Physiol Mol Biol Plants (November–December 2018) 24(6):1245–1259
https://doi.org/10.1007/s12298-018-0584-6

RESEARCH ARTICLE

Validation of QTLs for plant ideotype, earliness and growth habit traits in pigeonpea (*Cajanus cajan* Millsp.)

Prakash G. Patil[1,2] · Abhishek Bohra[1] · Naik S. J. Satheesh[1] · Jyotirmay Dubey[1] · Praveen Pandey[1] · Dibendu Dutta[1] · Farindra Singh[1] · I. P. Singh[1] · N. P. Singh[1]

Received: 7 November 2017 / Revised: 7 July 2018 / Accepted: 17 July 2018 / Published online: 4 September 2018

Abstract Pigeonpea productivity is greatly constrained by poor plant ideotype of existing Indian cultivars. Enhancing pigeonpea yield demands a renewed focus on restructuring the ideal plant type by using more efficient approaches like genomic tools. Therefore, the present study aims to identify and validate a set of QTLs/gene(s) presumably associated with various plant ideotype traits in pigeonpea. A total of 133 pigeonpea germplasms were evaluated along with four checks in the augmented design for various ideotype traits i.e. initiation of flowering (IF), days to 50% flowering (DFF), days to maturity (DM), plant height (PH), primary branches (PB), seeds per pod (SP) and pod length (PL). We observed significant genetic diversity in the germplasm lines for these traits. The genetic control of IF, DFF, DM and PH renders these traits suitable for detection of marker trait associations. By using residual maximum likelihood algorithm, we obtained appropriate variance–covariance structures for modeling heterogeneity, correlation of genetic effects and non-genetic residual effects. The estimates of genetic correlations indicated a strong association among earliness traits. The best linear unbiased prediction values were calculated for individual traits, and association analysis was performed in a panel of 95 diverse genotypes with 19 genic SSRs. Out of five QTL-flanking SSRs used here for validation, only ASSR295 could show significant association with FDR and Bonferroni corrections, and accounted for 15.4% IF, 14.2% DFF and 16.2% DM of phenotypic variance (PV). Remaining SSR markers (ASSR1486, ASSR206 and ASSR408) could not qualify false discovery rate (FDR) and Bonferroni criteria, hence declared as false positives. Additionally, we identified two highly significant SSR markers, ASSR8 and ASSR390 on LG 1 and LG 2, respectively. The SSR marker ASSR8 explained up to 22 and 11% PV for earliness traits and PB respectively, whereas ASSR390 controlled up to 17% PV for earliness traits. The validation and identification of new QTLs in pigeonpea across diverse genetic backgrounds brightens the prospects for marker-assisted selection to improve yield gains in pigeonpea.

Keywords Ideotype · Pigeonpea · QTL · Trait mapping · Germplasm · SSR

Introduction

Pigeonpea [*Cajanus cajan* (L.) Millsp.] belongs to the tribe Phaseoleae in the family Leguminosae (Lackey 1977; van der Maesen 2003). *C. cajan* is the only domesticated species subtribe Cajaninae and has a genome size of 833.07 Mbp (Varshney et al. 2012). It is a hardy, widely adapted, drought tolerant pulse crop that is cultivated globally on 5.40 million hectares with an average annual production of 4.48 million tonnes (FAOSTAT 2016). Due to the vast natural genetic variability in local germplasm and the presence of numerous wild relatives, India is considered as the primary center of origin (Van der Maesen 1980) and remains one of the largest pigeonpea producers accounting for 71% (3.88 mha) and 63.39% (2.84 mt) of

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12298-018-0584-6) contains supplementary material, which is available to authorized users.

✉ Abhishek Bohra
abhi.omics@gmail.com

[1] Crop Improvement Division, ICAR-Indian Institute of Pulses Research (IIPR), Kanpur 208024, India

[2] Present Address: ICAR-National Research Centre on Pomegranate, Solapur 413 255, India

the global area and production, respectively (FAOSTAT 2016). Pigeonpea cultivars have a narrow genetic base due to limited utilization of wild pigeonpea species in breeding programs. The breeding efforts aimed at improving pigeonpea led to the development and release of more than 100 improved varieties over the last 50 years in India (Singh et al. 2016). However, the genetic gain from conventional breeding has remained limited over same period of time (Varshney et al. 2013). This implies toward an urgent need to strengthen pigeonpea breeding programs with the modern genomic tools to improve their efficacy (Bohra et al. 2017).

Notwithstanding the substantial efforts directed towards pigeonpea genetic improvement, the crop productivity has remained near-stagnant over the last several decades. Poor productivity of pigeonpea is due to a range of constraints, of which poor plant type and longer crop duration of existing cultivars are of paramount significance (Odeny et al. 2007). Improving yield gains in pigeonpea calls for a renewed focus on restructuring the ideal plant type. Also, the development of short duration pigeonpea cultivars could bring significant increment in pigeonpea productivity (Saxena and Sharma 1990; Saxena 2008). In pigeonpea, days to flowering and days to maturity determine earliness and the two traits are important with respect to increasing cropping intensity in rotation systems such as pigeonpea—wheat system.

Plant ideotype breeding intends to deliver crop genotypes that are suitable for modern farming practices and it involves improvement in key traits such as harvest index and mechanical harvesting. In other words, ideotype breeding seeks accumulating favorable QTLs for various component traits in a given genotype (Wu 1998). Several plant attributes including plant height, number of branches, pods per plant and synchronous maturity collectively contribute to improved plant architecture. Besides, short-duration cultivars are important in light of the need for increasing cropping efficiency of farming system. Although conventional breeding in pigeonpea has delivered a range of cultivars catering to the demand of wider farming community, remodeling of plant type through deploying modern genomic tools has not yet been started (Kumawat et al. 2012).

To understand the genetic architecture of a particular trait, discovery of genomic regions/QTLs tightly associated with the trait remains crucial. To this end, Kumawat et al. (2012) identified a set of QTLs in pigeonpea for traits pertaining to plant type and earliness viz. plant height (qPH5.1), number of secondary branches/plant (qSB5.1), number of pods/plant (qPD5.1), days to flowering (qFL5.1) and days to maturity (qMT5.1). Interestingly, all these QTLs were mapped on to the linkage group 5 (LG 5) within the marker interval 'ASSR100-ASSR 206'. Two QTLs for earliness qFL5.1 and qMT5.1 had PV up to 25.9%, whereas QTLs for plant type traits (qPH5.1, qSB5.1 and qPD5.1) explained up to 27.5% PV. Later, Geddam and colleagues (2014) successfully validated QTLs for plant type and earliness traits in recombinant inbred (RIL) population in pigeonpea. Determinacy is another important trait in pigeonpea for which *CcTFL1* gene has been identified as a likely candidate gene (Mir et al. 2014).

Apart from bi-parental population, QTL-validation in diverse germplasm lines confirms the utility of these tools for accelerated and targeted improvement of the concerned traits. Given the fact that ideotype breeding involves reconstructing plant type via capturing favorable gene(s)/allele(s) (Dhanasekar et al. 2010), the present study reports identification of new and validation of previously reported QTLs for plant ideotype traits in pigeonpea. The information generated in this study could pave the way for an efficient ideotype breeding to deliver superior pigeonpea varieties.

Materials and methods

Plant material and trait phenotyping

Plant material comprised 133 diverse pigeonpea genotypes including four controls viz. UPAS 120, ICP 8863, IPA 203 and Dholi Dwarf (Suppl. Table 1). Experiment was conducted in augmented design (Federer 1956) with seven blocks at Indian Institute of Pulses Research (IIPR), Kanpur during 2016–2017. Controls were repeated in each block. Each block contained 19 test entries and four controls with a spacing of 60 × 30 cm and a row length of 5 m each. Recommended agronomic practices were followed to raise healthy crop. Phenotypic data on plant ideotype traits viz. plant height (PH), primary branches (PB), seeds per pod (SP), pod length (PL), and earliness traits such as initiation of flowering (IF), days to fifty percent flowering (DFF), days to maturity (DM) and determinacy [determinate (DT)/indeterminate (IDT)] were recorded on all entries. Five plants from the middle of each row were used for trait scoring. The earliness traits i.e. days to initiation of flowering and fifty percent flowering were scored as number of days counted from the date of sowing to the opening of the first flower and fifty percent flowering, whereas days to maturity were calculated from the date of sowing to the date when nearly 80% of the pods turned yellow. The remaining plant type traits viz. PH, PB DT, IDT, SP and PL were recorded at maturity and after harvesting. Finally, a sub set of 95 genotypes was chosen for validation of QTLs/genes using an association mapping approach (Table 1). On the other hand, 120 genotypes were

Table 1 Panel of 95 pigeonpea genotypes used for association analysis

Sl. no.	Genotype	Type	Sl. no.	Genotype	Type
1	D 20	Breeding line	49	IC 368982	Germplasm line
2	DSLR 129	Breeding line	50	IC 368995	Germplasm line
3	ICPL 87	Variety	51	IC 368996	Germplasm line
4	ICPL 51	Germplasm line	52	IC 368999	Germplasm line
5	ICPL 7124	Germplasm line	53	IC 8345	Germplasm line
6	ICPL 84023	Germplasm line	54	IC 16202	Germplasm line
7	ICPL 87154	Germplasm line	55	RVK 284	Breeding line
8	ICPL 91045	Germplasm line	56	Bennur Local	Landrace
9	ICPL 67B	Breeding line	57	VKS 11/24-2	Germplasm line
10	UPAS 120	Variety	58	TJT 501	Variety
11	PUSA 992	Variety	59	JKM 189	Variety
12	ICPL 11255	Germplasm line	60	ICP 3451	Germplasm line
13	ICPL 20338	Germplasm line	61	IPAC 79	Germplasm line
14	ICPL 20340	Germplasm line	62	PKV Tara	Variety
15	ICPL 88034	Germplasm line	63	IPA 2010-30-5	Germplasm line
16	IC 15707	Germplasm line	64	JSA 59	Breeding line
17	IC 16191-1	Germplasm line	65	RVK 275	Breeding line
18	IC 16192-1	Germplasm line	66	BSMR 853	Variety
19	IC 16198-1	Germplasm line	67	ICP 348	Germplasm line
20	IC 16206	Germplasm line	68	WRP 1	Variety
21	IC 22520	Germplasm line	69	ICPL 87119	Variety
22	IC 22540	Germplasm line	70	AKT 9913	Breeding line
23	IC 23686	Germplasm line	71	JBP 13A	Breeding line
24	IC 25049	Germplasm line	72	JBP 13B	Breeding line
25	IC 25053	Germplasm line	73	PDA 10A	Breeding line
26	IC 28186	Germplasm line	74	PDA 10B	Breeding line
27	IC 28202	Germplasm line	75	NDA 2	Variety
28	IC 33725	Germplasm line	76	MAL 13	Variety
29	IC 44865	Germplasm line	77	Banda Palera	Landrace
30	IC 52944	Germplasm line	78	P 3497	Germplasm line
31	IC 56060	Germplasm line	79	JAM 9-9	Breeding line
32	IC 56066	Germplasm line	80	ICP 6951	Germplasm line
33	IC 78357	Germplasm line	81	ICP 109893	Germplasm line
34	IC 94674	Germplasm line	82	ICP 109888	Germplasm line
35	IC 94677	Germplasm line	83	IPA 7-2	Breeding line
36	IC 94486	Germplasm line	84	DSLR 124	Germplasm line
37	IC 299033	Germplasm line	85	PANT A3	Variety
38	IC 299035	Germplasm line	86	JAM 9-19	Breeding line
39	IC 299052	Germplasm line	87	R 7-2	Germplasm line
40	IC 347147	Germplasm line	88	P 593	Germplasm line
41	IC 369012	Germplasm line	89	C 2291	Germplasm line
42	IC 369013	Germplasm line	90	DPP 3-42	Germplasm line
43	IC 369014	Germplasm line	91	ICP 8585	Germplasm line
44	IC 369015	Germplasm line	92	MSQ 14A	Germplasm line
45	IC 368952	Germplasm line	93	ICP 8863	Variety
46	IC 368963	Germplasm line	94	IPA 203	Variety
47	IC 368966	Germplasm line	95	Dholi Dwarf	Landrace
48	IC 368971	Germplasm line			

used to determine the discrimination efficiency of *CcTFL*1 marker for determinacy trait.

Marker genotyping

Genomic DNA was isolated from young leaves of pigeonpea genotypes by following modified cetyltrimethylammonium bromide (CTAB) method (Agbagwa et al. 2012). Polymerase chain reaction (PCR) templates were prepared by diluting stock DNA samples to 10 ng/µl. Association analysis was conducted using 19 genic-SSRs including five QTL flanking markers with known map positions (reported by Kumawat et al. 2012) and *CcTFL1* specific SNP markers for determinacy trait (reported by Mir et al. 2014) (Suppl. Table 2). The PCR for genic SSRs was performed in 10 µl reaction mixture that contained 10 ng genomic DNA, 1 µl 10X PCR buffer (15 mM of MgCl2), 1 mM dNTP mix, 10 µM of forward and reverse primer and 0.2 µl *Taq* DNA polymerase (3 U/µl). The PCR program was as follows: initial denaturation at 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min and final extension of 72 °C for 7 min. For allele specific *CcTFL1* marker assay, two external common primers (external common forward primer-TFL1_PCR_CF and external common reverse primer-TFL1_PCR_CR) flanking the SNP and one internal primer targeting one SNP allele "A-allele" (TFL1_PCR_A) and the other internal primer targeting the other SNP allele "T-allele" (TFL1_PCR_T were used. These four primers were multiplexed into a single PCR reaction to obtain co-dominant pattern. The 20 µl reaction mixture containing 20 ng genomic DNA, 2 µl 10X PCR buffer (15 mM of MgCl2), 2 mM dNTP mix, 5 µM of each two common external and two SNP specific internal primers and 0.3 µl *Taq* DNA polymerase (3 U/µl) was used. The touchdown PCR profile was followed with the following conditions: Initial denaturation at 94 °C for 5 min followed by 10 cycles of touchdown 60 °C, 30 s at 94 °C, annealing for 30 s at 55 °C (the annealing temperature for each cycle being reduced by 0.5 °C per cycle) and extension for 45 s at 72 °C. This was accompanied by 30 cycle of denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s, elongation at 72 °C for 45 s, and 10 min of final extension at 72 °C. All the PCR reactions were conducted in G-40402 thermo cycler (G-STORM, Somerset, UK). The amplification products of genic SSRs, *CcTFL1* assay were resolved on 3 and 2% agarose gels respectively using 1XTBE running buffer. Finally, all the images were analyzed in Quantity one software (Bio-Rad, CA 94547, USA) (Fig. S1).

Statistical analyses

Phenotypic data on seven plant ideotype traits were recorded on 133 test genotypes to obtain adjusted trait values and subjected to analysis of variance (ANOVA) as per augmented design (Federer 1956) using Windostat ver. 8.5. The descriptive statistics, frequency distribution graphs and box plot analysis were performed using GenStat ver. 18. The residual maximum likelihood (REML; Patterson and Thompson 1971) algorithm of GenStat ver. 18 was used to analyze all the traits measured in experiments, considering genotypes as random and block effect as fixed for calculating the best linear predictions (BLUPs). The BLUP values were used in association mapping since these will minimize the effects of environmental variations (Kump et al. 2011). Let y be a trait observations, then the linear mixed model has the form:

$$y_{ib} = \mu + g_i + k_b + e_{ib}$$

where y_{ib} is the phenotypic value of ith genotype and bth block; μ is the overall mean of the genotypes; g_i is the random effect of ith genotype; k_b is the fixed effect of bth block; and e_{ib} is the random residual error due to genotype and block. The block terms were treated as fixed effects, and their significance was assessed by Wald test statistics ($p < 0.05$), while genotype was considered as random effect and their significance was tested by likelihood ratio test. Variance components due to genotype (σg^2) and corresponding standard errors (SE) were estimated and then used for calculating heritability. To compare residual genetic variance–covariance matrix models for each trait and to identify best model, both Akaike Information Criteria (AIC, Akaike 1974) and Bayesian Information Criteria (BIC, Raftery 1986) were calculated. The minimization of AIC and BIC allowed us to select the significant factors, variance function and covariance structure to be considered in the model. Total five models were compared for variance–covariance (VCOV) structure: model 1-Identity (ID) considers homogeneous residual genetic variances, model 2-Diagonal (DIAG) considers heterogeneous residual genetic variance, models 3-AR1 and 4-AR2 are based on first and second order autoregressive for heterogeneous residual variances and model 5-Uniform (UNI) is based on residual variance. The estimated BLUPs were used for association analysis and marker validation. Further, pairwise genetic and phenotypic correlations were calculated using Pearson's correlation coefficient and correlation heat map was drawn using Heatmapper (Babicki et al. 2016, http://www.heatmapper.ca). For calculating pairwise genetic correlations for the plant ideotype traits, model-based predictions for each trait from corresponding linear mixed model were used.

Population structure analysis

The data of 19 SSRs and one SNP marker scored on 95 genotypes were used to infer genetic structure using model-based approach with STRUCTURE ver. 2.3.3 (Pritchard et al. 2000, http://pritch.bsd.uchicago.Edu/structure.html). The project was run with the admixture model and correlated allele frequency using burn in period of 20,000 and 200,000 Markov Chain Monte Carlo (MCMC) replications. Five independent runs were performed with each K value ranging from 1 to 10. Evanno's delta K value was calculated by using STRUCTURE HARVESTER program by processing the STRUCTURE results (Evanno et al. 2005). DARwin ver. 6.0.13 (Perrier and Jacquemoud-Collet 2006) was employed to generate genetic distance (GD) matrix, which was then used to creating dendrogram and factorial analysis.

Association analysis

Association analysis was performed with TASSEL ver. 2.1 (Bradbury et al. 2007). Marker significance was tested for genotypic and phenotypic data by following three models (GLM Q, K and Q + K). Markers with minor allele frequency (MAF) of less than 5% were excluded, and the remaining markers were then used for association analysis. For the mixed linear model (MLM), both K and Q matrices were incorporated (Lu et al. 2015), whereas information on only population structure (Q-matrix) was used as a covariate in general linear model (GLM). SPAGeDi (Hardy and Vekemans 2002) was used to calculate kinship (K) coefficients. A kinship coefficient computed as a correlation coefficient between allelic states (Loiselle et al. 1995). All negative kinship values were set to zero (Yu et al. 2006). Significance of Marker-Trait Association (MTA) was tested by two thresholds viz. Bonferroni correction and false discovery rate (FDR). Bonferroni-corrected threshold probability based on individual tests is calculated to correct for multiple comparisons, using 1/N ($\alpha = 0.05$), where N is the number markers tested. The Bonferroni threshold was 1/16 *0.05 = 0.003125, where 16 is the number of association tests for each trait in this study. The FDR was performed (FDR, $\alpha_c = 0.05$) according to Benjamini and Hochberg (1995). In addition to association analysis, single marker analysis (SMA) was performed for *CcTFL1* gene using SPSS ver. 16.

Results

Analysis of variance and estimates of genetic parameters

Highly significant mean sum of squares due to checks and test genotypes were observed for all the plant ideotype traits except PB (Table 2). Mean squares due to 'genotypes versus checks' were significant for all traits except PH and PB. The estimates of mean values with wider range were observed for all the traits across 137 genotypes (Table 3). These results suggested adequate variability among the pigeonpea genotypes examined. Higher values were obtained for both genetic coefficient of variation (GCV) and phenotypic coefficient of variation (PCV) in case of IF, DFF and PH. Lower GCV and higher PCV were observed for DM and PB, while SP and PL had lower GCV and PCV with greater differences between the two. The estimate of h^2 ranged from 0.13 (PB) to 0.73 (IF) (Table 3). PB and SP showed lower heritabilities (< 0.30) in contrast to the IF and DFF having higher heritabilities (> 0.60). On the other hand, traits like DM, PH, PL showed moderate heritability estimates (0.30–0.60). The coefficient of variation (CV) ranged from 10.06 (SP) to 27.50 (PB), the higher CV was noticed for PB followed by PH (25.33), IF (24.77), DFF (22.98) and DM (19.65), whereas SP and PL had lower CV (10.06 and 11.40).

Table 2 Analysis of variance for plant ideotype traits in pigeonpea germplasm

Sources of variation	*df*	Mean sum of squares						
		IF	DFF	DM	PH	PB	SP	PL
Blocks	6	33.66	63.19	675.99	608.92	6.65	0.072	0.24
Entries (genotypes + checks)	136	848.79***	977.12***	2006.79***	1124.89***	7.35	0.410**	0.94**
Genotypes	132	783.44***	913.79***	1823.07***	969.16**	7.22	0.40**	0.87*
Checks	3	3874.99***	3857.56***	9739.27***	8351.79***	10.50	0.48*	2.47**
Genotypes versus checks	1	396.63**	696.35**	3060.00*	0.009	14.39	1.39**	4.97**
Error	18	30.57	69.84	455.69	268.53	5.34	0.107	0.34

df degrees of freedom, *IF* days to initiation of flower, *DFF* days to 50% flowering, *DM* days to maturity, *PH* plant height (cm), *PB* number of primary branches, *SP* seeds per pod, *PL* pod length (cm)

*Significant at $p < 0.05$, **$p < 0.01$, ***$p < 0.001$ levels

Table 3 Descriptive statistics for plant ideotype traits in pigeonpea

Sl. no	Trait	Min	Max	Mean	GCV	PCV	h^2	GAM	CV %
1	IF	34	151	103.96 ± 2.04	0.25	0.25	0.73	0.50	24.77
2	DFF	46	216	119.68 ± 2,17	0.23	0.24	0.71	0.45	22.98
3	DM	118	285	199.92 ± 3.11	0.17	0.21	0.41	0.30	19.65
4	PH	37.6	195	115.46 ± 2.32	0.21	0.26	0.59	0.37	25.33
5	PB	3.4	17.4	9.34 ± 0.20	0.14	0.29	0.13	0.14	27.50
6	SP	2.6	4.8	3.58 ± 0.03	0.14	0.17	0.30	0.25	10.06
7	PL	3.4	7.3	5.20 ± 0.05	0.14	0.18	0.33	0.21	11.40

Where GCV and PCV [0.0–0.1: low; 0.1–0.2: moderate; > 0.2: high] h^2 [0.0–0.3: low; 0.3–0.6: moderate; > 0.6: high] GAM [0.0–0.1: low; 0.1–0.2: moderate; ≥ 0.2: high]

**GCV* genotypic coefficient of variance, *PCV* phenotypic coefficient of variance, h^2 broad sense heritability, *GAM* genetic advance as percent of mean, *CV* coefficient of variation, *IF* days to initiation of flower, *DFF* days to 50% flowering, *DM* days to maturity, *PB* number of primary branches, *SP* seeds per pod, *PH* plant height (cm), *PL* pod length (cm)

Variance components

Components of variance due to genotypes and residual (error) and their respective standard errors were estimated (Table 4). The variance components due to random factors genotype and their residual (error) were significant ($p < 0.05$) in comparison to fixed factor for all the traits. Variance due to genotype (σ_g^2) was higher than error variance (σ_e^2) for IF, DFF and PH traits. By contrast, estimates of σ_g^2 were smaller than those of σ_e^2 for DM, PB, SP and PL. Further, to study the variability for each trait based on residual VCOV matrix (G) and to identify best model, the structures of VCOV matrix were compared. The differences for AIC and BIC were used to select the best VCOV models. Out of five models tested (ID, DIAG, AR1, AR2 and UNI), the AIC and BIC values were lower for all the six traits in AR 1 model. On the other hand, the ID model suggested lower AIC and BIC values for SP, indicating the suitability of this model for this particular trait. Finally, appropriate models were used to obtain predicated values. These values were further used to obtain genetic correlations that were compared with phenotypic correlations (Fig. 1). Three traits IF, DFF and DM showed significant positive correlations for both phenotypic and genetic correlations (Suppl. Table 3). However, PB versus PH and SP versus PB showed significant positive correlations only for phenotypic correlations. The PL showed least correlations with all remaining traits for both phenotypic and genetic correlations. It was also evident from the heat map that most of traits with lower correlations were observed for genetic correlations compared to phenotypic correlations.

Population structure and association panel

A subset of 95 pigeonpea genotypes was selected from the collection of 137 genotypes. Suitability of this panel for association analysis was supported by frequency distribution graphs and clustering pattern inferred from both model- and distance-based approaches. The frequency distribution graph showed normal distribution for plant ideotype traits (Fig. 2). Genetic structure of the diversity panel was deciphered using genotypic data of 19 ASSRs and one SNP marker that span seven LGs. Maximum delta K (ad hoc quantity) was noted at K = 2, when analyzed for presumed K of 1–10. In other words, the analysis uncovered two sub-populations in the entire collection. The clustering patterns emanating from STRUCTURE analysis gathered support from other distance-based approaches including neighbor-joining (NJ) tree and factorial analysis (Fig. 3a–c).

Association analysis and validation of QTL linked markers

Association mapping technique was implemented to identify and validate QTLs controlling traits related to plant type, earliness and growth habit in pigeonpea. Nineteen genic SSRs and one *CcTFL*1-based SNP marker were used to elucidate the genetic structure of 95 pigeonpea genotypes. Association analysis was performed using the BLUP values, and three different models were used GLM Q, K and MLM QK. Markers with MAF of less than 5% were removed, and consequently, 16 of the total 20 markers were found to be appropriate for association analysis. By using the criterion $p < 0.05$, the association analysis could detect 10 SSRs tightly associated with plant ideotype traits (Table 5). Of these, four markers ASSR8, ASSR390, ASSR295 and ASSR610 remained significant across all three approaches. However, most of these MTAs did not show significance when subjected to the FDR and Bonferroni corrections. Two markers (ASSR 8 and ASSR390) for GML Q, three markers (ASSRs 8, 295, 390) and four markers (ASSRs 8, 295, 390, 610) for MLM K did qualify the FDR and Bonferroni correction. Out of five previously

Table 4 Fixed and random effects of linear mixed model, and comparison of structures of variance–covariance matrix (VCOV) for selected models considering each trait separately

Traits	Fixed effect	Random effect		Comparison of VCOV structure							
				ID		DIAG		AR1		AR2		UNI	
	k_b	σ_g^2	σ_e^2	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
IF	***	404.6 ± 92.9	136 ± 52.6	1117.7	1123.8	1292.9	1711.1	1103.9	1113.0	1104.3	1116.4	1151.6	1160.7
DFF	**	476.3 ± 109.9	177.4 ± 64.5	1148.3	1154.4	1300.9	1719.1	1130.5	1139.6	1130.6	1142.7	1183.8	1192.9
DM	***	411.5 ± 167.8	587.2 ± 146.9	1228.4	1234.5	1448.4	1866.6	1200.9	1209.9	1202.4	1214.5	1230.4	1239.5
PH	***	337.5 ± 84.5	214.9 ± 59.5	1122.5	1128.5	1290.1	1707.4	1118.1	1127.3	1119.5	1131.6	1162.9	1172.0
PB	***	0.667 ± 0.861	4.782 ± 0.96	434.7	440.7	629.9	1047.2	430.3	439.4	432.2	444.3	436.7	445.8
SP	**	0.035 ± 0.020	0.081 ± 0.019	− 152.9	− 146.9	30.60	446.99	− 152.0	− 143.0	− 150.2	− 138.1	− 150.9	− 141.9
PL	***	0.100 ± 0.044	0.197 ± 0.042	− 11.29	− 5.25	171.73	588.12	− 29.34	− 20.29	− 28.11	− 16.04	− 9.29	− 0.24

*k_b is the fixed effect of blocks, σ_g^2 is the genetic variance, σ_e^2 is the non-genetic error variance, *VCOV* variance and covariance matrix, *ID* identity model, *DIAG* diagonal variance model, *AR1 and AR2* first and second order autoregressive variance models, *UNI* uniform variance model, *IF* days to initiation of flower, *DFF* days to 50% flowering, *DM* days to maturity, *PH* plant height (cm), *PL* pod length (cm), *PB* number of primary branches, *SP* seeds per pod, *AIC* Akaike information criteria, *BIC* Bayesian information criteria

***Significant at ($p < 0.001$), **($p < 0.01$)

reported QTL/gene-linked markers tested here (ASSRs 295, 206, 408, 1486 and *CcTFL*1), only one marker ASSR295 remained significant with FDR and Bonferroni corrections and showed association with DM, IF and DFF with PVs being 16, 15, 14%, respectively. Interestingly, we identified two new markers ASSR8 and ASSR390 that qualified for FDR and Bonferroni correction in three approaches and showed associations with IF, DFF, DM and PB. However, ASSR610 qualified for FDR and Bonferroni only in MLM QK and explained 9% PV for DM. Since *CcTFL1*-based SNP marker could not be included in association analysis due to MAF < 5%, SMA was performed for validating this marker and the analysis revealed significant MTAs for IF, DFF, DM and PL with up to 6% PV.

Determination of favorable marker alleles

Box plot analysis was performed to figure out the allele of associated markers (ASSR8, ASSR390, ASSR295 and *CcTFL* 1) that contributed favorably to desired plant ideotype (Fig. 4). The box plot indicated $ASSR8_{150}$, $ASSR390_{170}$, $ASSR295_{140}$ and $CcTFL1_{734}$ as the contributing alleles. Among these, the allele $ASSR8_{150}$ caused reduction in PB, IF, DFF, DM. Similarly, $ASSR390_{170}$ allele contributed towards reduction in IF, DFF and DM. In case of $ASSR295_{140}$, the allele A3-140 (of total five alleles A1-160, A2-150, A3-140, A4-160/150 and A5-160/140) contributed to reduce IF, DFF and DM. The $CcTFL\ 1_{734}$ allele contributed to reduce IF, DFF, DM and PL.

We also examined efficiency of an SNP marker to discriminate DT and IDT types among 120 genotypes (Fig. 5). The SNP alleles were scored according to Mir and colleagues (2014). Out of 104 IDT types, 81 genotypes had specific allele (734 bp) thus reflecting an efficiency of 77.9%. On the other hand, 56.3% efficiency was observed for DT-specific allele where nine of the total 16 DT genotypes could show 167 bp specific allele.

Discussion

Considerable genetic variability was obtained for all plant ideotype traits analyzed in the current study. Significant differences with wider mean range for yield attributing traits remained in agreement with an earlier study involving advanced breeding lines of pigeonpea (Meena et al. 2017). Similar results were also reported in pigeonpea for plant height by other researchers (Bhadru 2010; Satheesh et al. 2013). Insights on relative contribution of genetic (GCV) and non-genetic (PCV) sources for the trait expression are of great importance (Patil et al. 2015). In this study, we observed higher GCV and PCV with least

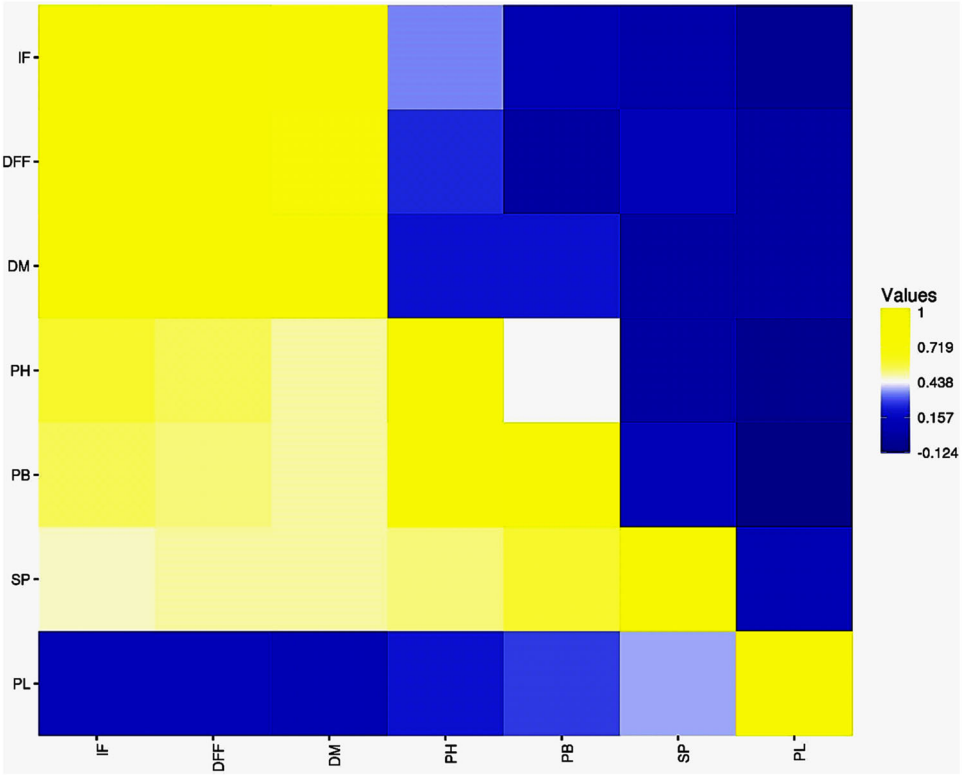

Fig. 1 Heat map of genetic (top) and phenotypic (bottom) correlations between each of the measured plant ideotype traits. The color assigned to a point in the heat map grid indicates the strength of a particular correlation between two traits. The level of correlation is indicated by yellow for positive correlations and blue for negative correlations, as depicted in the color key (color figure online)

differences for IF, DFF and PH indicating these characters to be predominantly under genetic control and consequent responsiveness of these traits for direct selection during trait improvement. Researchers have previously found similar patterns for DFF (Lakhote et al. 2015), and DFF and PH (Meena et al. 2017) in pigeonpea. Higher differences between GCV and PCV for DM, PB, SP and PL suggested profound impact of environment on these traits, which renders response of these traits to direct selection a bit unpredictable. Similarly, Meena et al. (2017) also concluded greater influence of environment on traits like number of seeds/pod based on higher difference between GCV and PCV in pigeonpea. Stronger influence of experimental circumstances on the expression of DM, PB, SP and PL was supported by REML analysis, which revealed lesser values of σ_g^2 than σ_e^2 for these traits. By contrast, higher σ_g^2 estimates for IF, DFF and PH indicated towards strong genetic control of these traits. Low heritability estimates for DM, SP, PL and PB also reflected influence of experimental circumstances. Similar results were obtained for such traits in pigeonpea by other researchers (Chetukuri et al. 2013). Higher CV values for PB, PH, IF, DFF and DM indicated these traits accounted for greater variability among the genotypes compared to SP and PL. The CV values for each trait in this study are congruent with the findings of Saroj et al. (2013).

The mixed model approach is suitable for evaluating the heterogeneity of genetic variances (Malosetti et al. 2013). Structures of VCOV matrix were compared for each trait among different models. All traits showed better fit to first order autoregressive heterogeneous variance model AR1, except SP than showed better fit to homogeneous residual genetic variance ID model. These models were used to model the traits for genetic correlations. The estimates of genetic correlations clearly indicated the strong association among plant earliness traits i.e. IF, DFF and DM. Analysis of phenotypic data demonstrates genetic control of these traits such as IF, DFF, DM and PH, and hence are suitable for detection of MTAs.

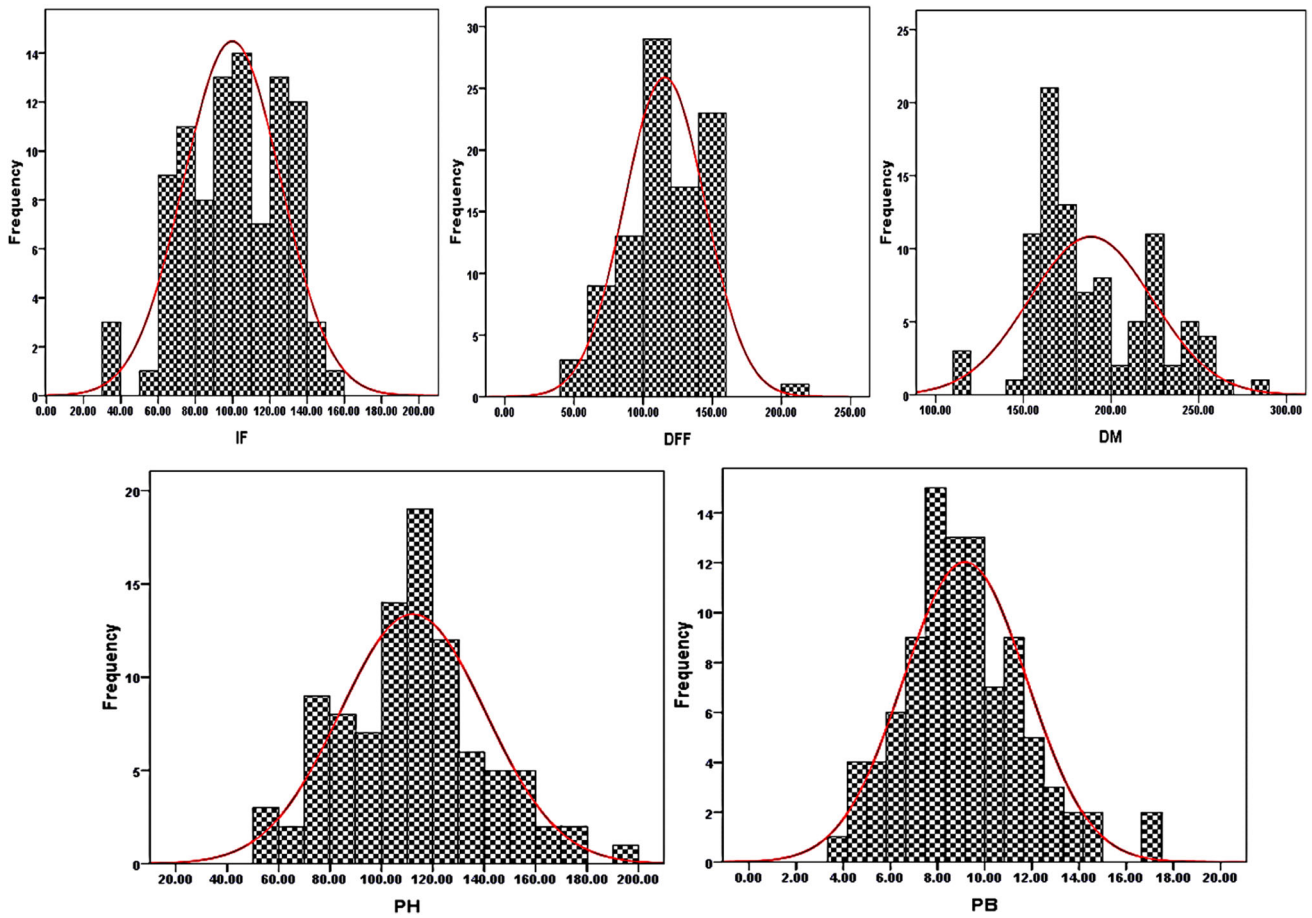

Fig. 2 Frequency distribution graph for important plant ideotype traits in 95 pigeonpea genotypes

Bimodal distribution patterns were obtained for PH and DF suggesting involvement of major genes in comparison to other traits like PB, PD and MT that showed normal distribution patterns with possibility of multiple genes (Kumawat et al. 2012). However, Geddam et al. (2014) reported normal distribution of phenotypic variation for DF, DM, PH, SB and PD, thus suggesting potential involvement of multiple genes in phenotypic manifestation of these attributes. Our results strongly corroborate with the distribution patterns reported earlier in pigeonpea. Analysis of genetic structure of the pigeonpea collection revealed two subpopulations in the collection. Our observation is congruent with a recent study that established two subpopulations based on the analysis of 94 pigeonpea genotypes with 60 SSRs (Bohra et al. 2017). Similarly, Patil et al. (2017) have also found two subpopulations among 89 pigeonpea genotypes using 65 SSRs.

The potential of association analysis in validation of putative QTLs has been shown in various crops including durum wheat (Amallah et al. 2016). In the current study, we conducted association analysis using the BLUP values for each trait. To control false positive associations, three models GLM Q, K and MLM QK were used. Lu et al. (2018) performed GWAS for seed dormancy in rice and the authors found the K and Q + K models having greater control of type I errors. Similar to the present study, the MLM approach has been used for trait mapping in other grain legumes such as common bean (Shi et al. 2011), chickpea (Thudi et al. 2014) and soybean (Hu et al. 2014). Based on these three association analysis approaches, total 10 SSRs were identified using the criteria of $p < 0.05$. However, only four SSRs could reach the level of significance across all three approaches and qualified for FDR and Bonferroni corrections. Of the five SSRs that flanked

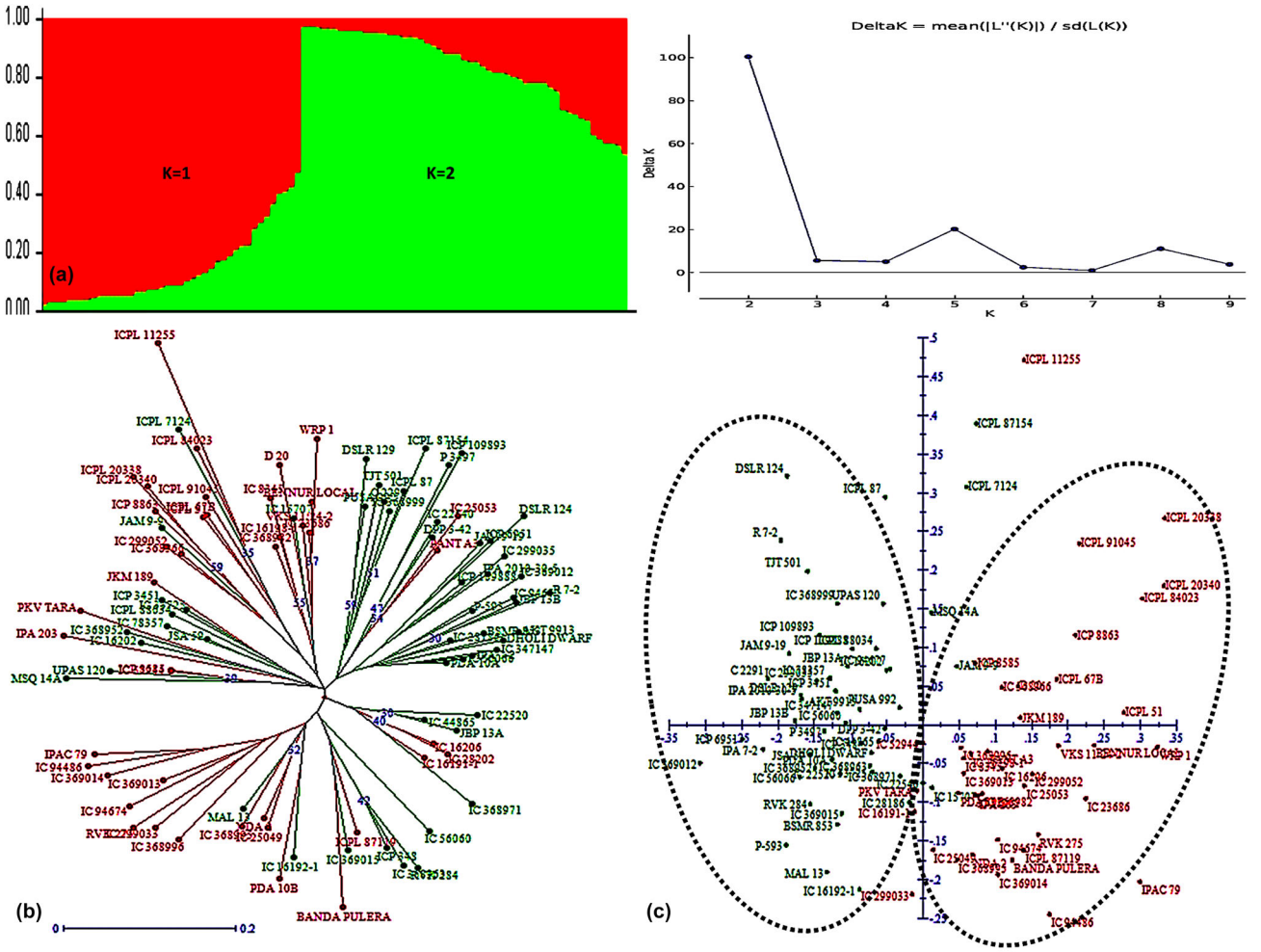

Fig. 3 **a** Population structure of 95 pigeonpea genotypes based on 19 EST–SSR markers (K = 2) and graph of estimated membership fraction for K = 2. The maximum of adhoc measure ΔK determined by structure harvester. **b**, **c** Unrooted neighbour joining tree and factorial analysis showing clustering of 95 pigeonpea genotypes into two subpopulations

QTL region in an earlier study (Kumawat et al. 2012), only ASSR295 was found to be significant with FDR and Bonferroni corrections and showed genetic linkages with DM, IF and DFF. The MTA explained up to 16% PV for these three traits. The remaining markers (ASSR1486, ASSR 206, and ASSR408), however, could only qualify for the criterion of $p < 0.05$, and hence declared as false positives in this study. In addition, here we identified new markers i.e. ASSR8 and ASSR 390 qualifying FDR and Bonferroni corrections in three approaches and showed association with IF, DFF, DM and PB; and ASSR610 could pass the same criteria only in MLM QK approach. ASSR8 explained overall PV of 22% (IF), 17% (DFF), 18% (DM) and 11% (PB). Similarly, ASSR390 explained overall PV of 17% (IF), 15% (DFF) and 13.3% (DM); whereas ASSR610 accounted for 9% of the PV for DM. It is interesting to note that we found an additional set of SSR markers relevant to the component traits, which could not be detected by Kumawat et al. (2012). In addition to trait-linked SSRs, we validated one candidate gene-based SNP

Table 5 Significancant marker trait associations in 95 pigeonpea genotypes

Sl. no	Locus	LG	Trait	Marker trait associations										
				GLM Q			MLM K			MLM QK			SMA	
				F	*p*	R^2	F	*p*	R^2	F	*p*	R^2	F	R^2
1.	ASSR8	1	IF	25.313	$9.99 \times 10^{-4a,b}$	0.2178	26.374	$1.60 \times 10^{-6a,b}$	0.224	25.2761	$2.5 \times 10^{-6a,b}$	0.2168	13.43***	0.209
			DFF	18.4087	$9.99 \times 10^{-4a,b}$	0.1691	19.0555	$3.35 \times 10^{-5a,b}$	0.1726	18.3825	$4.53 \times 10^{-5a,b}$	0.1683	9.88***	0.159
			DM	19.7969	$9.99 \times 10^{-4a,b}$	0.1798	20.2864	$1.97 \times 10^{-5a,b}$	0.1816	19.7347	$2.52 \times 10^{-5a,b}$	0.1786	10.23***	0.164
			PB	11.3835	5.0×10^{-3c}	0.1122	11.5698	$9.98 \times 10^{-4a,b}$	0.1059	11.4848	$1.0 \times 10^{-3a,b}$	0.1053	6.41**	0.103
			PH	–	–		7.1565	8.9×10^{-3c}	0.0728	7.66	6.9×10^{-3c}	0.0781	3.59*	0.052
2.	ASSR390	2	IF	9.1115	$9.99 \times 10^{-4a,b}$	0.1669	9.5699	$1.69 \times 10^{-4a,b}$	0.1721	9.0021	$2.73 \times 10^{-4a,b}$	0.162	6.82***	0.157
			DFF	8.2839	$9.99 \times 10^{-4a,b}$	0.1547	8.4405	$4.34 \times 10^{-4a,b}$	0.1495	8.0835	$5.91 \times 10^{-4a,b}$	0.1446	6.20**	0.142
			DM	6.9327	7.0×10^{-3c}	0.1328	7.2187	$1.2 \times 10^{-3a,b}$	0.1363	6.905	$1.6 \times 10^{-3a,b}$	0.1318	4.97**	0.113
3.	**ASSR295**	**4**	**IF**	**4.2011**	$\mathbf{3.1 \times 10^{-2c}}$	**0.1589**	**3.9965**	$\mathbf{5.0 \times 10^{-3a}}$	**0.147**	**4.2382**	$\mathbf{3.5 \times 10^{-3a}}$	**0.1554**	**3.41****	**0.114**
			DFF	**–**	**–**	**–**	**3.6253**	$\mathbf{8.8 \times 10^{-3a}}$	**0.1396**	**3.7084**	$\mathbf{7.8 \times 10^{-3a}}$	**0.1432**	**3.14***	**0.103**
			DM	**4.3612**	$\mathbf{1.9 \times 10^{-2c}}$	**0.1647**	**4.3556**	$\mathbf{2.9 \times 10^{-3a,b}}$	**0.1606**	**4.3787**	$\mathbf{2.8 \times 10^{-3a,b}}$	**0.1615**	**4.07****	**0.141**
			PH	**–**	**–**	**–**	**3.5118**	$\mathbf{1.04 \times 10^{-2c}}$	**0.1369**	**3.3817**	$\mathbf{1.27 \times 10^{-2c}}$	**0.1331**	**3.21***	**0.105**
4.	ASSR610	2	IF				–	–	–	5.8392	1.77×10^{-2c}	0.0578	4.25*	0.065
			DFF				–	–	–	4.3661	3.95×10^{-2c}	0.0447	3.68*	0.054
			DM	9.5341	1.7×10^{-2c}	0.0942	4.1566	4.43×10^{-2c}	0.0419	9.4665	$2.8 \times 10^{-3a,b}$	0.0919	4.23*	0.064
			PL				4.6721	3.33×10^{-2c}	0.0475	–	–	–	ns	ns
5.	**ASSR408**	**10**	**DFF**	**–**	**–**	**–**	**4.3086**	$\mathbf{4.07 \times 10^{-2c}}$	**0.0443**	**4.2449**	$\mathbf{4.22 \times 10^{-2c}}$	**0.0438**	**5.14****	**0.081**
			DM				**4.9836**	$\mathbf{2.8 \times 10^{-2c}}$	**0.0511**	**4.906**	$\mathbf{2.93 \times 10^{-2c}}$	**0.0505**	**4.65***	**0.072**
			PH				**4.3315**	$\mathbf{4.02 \times 10^{-2c}}$	**0.0415**	**4.3285**	$\mathbf{4.03 \times 10^{-2c}}$	**0.0418**	**3.59***	**0.052**
6.	**ASSR1486**	**4**	**IF**	**–**	**–**	**–**	**3.5094**	$\mathbf{3.41 \times 10^{-2c}}$	**0.05**	**3.4899**	$\mathbf{3.48 \times 10^{-2c}}$	**0.0494**	**2.81***	**0.055**
			DFF				**3.8752**	$\mathbf{2.43 \times 10^{-2c}}$	**0.058**	**3.8053**	$\mathbf{2.6 \times 10^{-2c}}$	**0.0572**	**2.75***	**0.053**
			DM				**3.1648**	$\mathbf{4.7 \times 10^{-2c}}$	**0.0562**	**3.2121**	$\mathbf{4.5 \times 10^{-2c}}$	**0.056**	**ns**	**ns**
7.	ASSR293	5	IF	–	–	–	5.4382	2.19×10^{-2c}	0.0412	5.4573	$2. \times 10^{-2c}$	0.0386	ns	ns
			DFF				5.6606	1.94×10^{-2c}	0.0452	5.537	2.08×10^{-2c}	0.0438	ns	ns
			PL				7.018	9.5×10^{-3c}	0.0714	6.9423	9.9×10^{-3c}	0.0626	9.01***	0.146
8.	**ASSR206**	**5**	**IF**	**–**	**–**	**–**	**4.4075**	$\mathbf{1.49 \times 10^{-2c}}$	**0.0838**	**4.5091**	$\mathbf{1.36 \times 10^{-2c}}$	**0.0849**	**3.15***	**0.064**
9.	ASSR23	6	PH	–	–	–	4.2763	1.68×10^{-2c}	0.0857	4.1124	1.95×10^{-2c}	0.0833	3.75*	0.081
10.	ASSR93	4	PB	–	–	–	2.7277	2.44×10^{-2c}	0.1337	2.7373	2.41×10^{-2c}	0.1352	2.31*	0.078

Table 5 continued

Sl. no	Locus	LG	Trait	Marker trait associations									
				GLM Q			MLM K			MLM QK			SMA	
				F	*p*	R^2	F	*p*	R^2	F	*p*	R^2	F	R^2
11.	***CcTFL* 1**	**9**	**IF**										**7.60****	**0.060**
			DFF										**6.45***	**0.051**
			DM										**8.59****	**0.067**
			PL										**11.35****	**0.086**

**IF* days to initiation of flower, *DFF* days to 50% flowering, *DM* days to maturity, *PH* plant height (cm), *PB* number of primary branches, *SP* seeds per pod, *PL* pod length (cm)

[a]The *p* values that passed FDR criteria of Benjamini and Hochberg (1995)

[b]The *p* values significant for Bonferroni threshold (< 0.003125)

[c]The *p* values nonsignificant for FDR and Bonferroni threshold

LG linkage group, ***significant at ($p < 0.001$), **($p < 0.01$) and *($p < 0.05$) (Bold font for markers flanking QTLs for plant ideoype traits)

for determinacy trait as reported earlier by Mir et al. (2014). As the SNP was excluded from association analysis, we validated this marker through SMA where it explained up to 9% PV for IF, DFF, DM and PL. These results implied towards the potential role of *CcTFL1* gene in regulating flowering and earliness traits in pigeonpea. Earlier, SMA has been used successfully for establishment of MTAs in pigeonpea and groundnut (Singh et al. 2013; Mace et al. 2006; Mondal and Badigannavar 2010). Finally, we checked the discrimination efficiency of *CcTFL* 1 for determinacy trait and found 77.9 and 56.3% efficiency for IDT (734) and DT (167 bp) specific alleles, respectively. The difference in efficiencies may be attributable to varying sizes of the samples containing IDT and DT types.

The box plots generated to detect favorable allele of trait-linked SSRs suggested that the alleles $ASSR8_{150}$, $ASSR390_{170}$, $ASSR295_{140}$ and $CcTFL1_{734}$ significantly contributed to earliness and plant type traits in pigeonpea. Similar approach has been used in other crops to discover the contributing allele of a linked marker for a trait, for instance SNP alleles for plant height, grain yield and release of glucose in wheat (Bellucci et al. 2015), fruit quality in tomato (Albert et al. 2016), lint percentage in upland cotton (Su et al. 2016) and dry matter content in cassava (Rabbi et al. 2017).

Conclusion

In summary, we attempted to validate a set of QTLs for plant ideotype traits across diverse collection of pigeonpea. These putative QTL regions were reported previously in biparental populations (Kumawat et al. 2012; Geddam et al. 2014). We could validate only one SSR (ASSR295) that showed significant association with earliness traits. Also, we detected new SSRs (ASSR8, ASSR390 and ASSR610), which accounted for higher PV in comparison to previously reported markers. Additionally, we confirmed association of *CcTFL*1 with earliness trait in pigeonpea, and discrimination efficiency of this marker was computed with respect to determinacy trait. The information

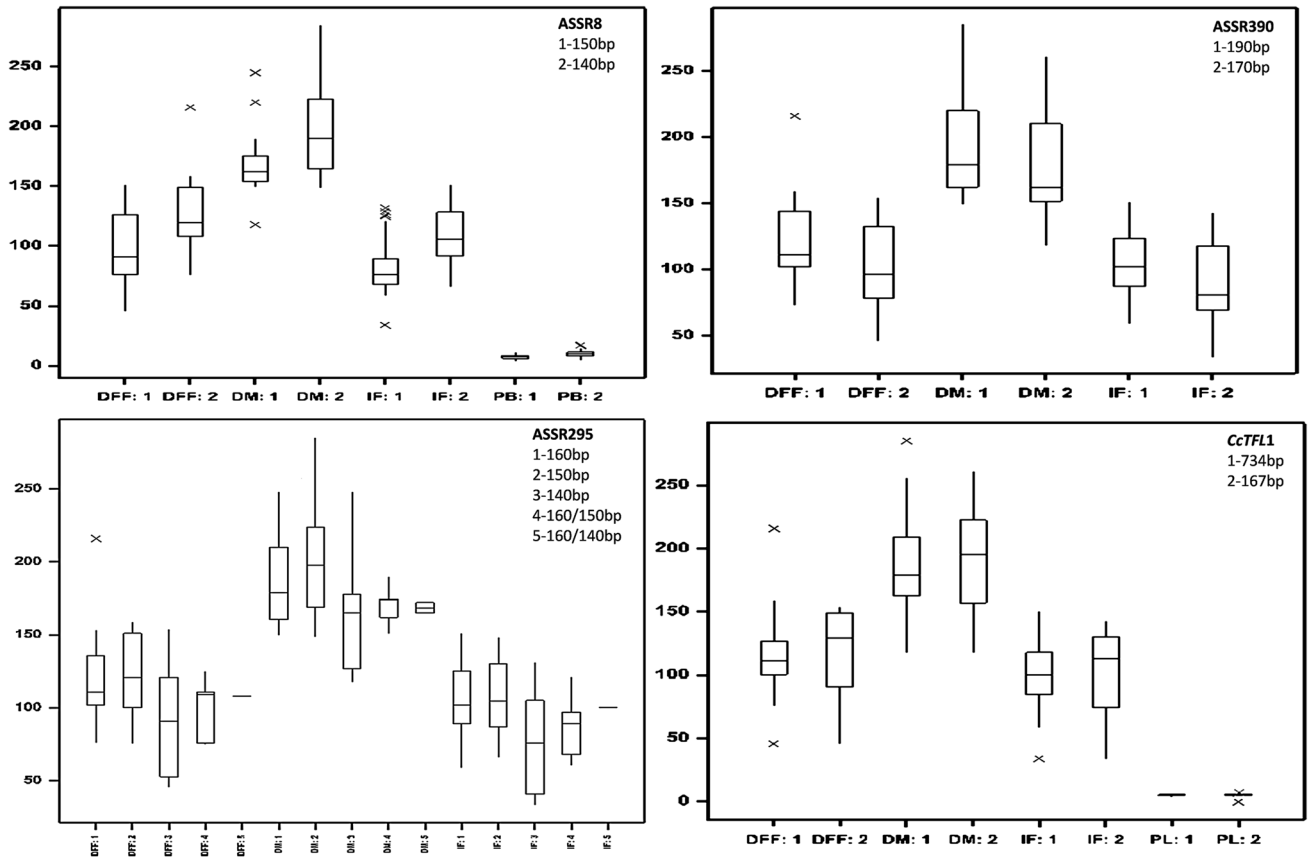

Fig. 4 Box plot showing phenotypic effects of major marker alleles for significantly associated EST–SSR loci for important plant ideotype traits

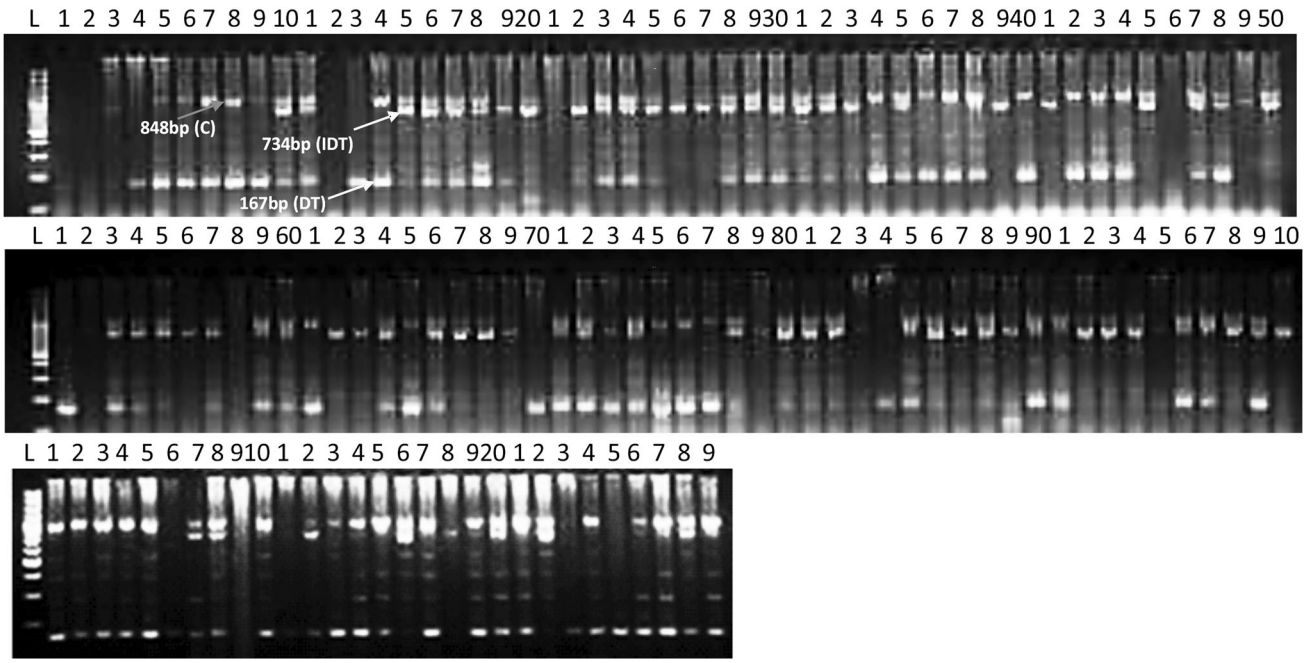

Fig. 5 Genotyping for *CcTFL1* specific SNP marker for determinacy trait in 120 pigeonpea genotypes

presented in this study could provide new avenues for exercising MAS to rapidly breed superior plant types of pigeonpea with early maturity and desirable growth habit.

Acknowledgements Financial support from Indian Council of Agricultural Research (ICAR), New Delhi is gratefully acknowledged.

Compliance with ethical standards

Conflict of interest Authors declare that there is no conflict of interest.

References

Agbagwa IO, Datta S, Patil PG, Singh P, Nadarajan N (2012) A protocol for high-quality genomic DNA extraction from legumes. Genet Mol Res 11:4632–4639

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 19:716–723

Albert E, Segura V, Gricourt J, Bonnefoi J, Derivot L, Causse M (2016) Association mapping reveals the genetic architecture of tomato response to water deficit: focus on major fruit quality traits. J Exp Bot 67:6413–6430

Amallah L, Taghouti M, Rhrib K, Gaboun F, Arahou M, Hassikou R, Diria G (2016) Validation of simple sequence repeats associated with quality traits in durum wheat. J Crop Sci Biotechnol 19:137–150

Babicki S, Arndt D, Marcu A, Liang Y, Grant JR, Maciejewski A, Wishart DS (2016) Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res 44 (W1):W147–W153

Bellucci A, Torp AM, Bruun S, Magid J, Andersen SB, Rasmussen SK (2015) Association mapping in Scandinavian winter wheat for yield, plant height, and traits important for second-generation bioethanol production. Front Plant Sci 6:1046

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc 57:289–300

Bhadru D (2010) Studies on genetic parameters and interrelationships among yield and yield contributing traits in pigeonpea (*Cajanus cajan* Millsp). Leg Res 33:23–27

Bohra A, Jha R, Pandey G, Patil PG, Saxena RK, Singh IP, Singh D, Mishra RK, Mishra A, Singh F, Varshney RK, Singh NP (2017) New hypervariable SSR markers for diversity analysis, hybrid purity testing and trait mapping in pigeonpea [*Cajanus cajan* (L.) Millspaugh]. Front. Plant Sci 8:377

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633

Chetukuri A, Vijayalakshmi P, Bhargavi V, Pavankumar D, Sreelaxmi A, Siddiq EA (2013) Correlation, variability and heritability in pigeonpea. Adv Biores 4:129–134

Dhanasekar P, Dhumal KN, Reddy KS (2010) Identification of RAPD markers linked to plant type gene in pigeonpea. Indian J Biotechnol 9:58–63

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

FAOSTAT (2016) Food and agriculture organization of United Nations. http://faostat.fao.org

Federer WT (1956) Augmented (or hoonuiaku) designs. Hawaii. Plant Res 2:191–208

Geddam SB, Raje RS, Prabhu KV, Singh NK, Chauhan DA, Jain P, Khare A, Yadav R, Tyagi A (2014) Validation of QTLs for earliness and plant type traits in pigeonpea (*Cajanus cajan* (L.) Millsp.). Indian J Genet Plant Breed 74:471–477

Hardy OJ, Vekemans X (2002) SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2:618–620

Hu Z, Zhang D, Zhang G, Kan G, Hong D, Yu D (2014) Association mapping of yield-related traits and SSR markers in wild soybean (*Glycine soja* Sieb. and Zucc.). Breed Sci 63:441–449

Kumawat G, Raje RS, Bhutani S, Pal JP, Mithra ASVRC, Gaikwad K, Sharma TR, Singh NK (2012) Molecular mapping of QTLs for plant type and earliness traits in pigeonpea (*Cajanus cajan* L. Millsp.). BMC Genet 13:84

Kump KL, Bradbury PJ, Wisser RJ, Buckler ES, Belcher AR, Oropeza-Rosas MA, Zwonitzer JC, Kresovich S, McMullen MD, Ware D, Balint-Kurti PJ, Holland JB (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. Nat Genet 43:163–168

Lackey JA (1977) A revised classification of the tribe Phaseoleae (Leguminosae, Papilionoideae), and its relation to canavanine distribution. Bot J Linn Soc 74:163–178

Lakhote SJ, Patil HE, Mali RA, Ingle P (2015) Genetic analysis for yield and yield contributing traits in vegetable type genotypes in pigeonpea (*Cajanus cajan* (L.). Int J Trop Agric 33:161–167

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic-structure of a tropical understory shrub, *Psychotria Officinalis* (Rubiaceae). Am J Bot 82:1420–1425

Lu Q, Zhang M, Niu X, Wang S, Xu Q, Feng Y, Wang C, Deng H, Yuan X, Yu H, Wang Y, Wei X (2015) Genetic variation and association mapping for 12 agronomic traits in indica rice. BMC Genom 16:1067

Lu Q, Niu X, Zhang M, Wang C, Xu Q, Feng Y, Yang Y, Wang S, Yuan X, Yu H, Wang Y, Chen X, Liang X, Wei X (2018) Genome-wide association study of seed dormancy and the genomic consequences of improvement footprints in rice (*Oryza sativa* L.). Front. Plant Sci 8:2213

Mace ES, Phong DT, Upadhyaya HD, Chandra S, Crouch JH (2006) SSR analysis of cultivated groundnut (*Arachis hypogaea* L.) germplasm resistant to rust and late leaf spot diseases. Euphytica 152:317–330

Malosetti M, Ribaut JM, Van Eeuwijk FA (2013) The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. Front Physiol 4:44

Meena SS, Verma SK, Choudhary R, Panwar RK, Singh JP (2017) Genetic variability and inter-relationship among yield characters in advance lines of pigeonpea [*Cajanus Cajan* (L.) Millsp.] grown at different altitudes. Chem Sci Rev Lett 6:1120–1128

Mir RR, Kudapa H, Srikanth S, Saxena RK, Sharma A, Azam S, Saxena K, Penmetsa RV, Varshney RK (2014) Candidate gene analysis for determinacy in pigeonpea (*Cajanus* spp.). Theor Appl Genet 127:2663–2678

Mondal S, Badigannavar AM (2010) Molecular diversity and association of SSR markers to rust and late leaf spot resistance in cultivated groundnut (*Arachis hypogaea* L.). Plant Breed 129:68–71

Odeny DA, Jayashree B, Ferguson M, Hoinsington D, Crouch J, Gebhardt C (2007) Development, characterization and utilization of microsatellite markers in pigeonpea. Plant Breed 126:130–136

Patil PG, Byregowda M, Vimarsha HS, Keerthi CM, Kundur PJ, Shashidhar HE (2015) Genetic variability on leaf morpho-anatomical traits in relation to sterility mosaic disease (SMD) resistance in pigeonpea. Aust J Crop Sci 9:621–628

Patil PG, Dubey J, Bohra A, Mishra RK, Saabale PR, Das A, Rathore M, Singh NP (2017) Association mapping to discover significant

marker-trait associations for resistance against Fusarium wilt variant 2 in pigeonpea [*Cajanus cajan* (L.) Millspaugh] using SSR markers. J Appl Genet 58:307–319

Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. Biometrika 58:545–554

Perrier X, Jacquemoud-Collet JP (2006) DARwin software. Paris: Centre de Cooperation Internationale en Recherche Agronomique Pour le De'veloppement (CIRAD)

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Rabbi IY, Udoh LI, Wolfe M, Parkes EY, Gedil MA, Dixon A, Ramu P, Jannink JL, Kulakow P (2017) Genome-wide association mapping of correlated traits in cassava: dry matter and total carotenoid content. Plant Genome 10:1–14

Raftery A (1986) A note on Bayes Factors for log-linear contingency table models with vague prior information. J R Stat Soc 48:249–250

Saroj SK, Singh MN, Kumar R, Singh T, Singh MK (2013) Genetic variability, correlation and path analysis for yield attributes in pigeonpea. Bioscan 8:941–944

Satheesh NSJ, Byregowda M, Dinesh HB, Miyan A, Pushpa HD (2013) Genetic analysis in advanced breeding lines (F_5) of pigeonpea [*Cajanus cajan* (L.) Millisp.]. Bioinfolet 10:196–200

Saxena KB (2008) Genetic improvement of pigeonpea-a review. Trop Plant Biol 1:159–178

Saxena KB, Sharma D (1990) Pigeonpea genetics. In: Nene YL, Hall SD, Sheila VK (eds) The pigeonpea. CAB International, Wallingford, pp 137–158

Shi C, Navabi A, Yu K (2011) Association mapping of common bacterial blight resistance QTL in Ontario bean breeding populations. BMC Plant Biol 11:52

Singh AK, Rai VP, Chand R, Singh RP, Singh MN (2013) Genetic diversity studies and identification of SSR markers associated with Fusarium wilt (*Fusarium udum*) resistance in cultivated pigeonpea (*Cajanus cajan*). J Genet 92:273–280

Singh IP, Bohra A, Singh F (2016) An overview of varietal development programme of pigeonpea in India. Legume Perspect 11:37–40

Su J, Fan S, Li L, Wei H, Wang C, Wang H, Song M, Zhang C, Gu L, Zhao S, Mao G, Wang C, Pang C, Yu S (2016) Detection of favorable QTL alleles and candidate genes for lint percentage by GWAS in Chinese upland cotton. Front Plant Sci 7:1576

Thudi M, Upadhyaya HD, Rathore A, Gaur PM, Krishnamurthy L, Roorkiwal M, Nayak SN, Chaturvedi SK, Basu PS, Gangarao NV, Fikre A, Kimurto P, Sharma PC, Sheshashayee MS, Tobita S, Kashiwagi J, Ito O, Killian A, Varshney RK (2014) Genetic dissection of drought and heat tolerance in chickpea through genome-wide and candidate gene based association mapping approaches. PLoS ONE 9:e96758

Van der Maesen LJG (1980) India is the native home of the pigeonpea. In: Arends JC, Boelema G, De Groot CT, Leeuwenberg AJM (eds) Liber Gratulatorius in nonerem HCD de Wit Landbouwhoge school Miscellaneous Paper no. 19, Wageningen, pp 257–262

Van der Maesen LJG (2003) Cajaninae of Australia (Leguminosae: Papilionoideae). Aust Syst Bot 16:219–227

Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA et al (2012) Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Nat Biotechnol 30:83–89

Varshney RK, Mohan S, Gaur PM, Gangarao NVPR, Pandey MK, Bohra A, Sawargaonkar SL, Chitikineni A, Kimurto PK, Janila P, Saxena KBC, Fikre A, Sharma MG, Rathore A, Pratap A, Tripathi S, Datta S, Chaturvedi SK, Mallikarjuna N, Anuradha G, Babbar A, Choudhary AK, Mhase MB, Bharadwaj CVG, Mannur DM, Harer PN, Guo B, Liang X, Nadarajan N, Gowda CLL (2013) Achievements and prospects of genomics-assisted breeding in three legume crops of the semi-arid tropics. Biotechnol Adv 31:1120–1134

Wu RL (1998) Genetic mapping of QTLs affecting tree growth and architecture in *Populus*: implication for ideotype breeding. Theor Appl Genet 96:447–457

Yu JM, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208