via SY indicating that simultaneous selection for increased SW and SY would lead to increased SPC.

## Materials and methods

### Plant material and field evaluation

Twenty three genotypes that were used in this study are presented in Table 6. Twenty-one of the genotypes are parents of different types of mapping populations, including IL, MAGIC, NAM, and RIL being developed at ICRISAT, Patancheru, India. These parental lines are of unknown SPC, except HPL 24, which together with two other genotypes, namely HPL 31 and HPL 26 are known high SPC breeding lines developed from the cross between *Cajanus cajan* variety 'Baigani' and an accession of *C. scarabaoiedes* (a wild relative of pigeonpea). HPL 31 and HPL 26 were included in the present study to allow comparison with parental lines of the mapping populations. To determine suitability of the mapping populations for genetic dissection of SPC and its relationships with agronomic traits, the 21 parental lines and the two high SPC breeding lines were evaluated under field conditions at ICRISAT, India. The experiment was laid out in a randomized complete design with two replications. Each of the 23 genotypes was planted in a single 4 m long row with inter- and intra-row spacing of 75 cm and 30 cm, respectively. All cultural practices were carried out as routinely done at ICRISAT.

### Estimation of seed protein content

To estimate SPC, 10 g of mature dry clean seeds of each plant were analyzed at the Central Analytical Services Laboratory at ICRISAT, India. Before grinding, seeds were oven-dried at 60°C for 48 hours.

The dried seed samples were ground into powder in a mill with Teflon chambers. The ground samples were again kept in an oven at 60°C overnight. Samples and appropriate blanks were digested simultaneously in duplicate (i.e. two independent analyses) using tri-acid digestion procedure as described in Upadhyaya et al. (2016). Briefly, 1.0 g of the ground seed sample was transferred to a 75 ml digestion tube containing 10 ml of tri-acid mixture of nitric, sulfuric and perchloric acids in the ratio of 10:0.5:2 (v/v). The contents were cold-digested overnight in a digestion chamber. Colorless and clear digest were obtained by keeping the samples at 120°C for 1 hour followed by digestion at 230°C for 2 hours. After cooling, the digests were dissolved in distilled water and volume topped up to 75 ml and then mixed well by shaking. Aliquots were obtained from the digests and used to estimate the total nitrogen (N) using a San++Automated Wet Chemistry

Analyzer (Skalar, Breda, The Netherlands). Seed protein content of a sample was estimated by multiplying its N (%) content by factor 6.25.

### Scoring for agronomic traits

Besides SPC, data were also collected on DTF, PltH, NPP, NSP, SW and SY per plant. The DTF was scored daily as described in Craufurd et al. (2001). Plant height was recorded as height in cm from the base to the tip of the plant. Number of pods per plant and number of seeds per pod were recorded as counts of number of pods on a plant and number of seeds per pod, respectively. Hundred-seed weight was recorded as weight of 100 dry, clean and healthy seeds in g, and SY was obtained by weighing all seeds from a plant in g.

### Data analysis

#### Genotypic and phenotypic variation

All statistical analyses were performed using SAS statistical software v9.4 (SAS Institute, 2015). Analysis of variance was carried out, and means were separated using Least Significance Difference (LSD) at 5%. Genotypic and phenotypic coefficients of variation were calculated as described in Singh and Chaudhary (1979) as follows:

$PCV(\%) = (\sqrt{\sigma^2 P} / \mu) \times 100$ , and  $GCV(\%) = (\sqrt{\sigma^2 G} / \mu) \times 100$ , where  $PCV$  and  $GCV$  are the phenotypic and genotypic coefficients of variation, respectively, and  $\sigma^2 P$  and  $\sigma^2 G$  are the phenotypic and genotypic variances, respectively. Phenotypic and genotypic coefficients of variations were categorized as low (<10%), moderate (10-20%), and high (>20%) (Subramanian and Menon, 1973).

Broad-sense heritability ( $H^2$ ) was estimated using the formula:  $H^2 = (\sigma^2 G / \sigma^2 P)$ , where  $\sigma^2 G$  and  $\sigma^2 P$  are genotypic and phenotypic variances respectively. The heritability was placed into three categories of low (0-0.3), moderate (0.3-0.6) and high (>0.6) (Johnson et al., 1955).

Genetic advance (GA) was obtained as:  $GA = H^2 \times \sqrt{\sigma^2 P} \times K$ , where  $H^2$  is the broad-sense heritability,  $\sigma^2 P$  is the phenotypic standard deviation and  $K$  is the selection differential (2.06 at 5%). GA was converted to percent genetic gain as:  $Genetic\ gain = GA \times 100$ , and categorized as low (0-10 %), moderate (10-20%) and high (>20%) (Johnson et al., 1955).

#### Genetic correlation and path analyses

Genotypic correlations were calculated according to Falconer and Mackay (1996) using the formula:  $r_G = \sigma_{Gxy} / (\sqrt{\sigma^2 Gx \times \sigma^2 Gy})$ , where  $\sigma_{Gxy}$  is genotypic covariance and  $\sigma^2 Gx$  and  $\sigma^2 Gy$  are genotypic variances of trait x and trait y, respectively.

Direct and indirect path coefficients were calculated using genotypic correlation coefficients following methods of

Wright (1921), with SPC considered as a response variable and DTF, PltH, NPP, NSP, SW and SY as causal variables.

### Conclusion

There is variation for SPC among the pigeonpea genotypes used as parents of the mapping populations at ICRISAT although no large differences were detected, which is a possible reflection of the low genetic diversity that has repeatedly been reported within the cultivated pigeonpea gene pool. Although the  $H^2$  and GCV for SPC were large, the genetic advance estimate was low resulting in low expected genetic gain. Nonetheless there is possibility of deriving desirable recombinants from biparental matings. Both favourable and unfavourable relationships exist between SPC and some of the agronomic traits with strong negative relationships of SPC with NPP and SW, which indicates that simultaneous selection for both high NPP and heavier seeds, or both NPP and high SY would lead to reduction in total SPC. However, simultaneous selection for high SY and high SPC, or for both high SW and high SY could result in increased SPC. An understanding of the genetic basis of the observed variation in SPC and its relationships with agronomic traits could facilitate the designing of efficient breeding strategies for improving SPC while maintaining other desirable agronomic attributes such SY and SW in pigeonpea. Because variations and relationships among traits are dependent upon the set of materials evaluated and the environment in which they are tested (Hamdi et al., 1991; Wray and Visscher, 2008), re-evaluating the 23 and other potentially useful genotypes for SPC and agronomic traits in multiple sets of environments may be necessary.

### Acknowledgements

We thank the United States Agency for International Development (USAID) and ICRISAT for funding various projects related to pigeonpea genomics at ICRISAT. This work has been undertaken as part of the CGIAR Research Program on Grain Legumes. ICRISAT is a member of CGIAR Consortium.

### References

- Afzal MA, Bashir MM, Luna NK, Bakr MA, Haque MM (2003) Relationship between seed size, protein content and cooking time of mungbean [*Vigna radiata* (L.) Wilczek] seeds. *Asian J Plant Sci.* 2:1008-1009.
- Akibode CS, Maredia M (2011) Global and regional trends in production, trade and consumption of food legume crops. Report submitted to the Standing Panel on Impact Assessment (SPIA) of the CGIAR Science Council, FAO, Rome.
- Asante IK, Adu-Dapaah H, Addison P (2004) Seed weight and protein and tannin contents of 32 cowpea accessions in Ghana. *Trop Sci.* 44:77-19.
- Baudoin JP, Maquet A (1999) Improvement of protein and amino acid content in seeds of food legumes - a case study in *Phaseolus*. *Biotechnol Agron Soc Environ.* 3:220-224.
- Brim CA, Burton JW (1978) Recurrent selection in soybeans. II. Selection for increased percent protein in seeds. *Crop Sci.* 19:494-498.

- Burstin J, Marget P, Huart M, Moessner A, Mangin B, Duchene C, Desprez B, Munier-Jolain N, Duc G (2007) Developmental genes have pleiotropic effects on plant morphology and source capacity, eventually impacting on seed protein content and productivity in pea. *Plant Physiol.* 144:768-781.
- Burton GW (1952) Quantitative inheritance in grasses. *Proceeding of the 6th International Grassland Congress* 1:277-283.
- Craufurd PQ, Soko HS, Jones JK, Summerfield RJ (2001) Inheritance of duration from sowing to first flowering in pigeonpea. *Euphytica* 119:323-333.
- Falconer DS, Mackey FC (1996) *Introduction to quantitative genetics*, 4th edn. Longman, London.
- FAOSTAT (2017) Food and Agricultural Organization of the United Nations Statistics Division. Available at: <http://www.fao.org/faostat/en/#data/QC>. Accessed 20 January 2018.
- Filho MM, Destro D, Miranda LA, Spinosa WA, Carrão-Panizzi MC, Montalván R (2001) Relationships among oil content, protein content and seed size in soybeans. *Braz Arch Biol Technol.* 44:23-32.
- Foley JA, Ramankutty N, Brauman KA, Cassidy ES, Gerber JS, Johnston M, Mueller ND, O'Connell C, Ray DK, West PC, Balzer C, Bennett EM, Carpenter SR, Hill J, Monfreda C, Polasky S, Rockström J, Sheehan J, Siebert S, Tilman D, Zaks DPM (2011) Solutions for a cultivated planet. *Nature* 478:337-342.
- Hamdi A, Erskine W, Gates P (1991) Relationships among economic characters in lentil. *Euphytica* 57:109-116.
- Iqbal A, Khalil IA, Ateeq N, Khan MS (2006) Nutritional quality of important food legumes. *Food Chem.* 97:331-335.
- Johnson HW, Robinson HE, Comstock RE (1955) Estimates of genetic and environmental variability in soybean. *Agron J.* 47:314-318.
- Kulwal PL, Mhase LB (2017) Protein content exhibits significant positive correlation with seed weight in chickpea germplasm collection. *Plant Genet Resour C.* 15:283-285.
- Lawn RJ, Rebetzke GJ (2006) Variation among Australian accessions of the wild mungbean (*Vigna radiata* ssp. *sublobata*) for traits of agronomic, adaptive, or taxonomic interest. *Aust J Agric Res.* 57:119-132.
- Leleji OI, Dickson MH, Crowder LV, Bourke JB (1972) Inheritance of crude protein percentage and its correlation with seed yield in beans, *Phaseolus vulgaris* L. *Crop Sci.* 12:168-171.
- Mligo JK, Craufurd PQ (2005) Adaptation and yield of pigeonpea in different environments in Tanzania. *Field Crops Res.* 94:43-53.
- Mula MG, Saxena KB (2010) Lifting the level of awareness on pigeonpea - a global perspective. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Odeny DA (2007) The potential of pigeonpea (*Cajanus cajan* (L.) Millsp.) in Africa. *Nat Resour Forum* 31:297-305.
- Rao PP, BIRTHAL PS, Bhagavatula S, Bantilan MCS (2010) Chickpea and pigeonpea economies in Asia: facts, trends and outlook. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- Rekha R, Prasanthi L, Sekhar MR, Priya MS (2013) Studies on selection indices in pigeonpea [*Cajanus cajan* (L.) Millsp.]. *Int J Appl Biol Pharm.* 4:291-294.
- Remanandan P, Sastry DVSSR, Mengesa MH (1988) ICRISAT pigeonpea germplasm catalog: evaluation and analysis. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India.
- SAS Institute (2015) *User's Guide: SAS/STAT® 14.1*. SAS Institute, Cary, NC.
- Sawargaonkar SL (2010) Study of heterosis, combining ability, stability and quality parameters in CGMS-based pigeonpea [*Cajanus cajan* (L.) Millsp.] hybrids. PhD Thesis. Marathwada Agricultural University.
- Saxena KB, Faris DG, Singh U, Kumar RV (1987) Relationship between seed size and protein content in newly developed high protein lines of pigeonpea. *Plant Foods Human Nutr.* 36:335-340.
- Saxena KB, Kumar RV, Rao PV (2002) Pigeonpea nutrition and its improvement. *J Crop Improv.* 5:227-260.
- Saxena, RK, von Wettberg E, Upadhyaya HD, Sanchez V, Songok S, Saxena KB, Kimurto P, Varshney RK (2014). Genetic diversity and demographic history of *Cajanus* spp. illustrated from genome-wide SNPs. *PLoS ONE* 9:e88568.
- Singh RK, Chaudhary BD (1979) *Biometrical methods in quantitative genetic analysis*. Kalyani Publishers, New Delhi.
- Subramanian S, Menon M (1973) Heterosis and inbreeding depression in rice. *Madras Agric J.* 60:1139-1144.
- Upadhyaya HD, Bajaj D, Narnoliya L, Das S, Kumar V, Gowda CLL, Sharma S, Tyagi AK, Parida SK (2016) Genome-wide scans for delineation of candidate genes regulating seed protein content in chickpea. *Front Plant Sci.* 7:302.
- Upadhyaya HD, Reddy KN, Gowda CLL, Silim SN (2007) Patterns of diversity in pigeonpea (*Cajanus cajan* (L.) Millsp.) germplasm collected from different elevations in Kenya. *Genet Resour Crop Evol.* 54:1787-1795.
- Wilcox JR, Cavins JF (1995) Backcrossing high seed protein to a soybean cultivar. *Crop Sci.* 35: 1036-1041.
- Wray NR, Visscher PM (2008) Estimating trait heritability. *Nature Education* 1:29.
- Wright S (1921) Correlation and causation. *J Agric Res.* 20:557-585.