Euphytica (2017) 213:215
DOI 10.1007/s10681-017-2006-1

Population structure and association mapping of traits related to reproductive development in field pea

Yunfei Jiang · Marwan Diapari · Rosalind A. Bueckert · Bunyamin Tar'an · Thomas D. Warkentin

Received: 10 April 2017 / Accepted: 19 August 2017 / Published online: 29 August 2017

Abstract Field pea (*Pisum sativum*) is an important pulse crop globally for human consumption and livestock feed. A panel of 92 diverse pea cultivars was evaluated across nine environments and genotyped using 1536 single nucleotide polymorphisms (SNPs) arranged in a GoldenGate array. Population structure analysis revealed three subpopulations roughly consistent with the cultivar origin. Phenotyping included days to flowering (DTF), duration of flowering (DOF), number of reproductive nodes, number of pods on the main stem, percentage of pods set, percentage of pods retained with seed and pollen germination reduction due to heat stress. Association analyses identified a total of 60 SNPs significantly associated ($-\log_{10}p \geq 4.3$) with these seven reproductive development-related traits. Among these 60 marker-trait associations, 33 SNPs were associated with the onset of flowering, 8 SNPs with pod development and 19 SNPs with the number of reproductive nodes. No SNP marker was significantly associated with in vitro pollen germination reduction caused by high temperature stress. We found that 12 SNPs associated with DTF and 2 SNPs associated with DOF overlapped with the SNP markers associated with the number of reproductive nodes. Genomic regions associated with variation for reproductive development-related traits identified in this study provide grounds for future genetic improvement in pea.

Keywords Association mapping · Population structure · *Pisum sativum* · Reproductive development

Y. Jiang and M. Diapari are joint first authors.

Y. Jiang · M. Diapari · R. A. Bueckert · B. Tar'an · T. D. Warkentin (✉)
Department of Plant Sciences, University of Saskatchewan, 51 Campus Drive, Saskatoon, SK S7N 5A8, Canada
e-mail: tom.warkentin@usask.ca

M. Diapari · B. Tar'an · T. D. Warkentin
Crop Development Centre, University of Saskatchewan, 51 Campus Drive, Saskatoon, SK S7N 5A8, Canada

Present Address:
M. Diapari
Genomics and Biotechnology, London Research and Development Centre, Agriculture and Agri-Food Canada, London, ON, Canada

Introduction

Field pea (*Pisum sativum*) belongs to the family Fabaceae, subfamily Papilionoideae and is one of the economically most important pulse crops. It is a diploid ($2n = 2x = 14$) with a haploid genome size of 4.45 Giga base pairs (Gbp; Dolezel and Greilhuber 2010). Field pea is an economically important pulse crop in Western Canada and is a rich source of food for humans and fodder for animals. Pea is of great importance to cropping systems due to its ability to

fix nitrogen. Canada was the largest producer of field pea in the world in 2014 with a total production of 3.4 million tonnes, followed by China, Russia, USA, and India (FAOSTAT 2016).

Flowering time in pea is under complex genetic control (Vanhala et al. 2016). Days to flowering (DTF) in pea was not strongly associated with the site of origin and genetic diversity of the genes (Vanhala et al. 2016). More than 20 loci associated with flowering time and inflorescence development in pea have been identified (Weller and Ortega 2015). Time of flowering in pea varies widely across different locations and years (Truong and Duthion 1993). Photoperiod, temperature (vernalization and post-vernalization), and genotype are crucial factors for time of flowering in pea (Alcade et al. 1999; Alcalde and Larrain 2006; Weller et al. 2009). Flowering is controlled by six genes including *Lf, Sn, E, Dne, Ppd,* and *Hr* (Murfet 1985; Weller et al. 1997). The *Hr* (late flowering) haplotype is a major factor of flowering time in pea, but other genes also play a role in flowering time (Vanhala et al. 2016). Pea accessions carrying loss-of-function *Sn* (STERILE NODES) alleles of the gene flowered early (Liew et al. 2014). Flowering at different nodes is not synchronous but is sequential due to the indeterminate growth habit of pea. Thus flowering duration of the crop across production regions and across cultivars varies greatly because of the large degree of variability in the number of reproductive nodes (Roche et al. 1998). The number of reproductive nodes in pea is an important component affecting seed yield (Roche et al. 1998). Pollen viability is another trait influencing seed yield in various crops. For example, yield has been reduced by loss of pollen viability in chickpea (*Cicer arietinum*; Devasirvatham et al. 2012), common bean (*Phaseolus vulgaris*; Konsens et al. 1991), and cowpea (*Vigna unguiculata*; Ahmed et al. 1992) in stress.

Association mapping and linkage analysis are two commonly used and complementary approaches to dissect complex traits (Yu et al. 2008; Varshney et al. 2012). Linkage analysis, or quantitative trait locus (QTL) mapping, usually detects broad chromosome regions associated [10–30 cM (centimorgans); a 10-cM chromosome interval corresponding to approximately 2.1 Mb and 400 genes in *Arabidopsis* and about 12.4 Mb and 310 genes in maize] with traits of interest with relatively low marker coverage, whereas linkage disequilibrium (LD) based association mapping has greater genetic resolution (many rounds of meiosis) compared to linkage mapping based on bi-parental mapping populations (Salvi and Tuberosa 2005; Myles et al. 2009; Morrell et al. 2012; Varshney et al. 2012). Several QTL mapping studies have been conducted for various traits of interest in pea: (1) Seed yield, seed protein concentration and days to maturity (Tar'an et al. 2004), (2) yield, yield components, and seed protein content (Krajewski et al. 2012), (3) seed yield, seed weight, seed number, harvest index, node of first flower, the number of flowering nodes, and total node number (Timmerman-Vaughan et al. 2005), (4) seed yield, seed protein content, seed weight, onset of flowering, plant height and lodging susceptibility (Ferrari et al. 2016), and (5) DTF duration of flowering (DOF), days to flowering termination, pod number, the number of reproductive nodes, seed number per pod, thousand seed weight (TSW) and yield (Huang et al. 2017).

Association mapping, also known as "linkage disequilibrium mapping", uses ancestral recombination events and natural genetic diversity within a population to dissect quantitative traits (Zhu et al. 2008; Myles et al. 2009; Morrell et al. 2012). Association mapping has been conducted in several legume crops including chickpea (Thudi et al. 2014; Diapari et al. 2014), soybean (*Glycine max*; Li et al. 2011), alfalfa (*Medicago sativa*; Sakiroglu et al. 2012), and common bean (Shi et al. 2011; Nemli et al. 2014). Several association mapping studies have been reported in pea: (1) 49 phenotypic traits related to seed nutrients (Kwon et al. 2012), (2) iron, zinc and selenium concentrations in seed (Diapari et al. 2015), (3) seed lipid content (Ahmad et al. 2015), (4) agronomic and quality traits such as disease resistance, flower color, seed type/color and seed mineral concentration (Cheng et al. 2015), and (5) partial resistance to *Aphanomyces euteiches* (Desgroux et al. 2016). In addition to linkage and association mapping conducted in this crop species, genotyping-by-sequencing based genomic selection for seed yield under severe terminal drought was reported in pea (Annicchiarico et al. 2017). Genomic areas associated with early flowering and high yield co-located under severe terminal drought (Annicchiarico et al. 2017).

Although several QTL mapping studies focusing on important agronomic traits have been documented in field pea, the complementary approach—association mapping is needed to dissect traits related to

reproductive development such as DTF, DOF, the number of reproductive nodes, the number of pods, percentage of pods set, percentage of pods retained with seed, or in vitro pollen germination reduction under normal conditions and, in particular, to an abiotic stress such as high temperature. The objectives for this study were to examine the G × E interaction in DTF, analyze the population structure of a panel of 92 pea cultivars and to identify DNA markers associated with reproductive development related traits using previously published gene-anchored SNP markers (Sindhu et al. 2014).

Materials and methods

Plant materials and growing conditions

A panel of 92 diverse field pea cultivars assembled at the Crop Development Centre (CDC), University of Saskatchewan and referred to as the pea association mapping (PAM) panel was grown at nine site-years (environments) and one controlled environment. Among these 92 cultivars, 31 were from Western Europe, 17 were from the CDC, 16 were from Eastern Europe, 14 were from Agriculture and Agri-Food Canada (AAFC), 10 were from USA, and 4 were from Australia. Association studies were also previously conducted in the PAM for iron, zinc and selenium concentration (Diapari et al. 2015), with the addition of two wild relative accessions [P651 (*Pisum fulvum*) and PI344538 (*Pisum sativum* subsp. *elatius*)].

Field trials were arranged using a randomized completed block design (RCBD) and grown at Sutherland (near Saskatoon; lat. 52°10′N, long. 106°41′W; Dark Brown chernozemic soil zone) and Rosthern (lat. 52°40′N, long. 106°20′W; Black soil zone), Saskatchewan in 2011, 2012, and 2013 and Rosthern in 2015. Seventy-five seeds of each cultivar were planted in a 1 m × 1 m microplot with 4 rows and 0.25 m spacing between rows. Six blocks were grown at each site-year at Sutherland and Rosthern in 2011 and 2012, and three blocks for each site-year at Rosthern in 2013 and 2015. The PAM panel was also evaluated near Yuma, Arizona with two seeding dates. The early seeding date was February 1 in 2012, and the late seeding date was February 27 in 2012. The weather conditions are listed in Table 1. Daily maximum temperatures were used as an indicator of heat stress (Bueckert et al. 2015), because fruit and flower abortion were observed under field conditions when the daily maximum temperatures exceeded 28 °C (Bueckert et al. 2015).

The PAM panel was also tested in a growth chamber (Conviron®) for a controlled temperature environment using a completely randomized design (CRD). A total of 184 pots (92 pea cultivars × 2 pots per cultivar) of 3.8 L volume (3 plants per pot) were seeded with Sunshine Gro® mix (Seba Beach, AB, Canada) and slow-release fertilizer (14-14-14, Type 100, Nutricote®, Brampton, ON, Canada). The pot dimensions were 15.9 cm depth and 16.5 cm diameter. Plants were thinned to two plants per pot approximately 2 weeks after seeding. Plants received the first application (500 mL per pot) of half strength modified Hoagland's culture solution (Hoagland and Arnon 1938) at 3 weeks after seeding and the second application (500 mL per pot) at the early flowering stage. Soil medium moisture was monitored carefully and plants were watered as necessary to avoid drought stress. Plants were grown at 24/18 °C day/night temperatures with the 16/8 h photoperiod and illumination levels of 450–500 μmol m^{-2} s^{-1} from cool fluorescent tubes.

Phenotyping

Seven phenotypic traits including DTF, DOF, the number of reproductive nodes, the number of pods on the main stem, percentage of pods set, percentage of pods retained with seed (Yuma only), and in vitro pollen germination reduction due to high temperature stress were collected. DTF was determined as the number of days from planting until 50% of the plants per plot were at the flowering stage at Rosthern in 2011, 2012, 2013 and 2015, Sutherland in 2011 and 2012, and Yuma in 2012 with two different seeding dates. For the controlled environment, DTF were determined when flower buds at the first reproductive node on the main stem were visible. DOF was calculated as days to flowering termination (DTFT) minus DTF, where DTFT was determined as the number of days from sowing until 50% of plants per plot reached flower termination. The number of reproductive nodes and pods on main stems were counted based on two randomly selected plants from each micro-plot at physiological maturity. The percentage of pods set, based on the expected or

Table 1 Seeding date, monthly mean temperature, monthly average maximum temperature, the number of days when the daily maximum temperatures were greater than 28 °C, and total precipitation at multiple site-years

Site (year)	Seeding date	Monthly mean temperature (°C)				Average maximum temperature (MT, °C)				Number of days when MT > 28 °C	Total precipitation (mm)				Total
		May	June	July	Aug	May	June	July	Aug	From May to Aug	May	June	July	Aug	
Rosthern (2011)[a]	May 15	10.8	15.5	18.0	16.3	17.9	21.1	23.8	22.9	7	24.6	131.4	83.4	31.5	270.8
Sutherland (2011)[b]	May 3	10.9	15.5	18.4	17.2	17.9	21.3	24.6	24.9	6	17.5	94.4	68.6	16.5	197.0
Rosthern (2012)	May 11	10.2	15.9	20.0	17.4	16.4	21.6	25.6	23.9	11	92.5	116.0	92.3	63.5	364.2
Sutherland (2012)	May 16	10.1	15.8	19.7	17.3	16.4	21.5	25.3	23.9	7	108.0	121.1	80.9	48.5	358.5
Rosthern (2013)	May 15	12.8	15.3	17.2	17.4	20.5	20.6	23.3	24.6	15	13.5	148.8	62.4	17.4	242.0
Rosthern (2015)	May 13	9.4	16.4	18.8	16.9	17.8	24.1	25.8	23.8	28	10.9	37.4	80.5	75.9	204.7
		Jan	Feb	Mar	Apr	Jan	Feb	Mar	Apr	From Jan to Apr	Jan	Feb	Mar	Apr	Total
Yuma (2012)[c]	Jan	15.9	16.0	18.8	23.4	22.8	23.2	26.6	31.2	43	0	0.8	0.3	1.3	2.4

[a] Data were calculated based on the average of Saskatoon and Prince Albert from Environment Canada, http://climate.weather.gc.ca

[b] Data were obtained from Environment Canada

[c] Data were obtained from Weather Underground, https://www.wunderground.com. Note that Yuma is an irrigated site

theoretical maximum number of pods, was calculated according to Eq. 1 at Rosthern and Sutherland in 2012:

$$\text{Pod set} = \text{Pods}/(\text{Rnode} \times 2) \quad (1)$$

where Pods is the number of pods on the main stem at harvest, and Rnode is the number of reproductive nodes with each node having two flowers per nodal raceme. Rnode × 2 is the theoretical or total number of potential pods possible.

At Yuma (2012) the number of reproductive nodes was not counted, so pod set was not calculated. Instead, the number of pods with seed plus dead pods retained on the main stem were added to give the total number of pods at harvest. The percentage of pods retained with seed were calculated from the total number of pods at harvest on the main stem, and provided an assessment of pod abortion in heat.

Pollen grains collected from the field conditions (Rosthern in 2015) and the controlled environment (University of Saskatchewan phytotron in 2015) were treated as controls. Pollen samples were incubated in the dark for 24 h in two incubators with one at the control or unstressed temperature of 24 °C and the other at the stress treatment temperature of 42 °C. Two replications were utilized for each cultivar at each temperature regime in each environment (736 samples in total—92 cultivars × 2 temperatures × 2 replications at each environment × 2 environments). The detailed information about the in vitro pollen germination assay was previously described with some modifications (Lahlali et al. 2014; Jiang et al. 2015). Pollen grains were incubated under light conditions in Lahlali et al. (2014) and Jiang et al. (2015), but dark conditions were used to incubate pollen grains in this study. In vitro pollen germination reduction caused by high temperature stress was calculated according to Eq. 2:

$$\text{Pollen germination reduction} = (\text{Control PGP} - \text{Heat stressed PGP})/\text{Control PGP} \quad (2)$$

where control PGP is the percentage of germinated pollen grains at controlled temperature, and heat stress PGP is the percentage of germinated pollen grains at high temperature.

Phenotypic data analysis

Homogeneity of variances for location effects was tested using Levene's test. The p values from the Levene's test for location effects on DTF and in vitro pollen germination reduction due to high temperature stress were less than 0.05, so the datasets were analyzed separately based on different environments. Although p values from the Levene's test for location effects on the number of reproductive nodes, the number of pods on main stem, and potential pod set were greater than 0.05, the G × E interaction terms were significant ($p < 0.05$), so data sets for each environment were analyzed separately. Variance components of genotype, environment, the G × E interaction, block within environment, and the residual were analyzed for each trait using the generalized linear model (GLM) and all factors were considered random effects. Broad sense heritability (H^2) was calculated as: $H^2 = \sigma_g^2/(\sigma_g^2 + \sigma_{ge}^2/n + \sigma^2/nb)$, where σ_g^2 is the genetic variance, σ_{ge}^2 is the variance of genotype and environment interaction, σ^2 is the error variance, n is the number of environments, and b is the number of replications within each environment (Wang et al. 2016).

The G × E interaction for DTF was assessed using the additive main effects and multiplicative interaction (AMMI) model using SAS (SAS Institute Inc.) as described in Kang et al. (2004). The AMMI model is a combination of analysis of variance (ANOVA) and principal component analysis (PCA) that partitions the variability of genotype, environment, and the G × E interaction model (Marjanović-Jeromela et al. 2011; Sabaghnia et al. 2008).

Association mapping

DNA extraction and genotyping using an Illumina GoldenGate array (Illumina Inc., San Diego, CA, USA) with a 1536-SNP pea OPA (Ps 1536) were previously described (Sindhu et al. 2014; Diapari et al. 2015). Among these 1536 SNPs, 1233 polymorphic SNPs were retained for further analysis, and the other 303 SNPs were removed due to being monomorphic or having a large ratio (35% or greater) of missing data (Diapari et al. 2015). The location of these SNPs on the pea genome was reported by Sindhu et al. (2014). Two cultivars, 40–10 and CDC Dundurn, were also removed from further analysis due to large proportions of missing data (82% missing data for 40–10 and 85% missing data for CDC Dundurn), thus 92 cultivars were retained in the population structure and association analysis. Population structure was analyzed using two different methods: (1) the Bayesian-based clustering approach with an admixture model using the Structure 2.3.4 software (Pritchard et al. 2000); (2) discriminant analysis of principal components (DAPC), a multivariate method, using the R package "Adegenet" (Jombart 2008; Jombart et al. 2010). DAPC partitions genetic variation into a between-group component and a within-group component. DAPC utilizes data transformation using Principal Component Analysis (PCA) before discriminant analysis (DA) is applied. In the Bayesian method using the Structure software, the number of sub-populations (k) was set to be from 1 to 10 with 5 simulations, 100,000 burn-in phase, and 100,000 Monte Carlo Markov Chain replicates (MCMC). Based on the maximum likelihood and delta K (ΔK) values, the number of sub-populations was determined using Structure Harvester (http://taylor0.biology.ucla.edu/structureHarvester/; Evanno et al. 2005; Earl and VonHoldt 2012). For the DAPC, the SNP dataset was first converted to the genepop format using the R package "diveRsity" before using the "Adegenet" R package. The optimum number of sub-populations was determined with the smallest number of Bayesian information criterion (BIC). Both these methods allow for a probabilistic assignment of individuals to each group.

Among these 1233 SNP markers, 943 markers were retained by considering a minor allele frequency (MAF) greater than 0.05 to remove low coverage marker sites. Estimates for clustering membership (Q matrix) from the DAPC method using the "Adegenet" R package were used as covariate in the general linear model (GLM) in association analysis using the software of Trait Analysis by Association, Evolution and Linkage (TASSEL version 5.2.19; Bradbury et al. 2007). The significant threshold ($-\log_{10}p = 4.3$) was adjusted using Bonferroni multiple test correction (0.05/943), for $p < 0.05$ and the denominator 943 was the total number of SNPs tested after marker filter alignment.

Results

Population structure

The optimum number of sub-populations (K) was determined using the largest value of Delta K using the

Bayesian clustering method in the Structure software. The optimum k was ambiguous between 2 and 3, since the values of Delta K were similar when K was equal to 2 and 3, and Delta K decreased dramatically when K continued to increase (Fig. 1a). Therefore, the identification of the clusters in this population was validated using the function of 'find.clusters' with the Adegenet R package. The optimum number of clusters in this population was 3, because BIC values decreased until k = 3 (Fig. 1b). These 92 pea cultivars were grouped into 3 clusters as shown in the DAPC scatterplot (Fig. 1c). The cultivars and their origins under individual clusters are listed in Table 2. Subpopulation 1 included the majority of cultivars released by Crop Development Centre, University of Saskatchewan and Agriculture and Agri-Food Canada (AAFC), as well as two cultivars (Highlight and SW-Marquee) from Western Europe (Table 2). Subpopulation 2 consisted of all of the Australian cultivars, the majority of cultivars from Eastern Europe and the USA, five AAFC cultivars, five cultivars from Western Europe, and two CDC cultivars (CDC-Vienna and MFR043) (Table 2). The majority of cultivars from Western Europe belonged to subpopulation 3, and this cluster also included five AAFC cultivars, two CDC cultivars (CDC Sage and CDC Striker), three cultivars from Eastern Europe, and three cultivars from the USA (Table 2).

Days to flowering

DTF varied depending on the growing environment (Table 3). In general, pea plants grown under environmentally controlled conditions took less time to reach flowering compared to plants grown under field conditions (Table 3). The range of DTF (the maximum DTF minus the minimum DTF) varied from 7 to 19 days under field conditions (Table 3). However, the range of DTF was 56 days in controlled conditions, reflected by the largest variance and coefficient of variance among the nine environments tested (Table 3). In general, plants at Sutherland took approximately 5–7 days longer to reach flowering compared to Rosthern (Table 3). Acceleration of DTF due to late seeding and exposure to warmer temperatures and a greater proportion of slightly longer days compared to early seeding was observed at Yuma in 2012, with mean DTF values of 63 and 57 days for early seeding and late seeding, respectively (Table 3). The broad sense heritability of DTF was 0.67 (Table 4). With three multiplicative components, this AMMI model explained 77.5% of the total residual variability contributed to the G × E interaction (Fig. 2b, c). Variability among environments was greater than variability among genotypes.

A total of 22 markers across seven linkage groups were significantly associated with DTF under environmentally controlled conditions (Table 5). PsC21767p87 on LG IV was significantly associated with DTF at Rosthern in 2013, and explained 20% of the variation for DTF (Table 5). The same marker was also significantly associated with DTF under environmentally controlled conditions, and this marker explained 24% of the variation in the trait.

Duration of flowering

DOF varied depending on the growing environment (Table 3). DOF at Rosthern in 2012 and 2015 was longer than DOF at Sutherland in 2012, with the averages of DOF being 25, 22, and 14 days at Rosthern in 2012, Rosthern in 2015, and Sutherland in 2012, respectively (Table 3). The broad sense heritability of DOF was 0.78 (Table 4). A total of 6 markers in linkage groups 2, 3 and 6 were significantly associated with DOF at Sutherland in 2012 (Table 5).

Number of reproductive nodes

The average number of reproductive nodes was 8 at Rosthern in 2012, and 7 at Sutherland in 2012 and Rosthern in 2015 (Table 3). The coefficient of variation for number of reproductive nodes (24.4–32.0%) was relatively large (Table 3), reflected in the relatively large error variance (55.1%) for this trait in Table 4. The broad sense heritability of the number of reproductive nodes was 0.80 (Table 4). Nineteen markers were significantly associated with the number of reproductive nodes at Rosthern in 2012 and 2015 (Table 5). PsC17990p348 on LG III, PsC4940p155 on LG VI, and unmapped PsC12883p342, were significantly associated with the number of reproductive nodes at Rosthern in 2012 and these three markers explained 20, 23 and 25% variation in this trait, respectively. Among these three markers, PsC17990p348 and PsC4940p155 were also significantly associated with DOF (Table 5). A total of 16 markers was significantly associated with the number

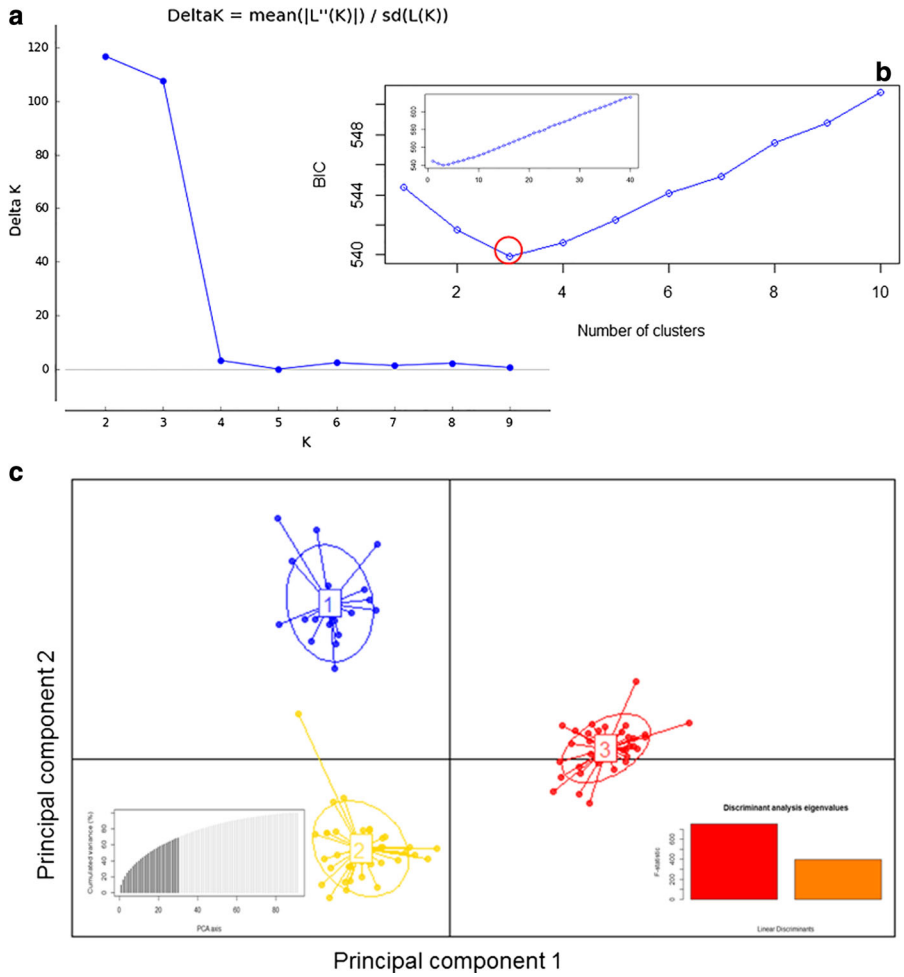

Fig. 1 Population structure analysis in a diversity panel of 92 pea cultivars. **a** Δk was used to determine the optimum k value for population structure using the Bayesian clustering method. **b** Inference of the number of clusters based on values of Bayesian information criterion (BIC). The selected number of clusters (3), corresponding to the smallest BIC, is circled in *red*. The inset represents the global results with the number of clusters up to 40, whereas the main figure indicates the detail of clusters up to 10. **c** Scatterplot of the discriminant analysis of principal components (DAPC) using the first two principal components (PCs) in a diversity panel of 92 pea cultivars. Ninety-two dots represent pea cultivars. The inset at the *bottom left* corner represents the number of principal components retained (35) in the DAPC; the inset at the *bottom right* corner represents the *bar* plot of eigenvalues for the discriminant analysis and two discriminant functions used in the DAPC

of reproductive nodes at Rosthern in 2015 and these markers explained 23–31% of the phenotypic variation in this trait (Table 5 and Fig. 3). Among these 16 markers detected at Rosthern in 2015, 11 out of 16 markers were co-localized with DTF (Fig. 3 and Table 5).

Pod number

The number of pods on the main stem varied greatly among cultivars in individual environments (Table 3). The number of pods ranged from 1 to 20 with an average of 7 at Rosthern in 2012, and varied from 1 to 18 with an average of 6 at Sutherland in 2012 (Table 3). Late seeding reduced the number of pods, with the average being 8 and 4 for early seeding and late seeding at Yuma in 2012, respectively (Table 3). The broad sense heritability of the number of pods was 0.51 (Table 4). The coefficient of variation for number of pods on main stem was relatively large (39.6–47.7%; Table 3. Two unmapped markers, PsC27644p242 and PsC6387p181, were significantly

Table 2 Pea cultivars and their origins under three individual clusters (k = 3) based on population structure analysis using discriminant analysis of principal components in a panel of 92 diverse pea cultivars

Cluster 1		Cluster 2		Cluster 3	
Cultivar (entry)	Origin	Cultivar (entry)	Origin	Cultivar (entry)	Origin
Hugo (72)	AAFC, Canada	Argus (71)	AAFC, Canada	Agassiz (58)	AAFC, Canada
P0309-09 (88)	AAFC, Canada	MI3360 (54)	AAFC, Canada	MP1401 (36)	AAFC, Canada
P0316-04 (90)	AAFC, Canada	MI3391 (53)	AAFC, Canada	Reward (26)	AAFC, Canada
P0321-08 (89)	AAFC, Canada	MPG87 (15)	AAFC, Canada	Stella (73)	AAFC, Canada
P0322-01 (91)	AAFC, Canada	Trapper (49)	AAFC, Canada	CDC Sage (41)	CDC, Canada
CDC 1-150-81 (56)	CDC, Canada	02H016P-03HO2004-06TGVP004 (74)	Australia	CDC Striker (27)	CDC, Canada
CDC 1-2347-144 (57)	CDC, Canada	03H107P-04HO2026 (75)	Australia	SGL-2024 (87)	Eastern Europe
CDC Acer (46)	CDC, Canada	03H267-04HO2006 (76)	Australia	TMP 15131 (4)	Eastern Europe
CDC Bronco (19)	CDC, Canada	Kaspa (32)	Australia	TMP 15213 (16)	Eastern Europe
CDC Centennial (20)	CDC, Canada	CDC Vienna (51)	CDC, Canada	Aragorn (94)	USA
CDC Golden (21)	CDC, Canada	MFR043 (34)	CDC, Canada	PS05100632 (77)	USA
CDC Leroy (47)	CDC, Canada	Naparnyk (48)	Eastern Europe	PS05100840 (78)	USA
CDC Meadow (23)	CDC, Canada	TMP 15116 (1)	Eastern Europe	Alfetta (38)	Western Europe
CDC Montero (40)	CDC, Canada	TMP 15121 (3)	Eastern Europe	Belote (65)	Western Europe
CDC Mozart (22)	CDC, Canada	TMP 15133 (5)	Eastern Europe	Bilboquet (64)	Western Europe
CDC Patrick (6)	CDC, Canada	TMP 15155 (7)	Eastern Europe	Carneval (35)	Western Europe
CDC Treasure (2)	CDC, Canada	TMP 15159 (8)	Eastern Europe	Carrera (55)	Eastern Europe
Cutlass (18)	CDC, Canada	TMP 15164 (9)	Eastern Europe	Cooper (28)	Western Europe
Highlight (52)	Western Europe	TMP 15169 (10)	Eastern Europe	Crackerjack (61)	Western Europe
SW-Marquee (39)	Western Europe	TMP 15179 (11)	Eastern Europe	Delta (45)	Western Europe
		TMP 15181 (12)	Eastern Europe	DS-Admiral (24)	Western Europe
		TMP 15202 (13)	Eastern Europe	Eclipse (25)	Western Europe
		TMP 15206 (14)	Eastern Europe	Espace (42)	Western Europe
		TMP 15221 (17)	Eastern Europe	FDP2010 (86)	Western Europe

Table 2 continued

Cluster 1		Cluster 2		Cluster 3	
Cultivar (entry)	Origin	Cultivar (entry)	Origin	Cultivar (entry)	Origin
		Fallon (84)	USA	Hardy (68)	Western Europe
		Lacy-Lady (82)	USA	Lido (92)	Western Europe
		Laxtons-Superb (80)	USA	Nitouche (29)	Western Europe
		Mini (83)	USA	Orb (37)	Western Europe
		PS05101142 (79)	USA	Polstead (59)	Western Europe
		Rally (85)	USA	Prelude (69)	Western Europe
		Superscout (81)	USA	Rialto (93)	Western Europe
		Cartouche (70)	Western Europe	Rocket (67)	Western Europe
		Lucy (63)	Western Europe	Rose (66)	Western Europe
		Radley (50)	Western Europe	SW-Sergeant (30)	Western Europe
		Rambo (33)	Western Europe	Terese (43)	Western Europe
		Torsdag (44)	Western Europe	Woody (62)	Western Europe

AAFC Agriculture and Agri-Food Canada, *CDC* crop development center

associated with the number of pods at Sutherland 2012, and these markers explained 23 and 20% of the variation in this trait (Table 5). The unmapped marker, PsC19517p115, was significantly associated with pod number at Yuma in 2012 with early seeding (Table 5).

Pod set

Pod set as a percentage of the total number of potential pods on the main stem had a wide range of variation among cultivars at individual environments. It ranged from 7 to 97% with a mean of 45% among these 92 cultivars at two locations in SK (Table 3). The coefficient of variation for potential pod set was relatively large (32.2–34.6%; Table 3), reflected in the relatively large error variance (74.2%) for this trait in Table 4. The broad sense heritability of potential pod set was 0.44 (Table 4). PsC8016p73, located on LG III, was significantly associated with pod set at Rosthern in 2012, and this marker explained 21% of the variation in this trait (Table 5). However, no marker was associated with pod set at Sutherland in 2012 (Table 5).

Pod set in heat

About 10% of the pods set aborted in early seeded pea plants at Yuma (2012) and 15% of the pods aborted in later seeded pea plants because heat stress was significant. Percentage of pods that contained at least one seed had a range of 42–100% with an average of 91.2% for all the pods of cultivars at early seeding in Yuma 2012, which was greater than late seeding with an average of 84.9% in pod retention (Table 3). The broad sense heritability of percentage of pod set was 0.26 (Table 4). Four markers, PsC11375p247 on LG I, PsC7631p74 on LG IV, PsC2509p330 on LG V and PsC1106p196 unmapped, were significantly associated with percentage of pods retained (Table 5). These

Table 3 Descriptive statistics for seven reproductive development related traits in a panel of 92 diverse pea cultivars

	Mean	SD	Variance	CV (%)	Mini-mum	Maxi-mum	Range[a]
Days to flowering							
Rosthern—2011 (6 reps)	51	0.9	0.8	1.7	48	55	7
Sutherland—2011 (3 reps)	58	1.9	3.5	3.2	56	63	7
Rosthern—2012 (3 reps)	56	2.5	6.3	4.5	47	62	15
Sutherland—2012 (3 reps)	61	2.2	4.8	3.6	57	70	13
Yuma—2012, early seeding (2 reps)	63	2.8	7.8	4.4	54	70	16
Yuma—2012, late seeding (2 reps)	57	4.3	18.1	7.5	48	64	16
Rosthern—2013 (3 reps)	50	2.7	7.6	5.5	43	58	15
Growth chamber—2015 (3 reps)	40	8.2	67.1	20.4	22	78	56
Rosthern—2015 (3 reps)	52	2.8	8.0	5.5	45	64	19
Duration of flowering (days)							
Rosthern—2012 (2 reps)	25	7.7	59.6	31.3	13	46	33
Sutherland—2012 (2 reps)	14	9.0	80.3	66.2	3	42	39
Rosthern—2015 (3 reps)	22	7.0	49.0	32.4	10	43	33
Number of reproductive nodes							
Rosthern—2012 (4 reps)	8	1.9	3.6	24.4	4	16	12
Sutherland—2012 (4 reps)	7	2.1	4.5	32.0	2	18	16
Rosthern—2015 (3 reps)	7	2.1	4.4	28.6	4	21	17
Number of pods							
Rosthern—2012 (4 reps)	7	2.8	8.0	39.8	1	20	19
Sutherland—2012 (4 reps)	6	2.8	7.9	47.7	1	18	17
Yuma—2012, early seeding (2 reps)	8	3.3	10.9	39.6	0	18	18
Yuma—2012, late seeding (2 reps)	4	2.1	4.4	46.8	0	11	11
Pod set (%)							
Rosthern—2012 (4 reps)	46.0	14.8	219.6	32.2	13	97	85
Sutherland—2012 (4 reps)	44.6	15.4	238.5	34.6	7	88	80
Pods retained with seed (%)							
Yuma—2012, early seeding (2 reps)	91.2	9.9	97.0	10.8	42	100	58
Yuma—2012, late seeding (2 reps)	84.9	15.3	234.2	18.0	20	100	80
Reduction of pollen germination percentage							
Growth chamber—2015 (2 reps)	0.90	0.19	0.037	21.2	0	1	1
Rosthern—2015 (2 reps)	0.74	0.29	0.084	39.1	0	1	1

SD standard deviation, *CV* coefficient of variance

[a] Range represents the range between the minimum value and the maximum value

four markers explained 19–26% of the phenotypic variation in this trait (Table 5).

Pollen germination reduction

Data of pollen germination reduction for samples collected from growth chambers were square transformed. Pollen samples collected from the field were more robust compared to pollen samples collected from the controlled environment, because the reduction percentage of in vitro pollen germination due to high temperature of pollen samples collected from the controlled environment was greater compared to the field conditions, with mean values of 90 and 74% at growth chamber conditions and Rosthern in 2015, respectively (Table 3). The broad sense heritability of

Table 4 Variance components of environment, genotype, and their interaction and broad sense heritability (H^2) on seven reproductive development related traits in a panel of 92 diverse pea cultivars

Source	Days to flowering		Duration of flowering (days)		Number of reproductive nodes		Number of pods on main stem		Pod set (%)		Pods retained with seed (%) at Yuma		Pollen germination reduction (%)	
	Variance	% of Total	Variance	% of Total	Variance	% of Total	Variance	% of Total	Variance	% of Total	Variance	% of Total	Variance	% of Total
Environment (E)	42.83***	76.5	32.18***	33.7	0.25 NS	5.4	1.93***	19.6	0.00NS	0.0	16.24NS	8.8	0.013***	17.3
Genotype (G)	2.03***	3.6	25.50***	26.7	1.23***	27.0	0.90***	9.2	24.38**	10.4	14.37NS	7.8	0.002NS	2.1
G*E	7.98***	14.2	9.63***	10.1	0.20NS	4.4	1.90***	19.3	18.03*	7.7	13.19NS	7.2	0.012*	16.1
Block(E)	0.10***	0.2	0.23NS	0.2	0.33***	7.2	0.27***	2.8	18.02***	7.7	5.18*	2.8	0.000NS	0.0
Error	3.07	5.5	28.02	29.3	2.56	56.1	4.84	49.2	173.47	74.2	135.55	73.5	0.047	64.6
Total	56.01		95.57		4.56		9.85		233.9		184.53		0.073	
Heritability (H^2)	0.67		0.78		0.8		0.51		0.44		0.26		0.1	

NS not significant at the 0.05 level

* Significant at the 0.05 level of probability

** Significant at the 0.01 level of probability

*** Significant at the 0.001 level of probability

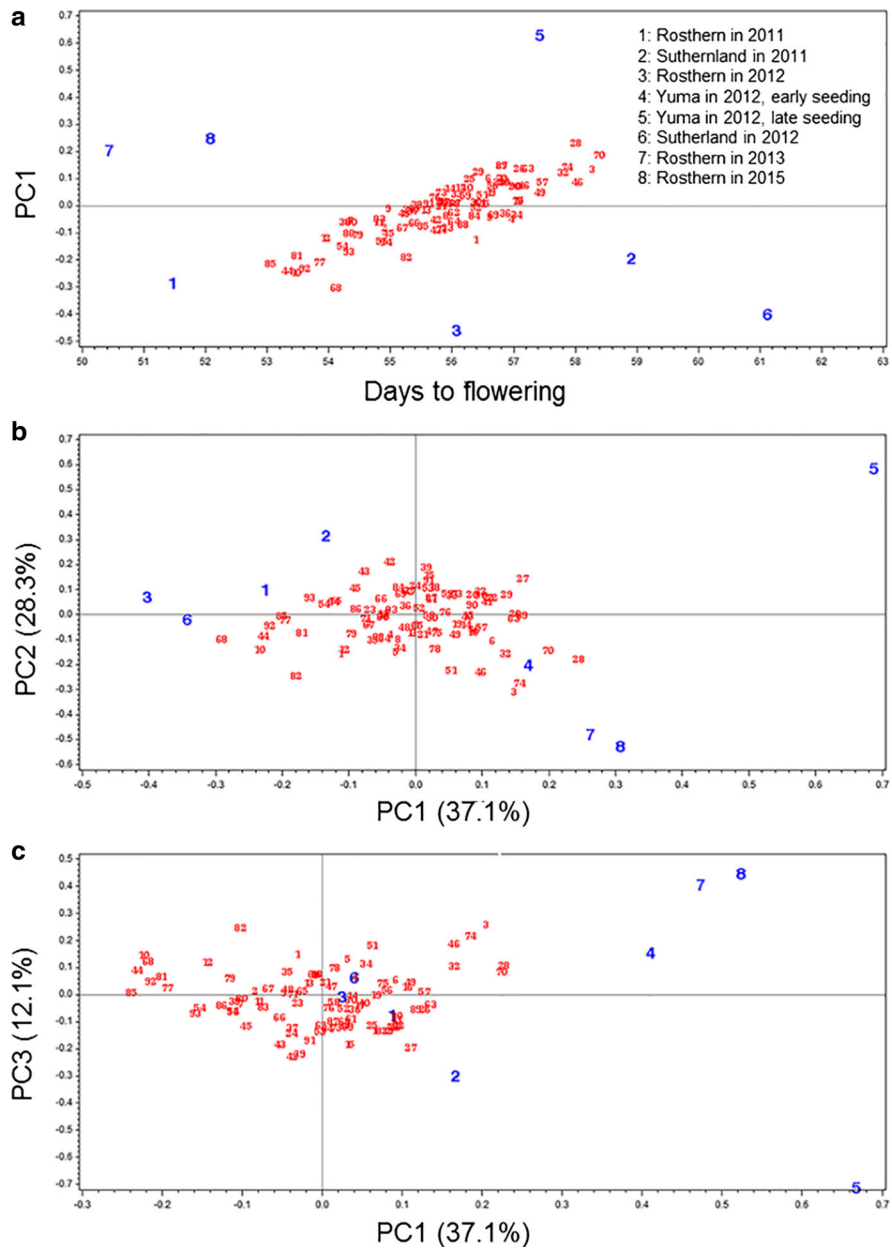

Fig. 2 Analysis of the genotype-by-environment interaction on days to flowering (DTF) using additive main effects and multiplicative interaction (AMMI) model in field pea in 8 site-years. **a** Biplot of AMMI analysis for Principal Component Analysis (PCA) component 1 versus average of DTF for 92 genotypes at 8 environments. **b** The second versus the first multiplicative component plot (PC2 vs. PC1). **c** The third versus the second multiplicative component plot (PC3 vs. PC2); Numbers in *blue* represent environments and numbers in red denote genotype code shown in Table 2

pollen germination reduction was only 0.10 (Table 4). The coefficient of variation for reduction in percentage of in vitro pollen germination (21.2–39.1%) was relatively large (Table 3), which was reflected in the large error variance (64.6%) for this trait in Table 4. No marker was associated with the reduction of in vitro pollen germination due to heat stress (incubation at 42 °C) for flower samples collected from both the field (Rosthern in 2015) and environmentally controlled conditions (Table 5).

Discussion

Both Bayesian and multivariate methods were employed to infer structure of the collection in the present study. Three groups were observed in this mapping panel using the Bayesian and DAPC methods, and these three groups roughly corresponded to the geographical origins of these cultivars. However, these clustering results differed to some extent from the 8 groups identified in a similar mapping panel

Table 5 Summary of significant ($-\log_{10}p \geq 4.3$) marker-trait associations identified by association analyses in a diversity panel of 92 pea cultivars

Trait	Location		Marker	Linkage group	Position (cM)	$-\log_{10}$ p	Marker r^2 [a]	Same marker for different traits
Days to flowering (DTF)	Rosthern-2011	1	PsC14322p98	Unmapped	22	4.5	0.20	–
	Rosthern-2012	2	PsC8005p478	Unmapped	80	4.3	0.20	–
	Rosthern-2013	3	PsC21767p87[b]	4	20	4.4	0.20	–
	Rosthern-2015	4	PsC7497p542	4	64	6.1	0.26	–
		5	PsC6902p242	Unmapped	73	5.3	0.20	–
	Growth chamber-2015	6	PsC20566p234	1	13	5.0	0.23	Rnode
		7	PsC22477p202	2	10	5.1	0.23	–
		8	PsC487p422	2	13	9.2	0.35	–
		9	PsC17431p188	2	32	5.3	0.24	Rnode
		10	PsC5013p645	2	70	4.3	0.21	–
		11	PsC5597p362	2	70	5.6	0.25	–
		12	PsC19344p128	3	17	4.6	0.23	Rnode
		13	PsC7220p181	3	91	8.3	0.35	–
		14	PsC27004p102	3	115	5.7	0.23	Rnode
		15	PsC15940p208	3	120		0.24	Rnode
		16	PsC3200p191	4	19	5.7	0.22	Rnode
		17	PsC21767p87[b]	4	20	5.1	0.24	Rnode
		18	PsC14392p100	4	59	4.5	0.21	–
		19	PsC908p622	7	46	4.5	0.21	–
		20	PsC10060p242	Unmapped	2	4.8	0.22	–
		21	PsC12051p325	Unmapped	12	8.7	0.36	
		22	PsC13289p266	Unmapped	18	5.5	0.26	Rnode
		23	PsC13955p450	Unmapped	21	9.3	0.38	Rnode
		24	PsC14155p82	Unmapped	21	6.3	0.28	Rnode
		25	PsC20385p319	Unmapped	42	5.7	0.22	Rnode
		26	PsC21134p263	Unmapped	43	5.7	0.22	Rnode
		27	PsC5231p296	Unmapped	66	4.6	0.21	–
Duration of flowering (DOF)	Sutherland-2012	28	PsC18479p162	2	6	5.5	0.23	–
		29	PsC19105p141	3	98	4.3	0.18	–
		30	PsC12831p152	3	124	6.6	0.26	–
		31	PsC17990p348	3	131	5.4	0.23	Rnode
		32	PsC4940p155	6	36	4.3	0.19	Rnode
		33	PsC6187p183	6	68	5.7	0.23	–
Number of pods	Yuma-2012, early seeding	34	PsC19517p115	Unmapped	40	5.2	0.24	–
	Sutherland-2012	35	PsC27644p242	Unmapped	55	6.5	0.23	–
		36	PsC6387p181	Unmapped	71	4.3	0.20	–
Pod set (%)	Rosthern-2012	37	PsC8016p73	3	16	4.8	0.21	–
Pods retained with seed (%) at Yuma	Yuma-2012, early seeding	38	PsC11375p247	1	87	4.6	0.20	–
		39	PsC7631p74	4	106	5.0	0.19	–
		40	PsC2509p330	5	83	4.5	0.26	–
		41	PsC1106p196	Unmapped	8	4.9	0.21	–

Table 5 continued

Trait	Location		Marker	Linkage group	Position (cM)	$-\log_{10} p$	Marker r^2 [a]	Same marker for different traits
Number of reproductive nodes (Rnode)	Rosthern-2012	42	PsC17990p348	3	131	4.3	0.20	DOF
		43	PsC4940p155	6	36	4.9	0.23	DOF
		44	PsC12883p342	Unmapped	16	6.4	0.25	–
	Rosthern-2015	45	PsC2491p801	1	1	5.0	0.24	–
		46	PsC20566p234	1	13	6.4	0.28	DTF
		47	PsC17776p66	2	15	5.2	0.24	–
		48	PsC17431p188	2	32	6.6	0.29	DTF
		49	PsC19344p128	3	17	6.4	0.30	DTF
		50	PsC27004p102	3	115	7.1	0.28	DTF
		51	PsC15940p208	3	120	6.1	0.28	DTF
		52	PsC3200p191	4	19	7.2	0.28	DTF
		53	PsC5316p234	5	75	8.2	0.34	–
		54	PsC13188p293	6	83	5.7	0.25	–
		55	PsC10016p165	Unmapped	2	5.1	0.23	–
		56	PsC13289p266	Unmapped	18	6.0	0.28	DTF
		57	PsC13955p450	Unmapped	21	7.2	0.31	DTF
		58	PsC14155p82	Unmapped	21	6.5	0.29	DTF
		59	PsC20385p319	Unmapped	42	7.2	0.28	DTF
		60	PsC21134p263	Unmapped	43	7.2	0.28	DTF

[a] r^2 (coefficient of determination) for the marker after fitting other model terms (population structure)

[b] Same significant marker for days to flowering at two different environments

where these same 92 pea cultivars plus two wild *Pisum* accessions [P651 (*Pisum fulvum*) and PI344538 (*Pisum sativum* subsp. *elatius*)] were analyzed using the Bayesian method, and these 94 cultivars were clustered into 8 groups (Diapari et al. 2015). The grouping assignments are consistent with Diapari et al. (2015) for the majority of individuals.

Molecular markers associated with important agronomic traits facilitate selection for interesting gene variants (Holland 2007). Comparing the QTL associated with flowering time (DTF and DOF) with the QTL associated with the number of reproductive nodes, 12 DTF QTL and 2 DOF QTL overlapped with the QTL associated with the number of reproductive nodes, i.e., flowering time QTL may play a crucial role in the number of reproductive nodes in field pea. Genes and environmental cues affecting the vegetative-to-reproductive transition can continue to affect post-flowering processes related to fertility and pod development (Weller and Ortega 2015). Flowering time is associated with fundamental decisions such as when and how to allocate resources. Weller and Ortega (2015) in a review of genetic control in flowering time in legume crops, indicated that flowering control is closely related to light perception/signaling, circadian clock, photoperiod response, signal integration and inflorescence development. The number of flowers/pods per node in pea is regulated by two genes, *Fn* and *Fna*. The single recessive mutations of *Fn* and *Fna* increase the number of flowers (reviewed in Benlloch et al. 2015). Flowering time genes such as *HR* and *SN* also affect flower/pod number in pea (Reid et al. 1996; Liew et al. 2014; Benlloch et al. 2015). Flower number is increased by dominant *HR* alleles and reduced by recessive *sn* alleles (Liew et al. 2014; Benlloch et al. 2015). The number of reproductive nodes is a function of the duration of meristem activity, which is also influenced by the flowering time genes *SN* and *HR* in field pea (Reid et al. 1996). Similar to the number of flowers per node, the number of reproductive nodes is increased by dominant *HR* alleles and decreased by

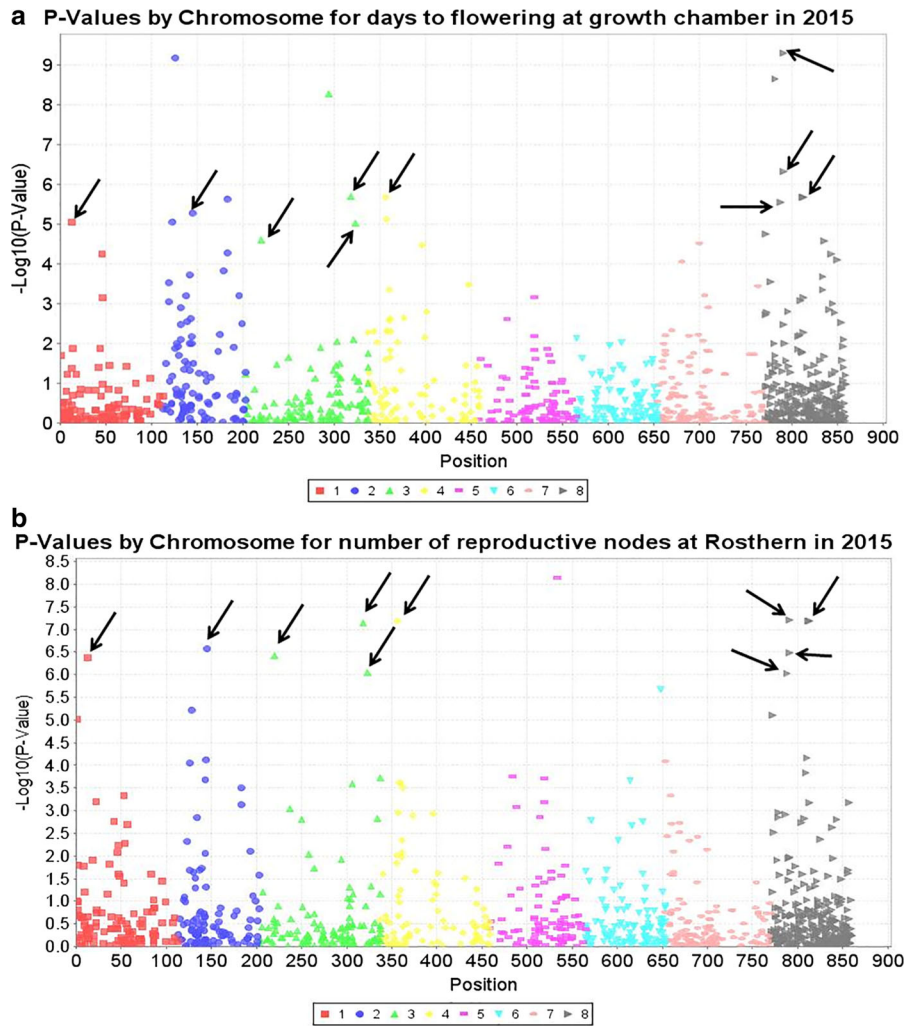

Fig. 3 Manhattan plots of $-\log_{10}$ (p value) of association mapping for days to flowering at growth chamber in 2015 and the number of reproductive nodes at Rosthern in 2015 in a diversity panel of 92 pea cultivars using the general linear model (GLM). The significant threshold was adjusted using the Bonferroni correction at $-\log_{10}(p < 0.05) = 4.3$. The *x*-axis represented seven linkage groups (LG I–LG VII) and unmapped markers are grouped in group 8. *Arrows* indicate the same significant markers in both traits as shown in Table 5

recessive *sn* alleles (Reid et al. 1996). To date, genes that specifically control the number of reproductive nodes have not been reported (Benlloch et al. 2015).

A total of 943 SNP markers from 1536 SNPs with known genetic positions from a consensus pea map (Sindhu et al. 2014) were used in the present study. An average SNP frequency of 1 SNP per 667 bp was detected (Sindhu et al. 2014). The synteny-based comparison between field pea and *Medicago truncatula* showed that the seven linkage groups, LG I, LG II, LG III, LG IV, LG V and LG VII, were collinear along their length with *Medicago* chromosomes 5, 1, 3, 8, 7 and 4, respectively; and the pea LG VI corresponded to the *Medicago* chromosomes 2 and 6 (Bordat et al. 2011; Sindhu et al. 2014). Flowering time QTL in *Medicago* are located at chromosomes 1, 4, 5, 7 and 8, which correspond to LG II, LG VII, LG I, LG V and LG IV of pea, respectively (Pierre et al. 2008). In the present study, significant markers related to DTF were identified on LG I, LG II, LG III, LG IV and LG VII. QTL associated with the onset of flowering were detected on LG II, LG IV and LG VI from three bi-parental populations in field pea (Ferrari et al. 2016). The flowering-time QTL on LG II was co-located with the QTL for grain yield in field pea, indicating a correlation between the onset of flowering and grain yield (Ferrari et al. 2016). However, due to a lack of common markers in the study conducted by Ferrari et al. (2016) and our study, a precise comparison of these locations is not possible. QTL associated with DTF were detected on LG II, III and VI in field pea (Huang et al. 2017). Significant markers associated

with DOF were detected on LG III in the present study and the study conducted by Huang et al. (2017). Significant markers associated with the number of reproductive nodes were observed on LG III and VI in the present study and Huang et al. (2017). QTL are often not consistently detected across different mapping populations (Holland 2007), partly due to genetic heterogeneity and different QTL that segregate in different mapping populations (Holland 2007; Myles et al. 2009).

The broad sense heritability of DTF, DOF, the number of reproductive nodes, and the number of pods were 0.67, 0.78, 0.80 and 0.51 in field pea in the present study. The broad sense heritability of DTF was documented to be 0.69, 0.77 and 0.87 in three bi-parental populations in field pea (Ferrari et al. 2016). Similarly, the broad sense heritability of DTF, DOF, the number of reproductive nodes, and the number of pods were 0.90, 0.49, 0.38 and 0.57, respectively in field pea in a bi-parental population across five site-years at Rosthern in 2012, 2013 and 2014, and Sutherland in 2013 and 2014 (Huang et al. 2017). Relatively high heritability in the traits related to reproductive development indicates that the phenotype is strongly linked to the genotype.

DTF was reduced for plants seeded late at the Arizona location in 2012, likely due to day length being longer and temperatures warmer for plants with the late seeding date. The regulation of flowering time via different seeding dates is of great interest since plants are able to avoid the most stressful phases during the growing season with appropriate seeding dates. The range of DTF under controlled conditions was greater in field conditions due to several late flowering cultivars being sensitive to light spectral quality in growth chambers. Artificial light in growth chambers usually has a greater red (660 nm) to far-red (735 nm) ratio (R:FR) compared to natural light with the R:FR ratio of $\sim$1.2 (Holmes and Smith 1975). For example, lentil plants flowered earlier when exposed to a smaller R:FR light ratio (Mobini et al. 2016). Results from the AMMI model demonstrated DTF was unstable in certain cultivars, whereas others had a greater stability in DTF, which suggested that stable flowering cultivars may be less photoperiod-sensitive. Late flowering genotypes are generally more sensitive to photoperiod and temperature compared to early flowering genotypes in field pea (Truong and Duthion 1993).

The genetic variation in an association mapping panel is generally much greater compared to linkage populations, and thus sample size for association mapping remains smaller than linkage populations (Zhu et al. 2008). Approximately 100 genotypes were investigated in many association mapping studies reviewed by Zhu et al. (2008). Population size and experimental design are two key factors affecting QTL detection using genome-wide association mapping (Wang et al. 2012). For example, population sizes greater than 384 individuals are needed to consistently detect the three major heading date QTL in barley (Wang et al. 2012). The Bonferroni adjustment was employed in the present study, and this multiple testing approach may overcorrect for the inflated false-positive rate, leading to a reduction in significant SNP markers detected compared to an approach without the multiple testing correction (Nyholt 2004). Some studies did not use the multiple testing corrections. For example, the p value of 0.01 (equivalent to the value of $-\log_{10}p$ of 2) was used to determine if a QTL was associated with a marker (Pillen et al. 2003; Agrama et al. 2007; Brazauskas et al. 2011; Nemli et al. 2014).

Additionally, the kinship analysis on a similar population conducted by Diapari et al. (2015) suggested that family relatedness among the majority of the cultivars (97.2% of 94 cultivars—the same 92 cultivars in the present plus two wild *Pisum* accessions) is loose, therefore the results from the GLM model are less conservative compared to the results from the mixed linear model (MLM) model in the present study. Significantly associated markers are more reliable if the same markers are detected in two or more separate environments compared to significant markers only detected at a single environment. Therefore, significant markers associated with the number of reproductive nodes, the number of pods on the main stem, and percentage of pods set and percentage of pods retained with seed are less convincing compared to the marker associated with DTF in two or more separate environments. Thus datasets of these traits (the number of reproductive nodes and the number of pods on the main stem) from more environments would be recommended to validate markers detected in only a single environment. Further, more precise phenotyping with lower values of coefficient of variation could have resulted in

greater values of coefficient of determination, thus improving the accuracy of the marker associations.

In vitro pollen germination has been used as a rapid screening criterion for stress tolerance in many legume crops such as chickpea (Devasirvatham et al. 2012), soybean (Salem et al. 2007), field pea (Petkova et al. 2009; Lahlali et al. 2014; Jiang et al. 2015), and groundnut (Kakani et al. 2002). QTL mapping of thermos-tolerance traits has been conducted, such as in vitro pollen germination in maize (Frova and Sari-Gorla 1994) and pollen viability in rice (Xiao et al. 2011). However, the present study failed to yield any significant markers associated with in vitro pollen germination, because phenotyping of this trait was both technically challenging and time-sensitive for a large population. Our results could be confounded by time of sampling and small differences in flower and anther development stages. These problems could first be alleviated with a repeated check cultivar in the field to standardize and calibrate the pollen germination data to improve the accuracy of phenotyping. It took approximately 4–5 h to run the in vitro pollen germination test on 92 cultivars with two temperature treatments (one replication). Pollen viability and vigor may be reduced between the first few cultivars compared to the last few cultivars.

A second improvement would be accurate flower sampling so anthers of all cultivar samples are at the exact same stage to achieve successful mapping of pollen vigor for a large population. Flower samples were collected by more than one person, which introduced additional sampling variation. The optimal stage of flower sampling was at Stage III (Jiang 2016; after anther dehiscence and before flowers were fully open), which can be easily confused with fully open flowers but with the standard petal closed again [Stage IV, V (Jiang 2016)] and pollen viability is now lower.

Although no significant markers associated with the heat-induced reduction of pollen germination in this study, the greater number of reproductive nodes (indeterminacy) could be an easy trait to explore and use for lengthening flowering duration. Previous research has shown indeterminacy is beneficial for recovering from environmental stress in chickpea and cowpea (Berger et al. 2006; Hall 2004). Heat and drought stress accelerate reproductive growth and reduce yield in soybean, so lengthening reproductive growth could counteract the negative effects of stress (Desclaux and Roumet 1996). Therefore, the number of reproductive nodes, DTF and DOF are recommended as criteria by which to screen heat tolerance. In Western Canada, short heat waves are the main manifestation of heat stress, so having more sequentially formed flowers and pods would be the primary strategy for heat resistant yield in future cultivars.

Percentage of pods set in Saskatchewan (Rosthern and Sutherland) was greater than losses associated with the death of developing pods (percentage of pods retained with seed) at Yuma. For the Northern Great Plains where reproductive growth proceeds at similar temperature ranges as early reproductive growth, pod set percentage is useful, indicating superior genetic performance where flower abortion is the main impediment to yield loss in stress. Including a second pod assessment such as the degree of pod abortion would be useful for regions that experience prolonged and significant heat stress in reproductive growth (e.g. Australia, late seeding in Arizona). In these regions, spring-sown pea flowers in cool temperatures and sets pods, but the pods develop and mature in warmer temperatures and drought, and stress will also cause some of these pods to die prematurely.

Conclusions

Ninety-two diverse pea cultivars clustered into three subpopulations roughly corresponding to the geographic origin of the individuals. Sixty SNPs showed marker-trait associations with seven reproductive development related traits. Fourteen flowering time QTL overlapped with the QTL associated with the number of reproductive nodes in field pea. One marker located at LG IV was significantly associated with DTF at two separate environments. Several markers were found to be significantly associated with other reproductive development related traits including the number of reproductive nodes, the number of pods, and percentage of pods set. Collectively, our findings unveiled the genetic basis of reproductive development related traits in pea, serving as an avenue for genetic improvement through marker-assisted breeding.

Acknowledgements This work was supported by Saskatchewan Agriculture Development Fund [Grant Number 20100033], Saskatchewan Pulse Growers Association [Grant Number AGR1116], Western Grains Research Fund [Grant Number 411939], and the Natural Sciences and Engineering

Research Council (NSERC) of Canada—Collaborative Research and Development (CRD) Program [Grant Number CRDPJ 439277]. We acknowledge the kind help from the staff at the Pulse Crop Breeding Program of the Crop Science Field Laboratory and the staff at the Crop Physiology Lab at University of Saskatchewan. We thank Amidou N'Diaye and Kirstin Bett for helpful discussions. We are grateful for the helpful comments from two anonymous referees that improved this manuscript.

References

Agrama H, Eizenga A, Yan G (2007) Association mapping of yield and its components in rice cultivars. Mol Breed 19:341–356

Ahmad S, Kaur S, Lamb-Palmer ND, Lefsrud M, Singh J (2015) Genetic diversity and population structure of *Pisum sativum* accessions for marker-trait association of lipid content. Crop J 3:238–245

Ahmed FE, Hall AE, DeMason DA (1992) Heat injury during floral development in cowpea (*Vigna unguiculata*, Fabaceae). Am J Bot 79:784–791

Alcade JA, Wheeler TR, Summerfield RJ, Norero AL (1999) Quantitative effects of the genes *Lf*, *Sn*, *E*, and *Hr* on time to flowering in pea (*Pisum sativum* L.). J Exp Bot 50:1691–1700

Alcalde JA, Larrain MF (2006) Timing of photoperiod sensitivity in relation to floral initiation in contrasting genotypes of pea (*Pisum sativum* L.). Field Crops Res 96:348–354

Annicchiarico P, Nazzicari N, Pecetti L, Romani M, Ferrari B, Wei Y, Brummer EC (2017) GBS-based genomic selection for pea grain yield under severe terminal drought. Plant Genome 10:1–13

Benlloch R, Berbel A, Ali L, Gohari G, Millán T, Madueño F (2015) Genetic control of inflorescence architecture in legumes. Front Plant Sci 6:543

Berger JD, Ali M, Basu PS, Chaudhary BD, Chaturvedi SK, Deshmukh PS, Yadav SS (2006) Genotype by environment studies demonstrate the critical role of phenology in adaptation of chickpea (*Cicer arietinum* L.) to high and low yielding environments of India. Field Crop Res 98:230–244

Bordat A, Savois V, Nicolas M, Salse J, Chauveau A, Bourgeois M, Potier J, Houtin H, Rond C, Murat F, Marget P, Aubert G, Burstin J (2011) Translational genomics in legumes allowed placing in Silico 5460 unigenes on the pea functional map and identified candidate genes in *Pisum sativum* L. G3 1:93–103

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 23:2633–2635

Brazauskas G, Lenk I, Pedersen MG, Studer B, Lübberstedt T (2011) Genetic variation, population structure and linkage disequilibrium in European elite germplasm of perennial ryegrass. Plant Sci 181:412–420

Bueckert RA, Wagenhoffer S, Hnatowich G, Warkentin TD (2015) Effect of heat and precipitation on pea yield and reproductive performance in the field. Can J Plant Sci 95:629–639

Cheng P, Holdsworth W, Ma Y, Coyne CJ, Mazourek M, Grusak MA, Fuchs S, McGee RJ (2015) Association mapping of agronomic and quality traits in USDA pea single-plant collection. Mol Breed 35:75

Desclaux D, Roumet P (1996) Impact of drought stress on the phenology of two soybean (*Glycine max* L. Merr) cultivars. Field Crop Res 46:61–70

Desgroux A, L'Anthoëne V, Roux-Duparque M, Rivière J, Aubert G, Tayeh N, Moussart A, Mangin P, Vetel P, Piriou C, McGee RJ, Coyne CJ, Burstin J, Baranger A, Manzanares-Dauleux M, Bourion V, Pilet-Nayel M (2016) Genome-wide association mapping of partial resistance to *Aphanomyces euteiches* in pea. BMC Genom 17:124

Devasirvatham V, Gaur PM, Mallikarjuna N, Tokachichu RN, Trethowan RM, Tan DKY (2012) Effect of high temperature on the reproductive development of chickpea genotypes under controlled environments. Funct Plant Biol 39:1009–1018

Diapari M, Sindhu A, Bett K, Deokar A, Warkentin TD, Tar'an B (2014) Genetic diversity and association mapping of iron and zinc concentrations in chickpea (*Cicer arietinum* L.). Genome 57:459–468

Diapari M, Sindhu A, Warkentin TD, Bett K, Tar'an B (2015) Population structure and marker-trait association studies of iron, zinc and selenium concentrations in seed of field pea (*Pisum sativum* L.). Mol Breed 35:30

Dolezel J, Greilhuber J (2010) Nuclear genome size: are we getting closer? Cytometry Part A 77:635–642

Earl DA, VonHoldt BM (2012) Structure Harvester: a website and program for visualizing structure output and implementing the evanno method. Conserv Genet Resour 4:359–361

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) Taken from http://faostat3.fao.org/browse/Q/QC/E 8 Oct 2016

Ferrari B, Romani M, Aubert G, Boucherot K, Burstin J, Pecetti L, Huart-Naudet M, Klein A, Annicchiarico P (2016) Association of SNP markers with agronomic and quality traits of field pea in Italy. Czech J Genet Plant Breed 52:83–93

Frova C, Sari-Gorla M (1994) Quantitative trait loci (QTLs) for pollen thermotolerance detected in maize. Mol Gen Genet 245:424–430

Hall AE (2004) Breeding for adaptation to drought and heat in cowpea. Eur J Agron 21:447–454

Hoagland DR, Arnon DI (1938) The water-culture method for growing plants without soil. University of California, Berkeley, pp 35–39

Holland JB (2007) Genetic architecture of complex traits in plants. Curr Opin Plant Biol 10:156–161

Holmes MG, Smith H (1975) Function of phytochrome in plants growing in natural-environment. Nature 254:512–514

Huang S, Gali KK, Tar'an B, Warkentin TD, Bueckert RA (2017) Pea phenology: crop potential in a warming environment. Crop Sci 57:1540–1551

Jiang Y (2016) Effect of heat stress on pollen development and seed set in field pea. University of Saskatchewan, Saskatoon

Jiang Y, Lahlali R, Karunakaran C, Kumar S, Davis AR, Bueckert RA (2015) Seed set, pollen morphology and

pollen surface composition response to heat stress in field pea. Plant Cell Environ 38:2387–2397

Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94

Kakani VG, Prasad PVV, Craufurd PQ, Wheeler TR (2002) Response of in vitro pollen germination and pollen tube growth of groundnut (*Arachis hypogaea* L.) genotypes to temperature. Plant Cell Environ 25:1651–1661

Kang MS, Balzarini MG, Guerra JLL (2004) Genotype-by-environment interaction. In: Saxton AM (ed) Genetic analysis of complex traits using SAS. SAS Institute, Cary

Konsens I, Ofir M, Kigel J (1991) The effect of temperature on the production and abscission of flowers and pods in snap bean (*Phaseolus vulgaris* L.). Ann Bot 67:391–399

Krajewski P, Bocianowski J, Gawłowska M, Kaczmarek Z, Pniewski T, Święcicki W, Wolko B (2012) QTL for yield components and protein content: a multienvironment study of two pea (*Pisum sativum* L.) populations. Euphytica 183:323–336

Kwon S, Brown AF, Hu J, McGee R, Watt C, Kisha T, Timmerman-Vaughan G, Grusak M, McPhee KE, Coyne CJ (2012) Genetic diversity, population structure and genome-wide marker-trait association analysis emphasizing seed nutrients of the USDA pea (*Pisum sativum* L.) core collection. Genes Genom 34:305–320

Lahlali R, Jiang Y, Kumar S, Karunakaran C, Liu X, Borondics F, Hallin E, Bueckert R (2014) ATR-FTIR spectroscopy reveals involvement of lipids and proteins of intact pea pollen grains to heat stress tolerance. Front Plant Sci 5:747

Li Y, Smulders M, Chang J, Qiu M (2011) Genetic diversity and association mapping in a collection of selected Chinese soybean accessions based on SSR marker analysis. Conserv Genet 12:1145–1157

Liew LC, Hecht V, Sussmilch FC, Weller JL (2014) The pea photoperiod response gene *STERILE NODES* is an ortholog of *LUX ARRHYTHMO*. Plant Physiol 165:648–657

Marjanović-Jeromela A, Nagl N, Gvozdanović-Varga J, Hristov N, Kondić-Špika A, Vasić M, Marinković R (2011) Genotype by environment interaction for seed yield per plant in rapeseed using AMMI model. Pesq agropec bras 46:174–181

Mobini SH, Lulsdorf M, Warkentin T, Vandenberg A (2016) Low red:far-red light ratio causes faster in vitro flowering in lentil. Can J Plant Sci 96:908–918

Morrell P, Buckler E, Ross-Ibarra J (2012) Crop genomics: advances and applications. Nature Rev Genet 13:85–96

Murfet IC (1985) *Pisum sativum*. In: Halevy AH (ed) CRC handbook of flowering, vol IV. CRC Press. Boca Raton, FL, pp 97–126

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES (2009) Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21:2194–2202

Nemli S, Asciogul TK, Kaya HB, Kahraman A, Eşiyok D, Tanyolac B (2014) Association mapping for five agronomic traits in the common bean (*Phaseolus vulgaris* L.). J Sci Food Agric 94:3141–3151

Petkova V, Nikolova V, Kalapchieva SH, Stoeva V, Topalova E, Angelova S (2009) Physiological response and pollen viability of *Pisum sativum* genotypes under high temperature influence. Acta Hortic 830:665–671

Pierre J, Huguet T, Barre P, Huyghe C, Julier B (2008) Detection of QTLs for flowering date in three mapping populations of the model legume species *Medicago truncatula*. Theor Appl Genet 117:609–620

Pillen K, Zacharias A, Léon J (2003) Advanced backcross QTL analysis in barley (*Hordeum vulgare* L.). Theor Appl Genet 107:340–352

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

Reid JB, Murfet IC, Singer SR, Weller JL, Taylor SA (1996) Physiological-genetics of flowering in *Pisum*. Semin Cell Dev Biol 7:455–463

Roche R, Jeuffroy M, Ney B (1998) A model to simulate the final number of reproductive nodes in pea (*Pisum sativum* L.). Ann Bot 81:545–555

Sabaghnia N, Sabaghpour SH, Dehghani H (2008) The use of an AMMI model and its parameters to analyse yield stability in multi-environment trials. J Agric Sci 146:571–581

Sakiroglu M, Sherman-Broyles S, Story A, Moore K, Doyle JJ, Charles BE (2012) Patterns of linkage disequilibrium and association mapping in diploid alfalfa (*M. sativa* L.). Theor Appl Genet 125:577–590

Salem MA, Kakani VG, Koti S, Reddy KR (2007) Pollen-based screening of soybean genotypes for high temperatures. Crop Sci 47:219–231

Salvi S, Tuberosa R (2005) To clone or not to clone plant QTLs: present and future challenges. Trends Plant Sci 10:297–304

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

Sindhu A, Ramsay L, Sanderson LA, Stonehouse R, Li R, Condie J, Shunmugam ASK, Liu Y, Jha AB, Diapari M, Burstin J, Aubert G, Tar'an B, Bett KE, Warkentin TD, Sharpe AG (2014) Gene-based SNP discovery and genetic mapping in pea. Theor Appl Genet 127:2225–2241

Tar'an B, Warkentin T, Somers DJ, Miranda D, Vandenberg A, Blade S, Bing D (2004) Identification of quantitative trait loci for grain yield, seed protein concentration and maturity in field pea (*Pisum sativum* L.). Euphytica 136:297–306

Thudi M, Upadhyaya HD, Rathore A, Gaur PM, Krishnamurthy L, Roorkiwal M, Nayak SN, Chaturvedi SK, Basu PS, Gangarao N, Fikre A, Kimurto P, Sharma PC, Sheshashavee 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

Timmerman-Vaughan GM, Mills A, Whitfield C, Frew T, Butler R, Murray S, Lakeman M, McCallum J, Russell A, Wilson D (2005) Linkage mapping of QTL for seed yield, yield components, and developmental traits in pea. Crop Sci 45:1336–1344

Truong H, Duthion C (1993) Time of flowering of pea (*Pisum sativum* L.) as a function of leaf appearance rate and node of first flower. Ann Bot 72:133–142

Vanhala T, Normann KR, Lundstrom M, Weller JL, Leino MW, Hagenblad J (2016) Flowering time adaption in Swedish landrace pea (*Pisum sativum* L.). BMC Genet 17:117

Varshney RK, Ribaut J, Buckler ES, Tuberosa R, Rafalski JA, Langridge P (2012) Can genomics boost productivity of orphan crops? Nat Biotechnol 30:1172–1176

Wang H, Smith KP, Combs E, Blake T, Horsley RD, Muehlbauer GJ (2012) Effect of population size and unbalanced data sets on QTL detection using genome-wide association mapping in barley breeding germplasm. Theor Appl Genet 124:111–124

Wang N, Chen B, Xu K, Gao G, Li F, Qiao J, Yan G, Li J, Li H, Wu X (2016) Association mapping of flowering time QTLs and insight into their contributions to rapeseed growth habits. Front Plant Sci 7:338

Weller J, Ortega R (2015) Genetic control of flowering time in legumes. Front Plant Sci 6:207

Weller JL, Reid JB, Taylor SA, Murfet IC (1997) The genetic control of flowering in pea. Trends Plant Sci 2:412–418

Weller JL, Hecht V, Liew LC, Sussmilch FC, Wenden B, Knowles CL, Vander SJK (2009) Update on the genetic control of flowering in garden pea. J Exp Bot 60:2493–2499

Xiao YH, Pan Y, Luo LH, Deng HB, Zhang GL, Tang WB, Chen LY (2011) Quantitative trait loci associated with pollen fertility under high temperature stress at flowering stage in rice (*Oryza sativa*). Rice Sci 18:204–209

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539–551

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20