Research Article **Open Access**

SNP Identification from RAD-Seq Data in Faba Bean (*Vicia faba* L.)

Zhu Yujie [1], Liu Tingfu [2], Li Hanmei [3], Wang Baogen [4], Wu Suqi [4], Li Sujuan [5], Wang Linlin [2], Li Guojing [4], Wu Xinyi [4] ✉

1 School of Agriculture and Food Science, Zhejiang A&F university, Hangzhou, 311300, China
2 Lishui Academy of Agricultural and Forestry Sciences, Lishui 323000, China
3 Lishui Vocational and Technical College, Lishui, 323000, China
4 Institute of Vegetables, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China
5 Central Laboratory of Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China
✉ Corresponding author email: wuxinyi@zaas.ac.cn

Abstract Faba bean is an important edible legume crop in China. Due to its huge genome size and no available reference genome, SNP marker is very limited in faba bean. To identify genome-wide SNP markers, we obtained 35.47 Gb data from eight landraces by RAD-sequencing, with an average of 4.77 Gb data for each accession. A total of 245443516 reads were generated, the single accession has an average of 30680439.5 reads, and the average length of the reads reaches to 144 bp. The Q20 and Q30 values were over 97.89% and 93.83%, respectively. The GC content between the reads varied from 38.05% to 40.09%. Using a special Bayesian method under the situation of no reference genome, we identified 3722 SNPs among the eight landraces. Regarding the single accession, the detected SNPs varied from 3278 to 3578, and the homozygous SNPs number was larger than that of heterozygous SNPs for most of accessions. For the SNP types, T:A->C:G type has the largest proportion (38.8%), followed by C:G->T:A (28.0%) and the smallest is T:A->A:T (7.5%). 31 SNPs were selected to convert into KASP markers, and they showed a success rate of 66.7% through amplifying on 46 accessions. The SNPs in this study provide a strong genetic tool for germplasm identification, gene mapping and molecular breeding in faba bean.

Keywords Faba bean; SNP; RAD-Sequencing; KASP

Faba bean (*Vicia faba* L.), belonging to *Leguminosae Papilionidae Vicia* genus, is an important edible legume crop in temperate and subtropical regions. Faba bean has been cultivated in China for over 2100 years (Wang et al., 2011; Yang et al., 2016). Faba bean is rich in protein and cellulose, and could be digested and absorbed easily, its dry seeds can be used as food, feed or processing into leisure food, while its fresh grains can be used as vegetables. Faba bean also plays an important role in crop rotation and nourishing system in the cropping structure adjustment due to the biological nitrogen fixation from the roots. In recent years, the annual planting area of faba bean in China is about 925,000 hm² and the annual yield is about 1,595,000 tons, ranking first both on production area and total yield in the world (http://faostat3.fao.org/).

Faba bean is a diploid crop (2n=2x=12), with a genome of about 13 Gb (Johnston et al., 1999). As one of the largest legume genomes, faba bean genome size is approximately 25 times larger than that of the model legume *Medicago truncatula*. Due to the larger genome, the genome research such as whole genome sequencing and marker development has been hindered seriously, thus causing limited genetic gain through the use of molecular markers. Recently, with the rapid development of next-generation sequencing technology and the significant decrease of sequencing cost, high-throughput sequencing has been widely used in the complex and larger genome crops such as wheat for marker development and gene mapping (Poland et al., 2012; Wang et al., 2014). In faba bean, Yang et al. (2012) identified 125,559 SSR sequences and developed 28503 SSR markers by sequencing the mixed genome of 247 germplasms using the 454 sequencing technology; Kaur et al. (2012) identified 304,680 specific genes and developed 802 SSR markers through transcriptome sequencing of two faba bean varieties using the 454 sequencing technology; Webb et al. (2016) detected 845 SNPs from transcriptome sequencing data of two faba inbred lines and constructed a genetic map containing 687 SNP markers. These markers resources laid a solid

foundation for faba bean genetic breeding. Single nucleotide polymorphism (SNP), the DNA sequence polymorphism caused by a nucleotide mutation on the genome, has become an ideal next generation molecular marker due to its advantages including wide distribution on the genome, high density, high stability and high-throughput. However, the SNP marker resource is relatively limited in faba bean until now excepting for the 845 SNPs reported in Vignal et al. (2002).

Reduced representation sequencing such as restriction site-associated DNA sequencing (RAD-Seq) is a next-generation sequencing technique that reduces genome complexity by digestion of genomic DNA by restriction enzymes, followed by the sequencing of fragments within a given size range, and thus to obtain partial sequences representing the whole genome information (Baird et al., 2008; Rowe et al., 2011). Due to its moderate sequencing depth, low cost and independent of reference genome, RAD-Seq has been widely used in marker development, genetic map construction and target gene mapping in multiple non-model species (Yang et al., 2013; Hegarty et al., 2013; Xu et al., 2014). In the current study, RAD-seq was used to sequence eight faba bean landraces to identify genome wide SNPs based on a special SNP detection technology under lack of reference genome (Xu et al., 2014), and also analyze faba bean SNPs characteristics. The obtained SNPs markers will provide a new resource for future gene mapping, genetic map construction and molecular marker-assisted breeding.

1 Results and Analysis

1.1 Sequencing data statistics

In the current study, eight faba bean accessions were sequenced using the Illumina Hiseq platform and a total of 35.47 Gb data was obtained. Total 245443516 high quality reads were generated, and the average length of each reads was 144 bp. For each accession, the smallest sequencing data was 3.83 Gb, the largest was 4.77 Gb, and the average sequencing data was 4.43 Gb; the read number ranged from 26415662 to 32822210, with the average of 30680439.5; Q20 and Q30 were greater than 97.89% and 93.83%, respectively, and the GC content varied from 38.05% to 40.09% (Table 1).

Table 1 Summary of sequencing data on 8 selected faba bean accessions

Genotypes	Total Reads	Total bases	Q20%	Q30%	GC%
FB017	30206500	4351379886	97.89	93.83	40.09
FB032	31125802	4488216216	97.93	93.96	38.76
FB036	31023152	4467907202	97.99	94.11	39.03
FB056	30804212	4441470823	97.97	94.05	39.71
FB076	31573060	4550456243	97.96	94.05	38.59
FB079	32822210	4767870400	98.02	94.17	39.15
FB080	26415662	3834851207	97.95	94.02	38.05
FB081	31472918	4572159352	98.07	94.29	39.42

1.2 Genome wide SNPs identification in faba bean

3722 SNPs were identified according to Xu et al. (2014) where a special Bayesian method was used to detect bottle gourd SNPs without reference genome. For single accession, the least 3278 SNPs were identified in FB076 while the most 3578 SNPs were identified in FB079. Among the eight accessions, the homozygous SNPs number ranged from 1579 to 2033 and the heterozygous SNPs number varied from 1245 to 1804, the number of homozygous SNPs was larger than that of the heterozygous SNPs in most accessions excluded FB080 and FB056 (Table 2). Among the 6 SNP mutational types, T:A->C:G mutational types accounted for the largest proportion (mean 38.8%), followed by C:G->T:A (mean 28.0%), and the smallest proportion was T:A->A:T (mean 7.50%) (Table 3).

Table 2 Summary of SNP identified among the 8 selected faba bean accessions

Genotypes	Total mutations	Homozygous SNP	Heterozygous SNP
FB017	3352	1798	1554
FB032	3475	2005	1470
FB036	3477	1881	1596
FB056	3575	1771	1804
FB076	3278	2033	1245
FB079	3578	1940	1638
FB080	3369	1579	1790
FB081	3533	1858	1675

Table 3 Summary of SNP mutation types identified in the current study

Genotypes	C:G->A:T	C:G->G:C	C:G->T:A	T:A->A:T	T:A->C:G	T:A->G:C
FB017	299	125	955	250	1335	452
FB032	293	124	974	258	1368	476
FB036	281	131	970	250	1377	460
FB056	316	130	1018	271	1380	476
FB076	276	102	908	254	1301	458
FB079	317	121	1002	277	1395	482
FB080	311	118	946	264	1254	449
FB081	315	132	993	259	1369	486

1.3 SNPs validation

To verify the effectiveness of SNPs detected in this study, 56 SNPs were selected to develop KASP markers after filtering the SNPs with missing rate ≤20%, MAF value ≥0.05 and SNP occurrence frequency ≥40. Finally, 31 SNPS were converted into KASP markers (Table 4), with a conversion success rate of 55.3%. These 31 KASP markers were used to genotype 46 faba bean germplasms, and the results showed that 22 markers detected successful signal, of which 14 markers showed no polymorphism, 4 markers revealed two genotypes and other 4 markers displayed three genotypes (Figure 1).

2 Discussion

With the cost reducing rapidly, genome resequencing technique has been widely used to identify genome-wide SNPs in many crops. However, due to larger genome size and lack of reference genome, the SNP mining research in faba bean lags behind other legume crops such as soybean, common bean and cowpea (Wu et al., 2010; Song et al., 2015; Muňoz-Amatriaín et al., 2017). Ocana et al. (2015) and Webb et al. (2016) have identified 39060 SNPs and 845 SNPs by transcriptome sequencing of different inbred lines, respectively. Webb et al. (2016) also converted the SNPs into KASP markers and constructed the first genetic map based on SNPs markers in faba bean. In the current work, eight faba bean germplasm were sequenced using the RAD-Seq technique, and each one got about 4.77Gb data, which covering about 36.7% of the whole genome, indicating a low-depth genome coverage. To improve the accuracy of SNP identification, a special Bayesian method developed by Xu et al. (2014) was used to detect SNPs and finally a total of 3722 SNP markers were identified. The number of SNPs detected in different accessions varied highly, and the maximum difference between two accessions reaches to 300 SNPs. Further analysis showed that there was very weak correlation between the sequence data and the identified SNPs in a single accession, for example, FB076 with a larger sequencing data volume identified the least SNPS (Table 1 and Table 2). Among the 6 SNP mutation types, T:A->C:G mutation types accounted for the largest proportion (mean 38.8%), followed by C:G->T:A (mean 28.0%), and the smallest proportion was T:A->A:T(mean 7.50%). However, the study in Ocana et al. (2015) showed that C/T type SNPs accounted for the largest proportion (18.7%), following by A/G (16.5%), T/C (15.3%) and G/A (14.2%), while other SNPs accounted for only 5%, indicating that the SNPs distribution on the genome is not evenly and different SNPs types varied highly in different cultivars. Different from Ocana et al. (2015) and Webb et al. (2016) where the SNPs detecting from transcriptome sequencing, the SNPs in the current study derived from genome sequence, thus could identify the SNPs both in inter-genes and intra-genes regions.

Table 4 The 31 KASP markers developed in the current study

ID	Primer AlleleX	Primer AlleleY	Primer Common
FB1 002	GAAGGTGACCAAGTTCATGCTTGACTT CAAAGACCTGGGTTCAAG	GAAGGTCGGAGTCAACGGATTGTGAC TTCAAAGACCTGGGTTCAAA	CCTGTATGGAAAAATAA GATTGGAATTTTGCAA
FB1 003	GAAGGTGACCAAGTTCATGCTGTCTCT CATACCAAAGGCACGG	GAAGGTCGGAGTCAACGGATTAAGTC TCTCATACCAAAGGCACGA	TTTTGTTTGGCCGTTTCA ATGAGCTCTA
FB1 005	GAAGGTGACCAAGTTCATGCTTAATC CACTAATGCTTTGTTTTTCTTTTATATC	GAAGGTCGGAGTCAACGGATTTAATC CACTAATGCTTTGTTTTTCTTTTATATT	GAGTGCGCATTCTGAATA TATAATAATAATCAAAT
FB1 008	GAAGGTGACCAAGTTCATGCTAATGG GTACTTACCGCTATTTATAGTG	GAAGGTCGGAGTCAACGGATTCAATG GGTACTTACCGCTATTTATAGTA	GCATTAATGTCCTCGAA TTTGGGGATTT
FB1 009	GAAGGTGACCAAGTTCATGCTATCCC AGGTATTGTTCTTCAGGAAG	GAAGGTCGGAGTCAACGGATTATATC CCAGGTATTGTTCTTCAGGAAA	TTTGAGCTGTTGTCATCA CAAGTTGGAA
FB1 010	GAAGGTGACCAAGTTCATGCTCATAC ATATGCTGAATGAAACCCTGAAA	GAAGGTCGGAGTCAACGGATTATACA TATGCTGAATGAAACCCTGAAC	GTGTTTTTGTGGCAGCTT TTTCGTGTAA
FB1 011	GAAGGTGACCAAGTTCATGCTCGAAT GGATTGCAAATATAGTGCCAG	GAAGGTCGGAGTCAACGGATTTCGAA TGGATTGCAAATATAGTGCCAA	ATGCATACTCTAAGAGAT CCATTTTTCTTAATTA
FB1 015	GAAGGTGACCAAGTTCATGCTACTGC CTTCACAAATAAACTTCAAAGTTTT	GAAGGTCGGAGTCAACGGATTCTGCCT TCACAAATAAACTTCAAAGTTTC	CCTGGAATGACGTTTTGG TGATGGTT
FB1 016	GAAGGTGACCAAGTTCATGCTAGGAT ACCAGTTTGATACCAAGTGTAA	GAAGGTCGGAGTCAACGGATTAGGAT ACCAGTTTGATACCAAGTGTAT	AAATTCTTGTCTTTCTTTT ACCCTCCTCTTT
FB1 017	GAAGGTGACCAAGTTCATGCTGAAAA GTGTTGGTGTGAAAAGTCTTTGA	GAAGGTCGGAGTCAACGGATTAAAAG TGTTGGTGTGAAAAGTCTTTGG	CGTAGATACTTGCTCTCT TCAAAACAAAAAAT
FB1 019	GAAGGTGACCAAGTTCATGCTTATTAT TTTCCAATTCACGTTCATATGCGT	GAAGGTCGGAGTCAACGGATTTTATTT TCCAATTCACGTTCATATGCGC	GTATGGTGCGTCTGCAGC CTTGAA
FB1 020	GAAGGTGACCAAGTTCATGCTTTCAA ATTATTCTCACAAAGGCGTATAACT	GAAGGTCGGAGTCAACGGATTCAAAT TATTCTCACAAAGGCGTATAACG	CTCCTGACGGTTAGTCGA GGATAAA
FB1 021	GAAGGTGACCAAGTTCATGCTTCTAA GGCAATCCAAGGGTGCG	GAAGGTCGGAGTCAACGGATTGTCTA AGGCAATCCAAGGGTGCA	TGACCAGGCTTATGACGT TCATCAATTT
FB1 022	GAAGGTGACCAAGTTCATGCTAGTTA ATTGACCAAGAGACTAAAAATCAGA	GAAGGTCGGAGTCAACGGATTAGTTA ATTGACCAAGAGACTAAAAATCAGT	TTTGGTGCTACAATTGAC TTTTTCATCAGTA
FB1 024	GAAGGTGACCAAGTTCATGCTTTCGA ATGAACAAAAGTTTTCTAGTCTCAA	GAAGGTCGGAGTCAACGGATTCGAAT GAACAAAAGTTTTCTAGTCTCAG	CTGCAAGATTCACCATTT GAAGTGACAA
FB1 025	GAAGGTGACCAAGTTCATGCTGGCCT GTTTGGACTTTTGCCAG	GAAGGTCGGAGTCAACGGATTTTGGC CTGTTTGGACTTTTGCCAA	CGGTCAACGCTTCAGCAC GCTA
FB1 026	GAAGGTGACCAAGTTCATGCTTTCACC ACCATCTTCAAGATAATCTTC	GAAGGTCGGAGTCAACGGATTATTCA CCACCATCTTCAAGATAATCTTA	CAAAGATGAATGGGGTT GTTGAATCTATTAAT
FB1 028	GAAGGTGACCAAGTTCATGCTGGAGT GCAGTCCATCACACGG	GAAGGTCGGAGTCAACGGATTTGGAG TGCAGTCCATCACACGA	TTAGAGTGGTTGAAAGA GATTGAGAGAATATT
FB1 033	GAAGGTGACCAAGTTCATGCTAATTC ATTGATTCTTTGGACCAAAAACCTA	GAAGGTCGGAGTCAACGGATTAATTC ATTGATTCTTTGGACCAAAAACCTT	GTTAGCTTTGCACTTATG TATGGAATGATTA
FB1 034	GAAGGTGACCAAGTTCATGCTCTGCC AGCTGAACCACTCAAC	GAAGGTCGGAGTCAACGGATTCCTGC CAGCTGAACCACTCAAT	CGTTGGTTTTAATATTTG TTGTTTATTAACCATATT
FB1 035	GAAGGTGACCAAGTTCATGCTATGCA ATCCTCCGTCCCAGAC	GAAGGTCGGAGTCAACGGATTCATGC AATCCTCCGTCCCAGAT	GGAGGGATATGGCCGGA GAGATTT
FB1 036	GAAGGTGACCAAGTTCATGCTTTGGG ATTGGATGAAGAGGGTTACT	GAAGGTCGGAGTCAACGGATTGGGAT TGGATGAAGAGGGTTACG	TAACCTCAACTCCCAGCA TCCACAT
FB1 037	GAAGGTGACCAAGTTCATGCTACAAG TCTTGTACCTTACATTTCAACG	GAAGGTCGGAGTCAACGGATTCAACA AGTCTTGTACCTTACATTTCAACA	GGTTGGTCATTCTACTGG CAATTTCTTT
FB1 038	GAAGGTGACCAAGTTCATGCTTTTATC TTCAGCCATCTCGGATTCG	GAAGGTCGGAGTCAACGGATTTTTTAT CTTCAGCCATCTCGGATTCA	CTGATGAGCTAATCTTGC CGAGGATA
FB1 039	GAAGGTGACCAAGTTCATGCTTTCCTT GAGTGATTTGGCTTAATTCTG	GAAGGTCGGAGTCAACGGATTGATTC CTTGAGTGATTTGGCTTAATTCTT	CTTACATCCCCAGGAACG ATTTGCAAT

Continued Table 4

| ID	Primer AlleleX	Primer AlleleY	Primer Common
FB1 041	GAAGGTGACCAAGTTCATGCTCTCAA GATCCGGCTTCGCATTCT	GAAGGTCGGAGTCAACGGATTTCAAG ATCCGGCTTCGCATTCC	TCGGTAGAATTGGAAGC CGCGTTAT
FB1 046	GAAGGTGACCAAGTTCATGCTACAGT GTTAATGCAATGAACAGCACAA	GAAGGTCGGAGTCAACGGATTCAGTG TTAATGCAATGAACAGCACAG	GTTGTATCACGAGGTACA GGAATCGAA
FB1 047	GAAGGTGACCAAGTTCATGCTCGGCA TACCTAACTTCTGAGCG	GAAGGTCGGAGTCAACGGATTATACG GCATACCTAACTTCTGAGCA	TTCTCTACCATTGGAACC CTTTTGCAA
FB1 050	GAAGGTGACCAAGTTCATGCTTTCCAT TGTAGAAAACTAACTTATATTCGG	GAAGGTCGGAGTCAACGGATTCTTCCA TTGTAGAAAACTAACTTATATTCGA	GGTATTTAACAAAATATA TCATCCACTAAAACTTAT
FB1 052	GAAGGTGACCAAGTTCATGCTAGGGA TTATTCTTCCCGGCAGC	GAAGGTCGGAGTCAACGGATTAGGGA TTATTCTTCCCGGCAGG	CTAGACGGCTGAAGATA AAAGCATCTT
FB1 055	GAAGGTGACCAAGTTCATGCTAAGCA ATGGTGTTTAGGATCAAAGAAATT	GAAGGTCGGAGTCAACGGATTGCAAT GGTGTTTAGGATCAAAGAAATG	ACGAGGGTATGATGAGC GTCCTTTT
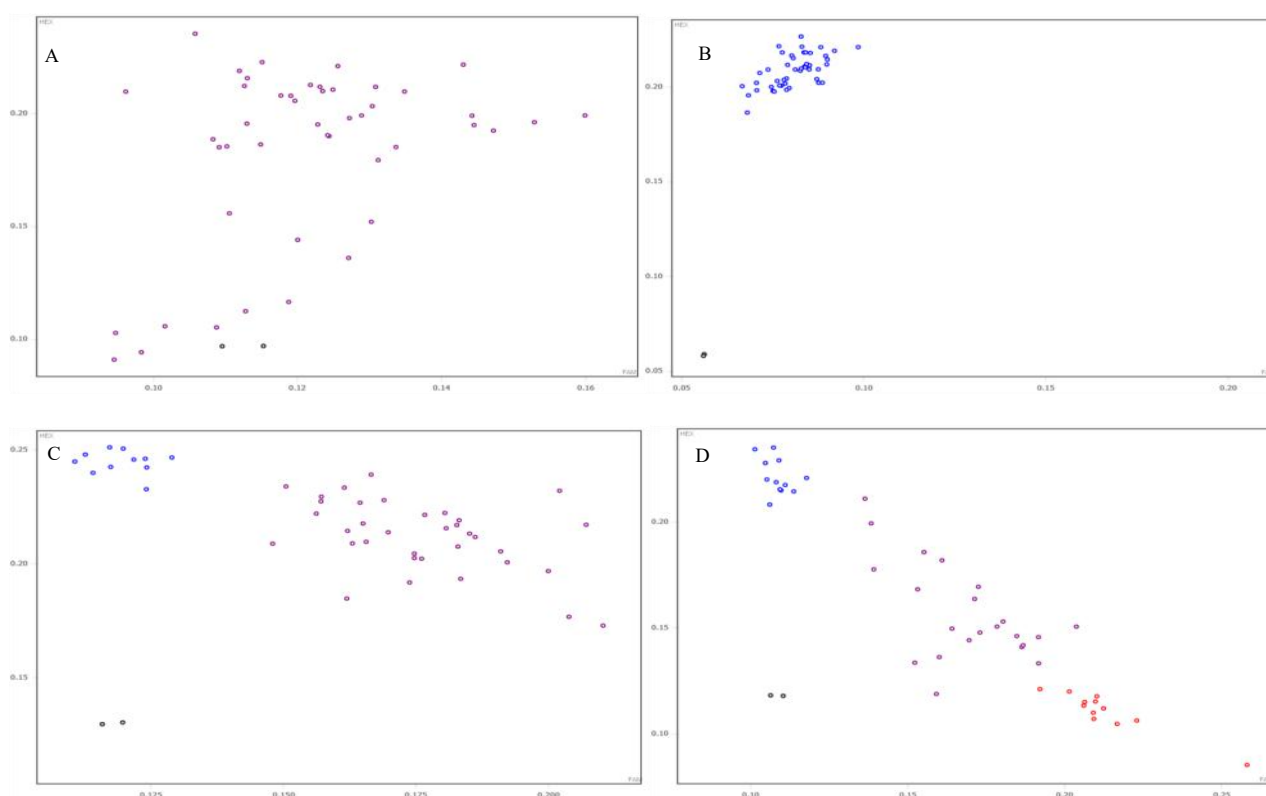
Figure 1 The genotype signals detedted by the KASP markers in this study

Note A: failure; B: single allele genotype; C: two alleles genotype; D: three alleles genotype

Faba bean is an often cross-pollinated crop with a natural crossing rate of 4%~84% (Kaur et al., 2012). Among the SNPs identified in this study, about 46% SNPs were heterozygous mutation (Table 2), indicating that higher heterozygous rate appeared in all the 8 landraces, which reminding that heterozygous genetic background should be paid more attention when analyzing the genetic diversity of faba bean germplasms. In this study, 31 SNPs were converted into KASP markers, of which 66.7% markers detected corresponding genotypes in different germplasm, which was close to the success rate of 67.6% for the KASP markers in Webb et al. (2016). As the KASP conversion rate in faba bean SNP is relatively low, the major reason is lack of reference genome, thus affecting the accuracy of SNP calling. Anyway, these SNPs provided a useful resource for genetic diversity assessment, gene mapping and molecular breeding in faba bean.

3 Materials and Methods

3.1 Materials
The faba bean germplasms used in this study was collected and preserved by Lishui Agricultural and Forestry Science Research Institute, among which 8 landraces (FB017, FB032, FB036, FB056, FB076, FB080, FB081) were used to sequence, and 46 germplasm were used to verify the accuracy of SNPs markers.

3.2 DNA extraction
The one week old young leaves were taken and grinded in liquid nitrogen, and then genomic DNA was extracted with a DNA extraction kit (Tiangen Biochemical Technology (Beijing) Co., LTD.). DNA quality was measured by agarose gel electrophoresis, and DNA concentration was detected by NanoDrop 2000 ultramicro spectrophotometer.

3.3 RAD library construction and sequencing
The RAD library was constructed according to the method of Baird et al. (2008). Genomic DNA was digested with *EcoRI*, and the 3' end of the digested fragment was ligated to a unique nucleotide multiplex identifier (MID) for barcoding. The genomic DNA libraries were enriched by PCR amplification and the PCR fragments were purified by agarose gel electrophoresis. Finally the single-end sequencing was performed using Illumina HiSeq 2000 platform.

3.4 SNP identification
The raw sequences were generated by Illumina base calling software CASAVA v1.8.2 (http://support.illumina.com/sequencing/sequencing_software/casava.ilmn), and quality analysis was performed using the Trimmomatic (http://www.usadellab.org/cms/uploads/supplementary/Trimmomatic) under the default parameters. After removing the reads ≤ 85 bp, the high quality reads were clustered into read tags (RAD-tags) using *ustacks* (Catchen et al., 2011) based on sequence similarity. According to the SNP calling method in Xu et al. (2014), the RAD-tags were then collapsed into clusters using *cstacks* under default parameters for SNP calling, finally the SNP genotypes were corrected again by the Bayesian algorithm (Hohenlohe et al., 2010).

3.5 KASP marker development and genotyping
KASP primers were designed using the Kraken™ software (LGC, Biosearch Technologies) based on the 100 bp fragment covering the SNP site. Each KASP marker includes two allele-specific forward primers and one common reverse primer. KASP genotyping was performed using the IntelliQube workstation (LGC, Biosearch Technologies) in the central laboratory of Zhejiang Academy of Agricultural Sciences. The PCR reaction was performed in a 1.6 μL volume containing 0.8 μL DNA template, 0.8 μL mixture of 2 x Master mix and Primer mix. PCR program includes 36 cycles following this protocol: 15 min at 95°C; 10 touchdown cycles of 20 s at 94°C and 60 s at 61°C~55°C (with the annealing temperature for each cycle reduced by 0.6°C per cycle); and 26 cycles of 20 s at 94°C and 60 s at 55°C. The genotype signals were detected by the own analysis system in IntelliQube.

References
Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A., Selker E.U., Cresko W.A., Johnson E.A., 2008, Rapid SNP discovery and genetic mapping using sequenced RAD markers, PLoS ONE, 3, e3376
https://doi.org/10.1371/journal.pone.0003376

Catchen J.M., Amores A., Hohenlohe P., Cresko W., and Postlethwait J.H., 2011, Stacks: building and genotyping loci de novo from short-read sequences, G3, 1: 171-182

https://doi.org/10.1534/g3.111.000240

Hegarty M, Yadav R, Lee M, Armstead I., Sanderson R., Scollan N., Powell W., and Skøt L., 2013, Genotyping by RAD sequencing enables mapping of fatty acid composition traits in perennial ryegrass (*Lolium perenne* (L.)), Plant Biotechnol J., 11: 572-581

https://doi.org/10.1111/pbi.12045

Hohenlohe P.A., Bassham S., Etter P.D., Stiffler N., Johnson E.A., and Cresko W.A., 2010, Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags, PLoS Genet, 6, e1000862

https://doi.org/10.1371/journal.pgen.1000862

Johnston J.S., Bennett M.D., Rayburn A.L., Galbraith D.W., and Price H.J., 1999, Reference standards for determination of DNA content of plant nuclei, Am J Bot, 86: 609

https://doi.org/10.2307/2656569

Kaur S., Pembleton L., Cogan N., Savin K., Leonforte T., Paull J., Materne M., and Forster J., 2012, Transcriptome sequencing of field pea and faba bean for discovery and validation of SSR genetic markers, BMC Genomics, 13: 104

https://doi.org/10.1186/1471-2164-13-104

Muňoz-Amatriaín M., Mirebrahim H., Xu P., Wanamaker S.I., Luo M.C., Alhakami H., Alpert M., Atokple I., Batieno B.J., Boukar O., Bozdag S., Cisse N., Drabo I., Ehlers J.D., Farmer A., Fatokun C., Gu Y.Q., Guo Y.N., Huynh B.L., Jackson S.A., Kusi F., Lawley C.T., Lucas M.R., Ma Y., Timko M.P., Wu J., You F., Barkley N.A., Roberts P.A., Lonardi S., and Close T.J., 2017, Genome resources for climate-resilient cowpea, an essential crop for food security, Plant J., 89: 1042-1054

https://doi.org/10.1111/tpj.13404

Ocaña S., Seoane P., Bautista R., Palomino C., Claros G.M., Torres A.M., and Madrid E., 2015, Large-Scale Transcriptome Analysis in Faba Bean (*Vicia faba* L.) under Ascochyta fabae Infection, PLoS ONE, 10(8): e0135143

https://doi.org/10.1371/journal.pone.0135143

Poland J.A., Brown P.J., Sorrells M.E., and Jannink J.L., 2012, Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach, PLoS ONE, 7: e32253

https://doi.org/10.1371/journal.pone.0032253

Rispail N., Kal P., Kiss G.B., Ellis N.T.H., Gallardo K., Thompson R.D., Prats E., Larrainzar E., Ladrera R., González E.M., Igor C.A., Ferguson B.J., Gresshoff P.M., and Rubiales D., 2010, Model legumes contribute to faba bean breeding, Field Crop Res., 115: 253-269

https://doi.org/10.1016/j.fcr.2009.03.014

Rowe H.C., Renaut S., and Guggisberg A., 2011, RAD in the realm of next-generation sequencing technologies, Mol. Ecol, 20: 3499-3502

https://doi.org/10.1111/j.1365-294X.2011.05197.x

Song Q., Jia G., Hyten D., Jenkins J., Hwang E., Schroeder S.G., Osorno J.M., Schmutz J., Jackson S.A., McClean P.E., and Cregan P.B., 2015, SNP Assay Development for Linkage Map Construction, Anchoring Whole-Genome Sequence, and Other Genetic and Genomic Applications in Common Bean, G3, 5: 2285-2290

https://doi.org/10.1534/g3.115.020594

Vignal A., Milan D., SanCristobal M., Eggen A., 2002, A review on SNP and other types of molecular markers and their use in animal genetics, Genet Sel Evol, 34: 275

https://doi.org/10.1186/1297-9686-34-3-275

Wang H., Guan J., Sun X., Ma Y., and Zong X., 2011, Genetic diversity analysis and ISSR analysis of genetic similarity of broad bean germplasm resources in the world, Chinese Agricultural Sciences, 44(5): 1056-1062

Wang S.C., Wong D., Forrest K., Allen A., Chao S., Huang B.E., Maccaferri M., Salvi S., Milner S.G., Cattivelli L., Mastrangelo A.M., Whan A., Stephen S., Barker G., Wieseke R., Plieske J., Lillemo M., Mather D., Appels R., Dolferus R., Guedira G.B., Korol A., Akhunova A.R., Feuillet C., Salse J., Morgante M., Pozniak C., Luo M.C., Dvorak J., Morell M., Dubcovsky J., Ganal M., Tuberosa R., Lawley C., Mikoulitch I., Cavanagh C., Edwards K. J., Hayden M., and Akhunov1 E., 2014, Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array, Plant Biotechnology Journal, 12: 787-796

https://doi.org/10.1111/pbi.12183

Webb A., Cottage A., Wood T., Khamassi K., Hobbs D., Gostkiewicz K., White M., Khazaei H., Ali M., Street D., Duc G., Stoddard F., Maalouf F., Ogbonnaya F., Link W., Thomas J., and Sullivan D., 2016, A SNP-based consensus genetic map for synteny-based trait targeting in faba bean (*Vicia faba* L.), Plant Biotechnology Journal, 14: 177-185

https://doi.org/10.1111/pbi.12371

Wu X., Ren C., Joshi T., Vuong T., Xu D., and Nguyen H.T., 2010, SNP discovery by high-throughput sequencing in soybean, BMC Genomics, 11: 469

https://doi.org/10.1186/1471-2164-11-469

Xu P, Xu S, Wu X, Tao Y, Wang B., Wang S., Qin D., Lu Z., and Li G., 2014, Population genomic analyses from low-coverage RAD-Seq data: a case study on the non-model cucurbit bottle gourd, Plant J., 77: 430-442

https://doi.org/10.1111/tpj.12370

Yang H, Tao Y, Zheng Z., Zhang Q., Zhou G., Sweetingham M.W., Howieson J.G., Li C., 2013, Draft genome sequence, and a sequence-defined genetic linkage map of the legume crop species *Lupinus angustifolius* L., PLoS ONE, 8: e64799
https://doi.org/10.1371/journal.pone.0064799

Yang S., Liu R., Yang T., Zhang H., Du M., and Zong X., 2016, Accurate evaluation of seed phenotypic traits of broad bean germplasm resources, Chinese Vegetables, (10): 32-40

Yang T., Bao S., Ford R., Jia T., Guan J., He Y., Sun X., Jiang J., Hao Ju., Zhang X., and Zong X., 2012, High-throughput novel microsatellite marker of faba bean via next generation sequencing, BMC Genomics, 13: 602
https://doi.org/10.1186/1471-2164-13-602