BSA-seq and genetic mapping reveals AhRt2 as a candidate gene responsible for red testa of peanut

Kun Zhang1,4,6 · Mei Yuan3 · Han Xia1,2 · Liangqiong He5 · Jing Ma1,2 · Mingxiao Wang1,2 · Huiling Zhao1,2 · Lei Hou1,2 · Shuzhen Zhao1 · Pengcheng Li1 · Ruizheng Tian1 · Jiaowen Pan1 · Guanghui Li1 · Mahendar Thudi1,7 · Changle Ma2 · Xingjun Wang1,2 · Chuanzhi Zhao1,2

Received: 24 November 2021 / Accepted: 28 January 2022 / Published online: 15 February 2022 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The candidate recessive gene AhRt2 responsible for red testa of peanut was identified through combined BSA-seq and linkage mapping approaches.

Abstract The testa color of peanuts (Arachis hypogaea L.) is an important trait, and those with red testa are particularly popular owing to the high-anthocyanin content. However, the identification of genes underlying the regulation of the red testa trait in peanut are rarely reported. In order to fine map red testa gene, two F2:4 populations were constructed through the cross of YZ9102 (pink testa) with ZH12 (red testa) and ZH2 (red testa). Genetic analysis indicated that red testa was controlled by a single recessive gene named as AhRt2 (Red testa gene 2). Using BSA-seq approach, AhRt2 was preliminary identified on chromosome 12, which was further mapped to a 530-kb interval using 220 recombinant lines through linkage mapping. Furthermore, functional annotation, expression profiling, and the analyses of sequence variation confirmed that the anthocyanin reductase namely (Arahy.IK60LM) was the most likely candidate gene for AhRt2. It was found that a SNP in the third exon of AhRt2 altered the encoding amino acids, and was associated with red testa in peanut. In addition, a closely linked molecular marker linked with red testa trait in peanut was also developed for future studies. Our results provide valuable insight into the molecular mechanism underlying peanut testa color and present significant diagnostic marker resources for marker-assisted selected breeding in peanut.

Introduction

Peanut (Arachis hypogaea L.) is widely grown in more than 100 countries, with a total production of approximately 48.8 million tons in 2019 (http://www.fao.org/faostat/en/#data/QC). Peanut is known as an important cash crop not only because of its high quality cooking oil but also for a variety of snacks. Apart from these applications, it is also known as a rich source of high nutritional value due to various nutritional ingredients, such as vitamin B1, B3, B9 and E, biotin, resveratrol, isoflavones, phytic acid, anthocyanin and procyanidins. Studies have shown that most of these nutritional components are accumulated in the testa (seed...
The testa color is an important trait of peanut with numerous variations, such as white, pink, red and black (or deep purple). The majority of peanut varieties have been identified with pink testa color. Anthocyanin content and composition are important factors to determine the color of testa. In higher plants, there are six kinds of anthocyanins including delphinidin, cyanidin, pelargonidin, peonidin, petunidin and malvidin. Significant differences were found in the types and content of anthocyanidin considering different testa color peanuts, and the content of delphinidin, cyanidin and pelargonidin were closely related to the testa color (Li et al. 2021). Anthocyanins have showed strong antioxidant capacity and important nutritional value (Shin et al. 2006; Winkel-Shirley 2001). Importantly, high-anthocyanin varieties have become one of the important directions in rice and wheat breeding (Giordano et al. 2017; Ito & Lacerda 2019). Testa color phenotype arises one generation later than other traits because the testa was developed by the integument and had the same genotype as its maternal plant (Chen et al. 2021; Zhao et al. 2020). For traditional breeding methods, phenotype of the testa color can only be identified after harvest, which prolonged the screening time. Testa color selection can be significantly assisted using marker-assisted selection (MAS), which prioritizes genotype-specific selection of target progenies. Understanding the genetics and developing DNA markers associated with the testa color trait is essential for MAS.

With the rapid development of next-generation sequencing (NGS), considerable progress has been made in peanut whole genome sequencing including both wild-type and cultivated peanuts (Bertioli et al. 2016, 2019; Chen et al. 2019; Yin et al. 2018; Zhuang et al. 2019). The availability of complete genome sequence of peanut provided ideal resource for genome-wide identification of SSR and SNP markers in silico (Zhao et al. 2017; Ma et al. 2020). In addition, large-scale 58 K SNP Array (Axiom_Arachis, v1) and 48 K SNP Array (Axiom_Arachis2, v2) also facilitated new avenues for constructing high-resolution linkage maps, genetic analysis and QTL mapping (Clevenger et al. 2017, 2018b; Nabi et al. 2021). Moreover, great efforts have been made in the genetic map construction, fine mapping and MAS of peanut (Agarwal et al. 2018; Han et al. 2018). The major gene related to the black testa color of peanut have been identified through BSA-seq and eQTL approaches (Zhao et al. 2020; Huang et al. 2020). The previous studies suggested that peanut red testa was controlled by one dominant gene (R1) and two recessive genes (r2, r3), and all these three genes appeared to inherit independently (Branch 2011). Similarly, an important dominant gene AhRt1 controlling red testa color was recently fine-mapped on chromosome A03, which supplied the closely connected markers for MAS breeding. (Chen et al. 2021; Zhuang et al. 2019). However, there is no explicit report available on the recessive gene controlling the red testa color in peanut. In the present study, one recessive gene controlling red testa, AhRt2, was fine-mapped to a 0.5 Mb genomic region on chromosome 12 using the BSA-seq and linkage mapping approaches. An anthocyanidin reductase (ANR) gene was suggested to be the possible candidate gene. A “G/A” SNP is in the third exon of the ANR gene. In addition, we also developed tightly linked molecular markers which could be used in future MAS breeding programs.

### Materials and methods

#### Development of mapping population

The pink testa peanut cultivar Yuanza 9102 (YZ9102) was used as the female parent to cross the red testa peanut cultivars Zhonghua 12 (ZH12) and Zhanhong 2 (ZH2), respectively, for inheritance study and construction of the mapping population. (Table 1). After received pollen from the male parents, the hybrid seeds collected from the female parents are “F1 seeds.” The plants from the germination of “F1

| Population                        | Generation | Total seeds | Pink | Red | Expected ratio | $\chi^2$ | $p$ value |
|----------------------------------|------------|-------------|------|-----|----------------|--------|----------|
| YZZH12 (YZ9102 × Zhong-Hua 12)   | F$_{2\times4}$ | 220         | 143  | 77  | 5:3            | 0.59   | 0.44     |
| YZZH2 (YZ9102 × Zhan-hong 2)     | F$_{2\times4}$ | 271         | 176  | 95  | 5:3            | 0.69   | 0.41     |

*Expected ratio: Dominant homozygous: Heterozygous: Recessive homozygous = 3:2:3, Dominant trait individual: Recessive trait individual = 5:3

$\chi^2$: $\chi^2 < 3.84$ is considered as significant correlated

$p$ value: $p$ value > 0.05 indicate no statistically significant differences
seeds” are “F1 plants.” The self-pollinated seeds collected from an F1 individual plant are corresponding to F2 family. Two populations (YZZH12 and YZZH2) were developed through single-seed-descent method. The YZZH12 was used for BSA-seq and the further fine-mapping analysis, and the two populations was used for the validation of diagnostic markers. The F1, F2 and F2:3 plants of two populations were planted in the field from 2017 to 2019 at Laixi Experimental Station (36°48′47.41″ N, 120°30′22.54″ E), Shandong Peanut Research Institute (SPRI), Shandong, China. The F2:4 plants were planted in the field in 2020 at Jiyang Experimental Station of Shandong Academy of Agricultural Sciences (SAAS), Shandong, China (36°58′34.53″ N, 116°59′1.29″ E).

Measurement of anthocyanin

The anthocyanin content of peanut testa was measured using previous methods with slight modifications (Mancinelli et al. 1991; Teng et al. 2005; Zhao et al. 2020). At least three sets of more than eight seeds per sample were used. The growth period of these seeds was 50 days after pegging, and the peanut testa was freshly peeled and quickly frozen into liquid nitrogen. In brief, frozen peanut testa (approximately 50 mg) was ground in a 5 mL centrifuge tube using liquid nitrogen. Then, homogenized testa was extracted at 4 °C by adding 700 µl acidic methanol (the volume ratio of methanol to HCl is 99:1). After overnight incubation, the homogenates were centrifuged for 1 min at 12,000 rpm for 10 min. The supernatant (approximately 600 µl) was collected and mixed with 1 mL trichloromethane and 400 µl distilled water, and centrifuged at 4 °C at 12,000 rpm for 10 min. Then, the absorbance of the supernatant was measured in spectrophotometer (U-3000, HITACHI, Japan) at 530 and 657 nm, respectively. The relative anthocyanin content was calculated according to the absorbance with the formula of \[ A = (1/4 \times A_{530}) - (1/4 \times A_{657}) \] then and normalized by sample weight.

Whole genome sequencing and BSA-seq analysis

The genomic DNA was extracted from the leaves using the DNA Extraction Kit (DP305) of TIANGEN Biotech (Beijing, China) according to the manufacturer’s instructions. DNA quality was determined using the BioPhotometer plus spectrophotometer (Eppendorf AG, Hamburg, Germany) and 1% agarose gel electrophoresis. For BSA-seq, the population of YZZH12 was used. Two DNA pools were constructed by mixing equal amounts of DNA from 30 red testa F2:4 individuals (Red-pool) and 30 pink testa F2:4 individuals (Pink-pool). The two DNA pools of YZ9102 and ZH12 were sequenced on BGISEQ-500 platform at the Beijing Genomics Institute (BGI). After sequencing, clean reads were obtained by removing low-quality and short reads using Soapnuke program (Chen et al. 2018), and mapped on reference genome of cultivated peanut Tifrunner (https://peanutbase.org/peanut_genome) using BWA software with the SAM tools (Li and Durbin 2009). Single-nucleotide polymorphism (SNP) and Insertion/Deletion (InDel) were called, and filtered by removing heterozygous and missing SNPs as well as Indels in the pools and parental lines using GATK software (McKenna et al. 2010). The SNP-index represents the ratio of reads harboring SNPs among the entire number of reads (Abe et al. 2012). The ΔSNP-index was the difference of the SNP-indices between bulks. To identify candidate regions associated with the red testa trait, the ΔSNP-index of each locus was calculated by subtracting the SNP-index of the Pink-pool from that of the Red-pool according to previous method (Takagi et al. 2013). To confirm the results of ΔSNP-index, Euclidean Distance (ED) algorithm was further preformed to identify the SNPs and Indels associated with the red testa trait using the equation reported previously (Lei et al. 2020; Hill et al. 2013). The greater the ΔSNP-index and ED, the more likely the SNPs and Indels contributes to the trait of red testa color or is linked to a gene that controls the trait.

Marker development, genetic map construction and mapping

To validate the BSA-seq results and further narrow down the region, 21 Indels in the candidate region were selected according to the comparative genomic information among the parents. Twenty-one primer pairs were designed to the flanking sequences of the targeted SNPs using Primer Premier 5.0 (http://www.premierbiosoft.com/primerdesign/). The polymorphism of these Indels was confirmed through polyacrylamide gel electrophoresis as described previously (Zhao et al. 2017). The Indel markers showing polymorphism among the parents were further used to genotype YZZH12. The sequences of primers used for mapping are listed in Supplementary Table S1. Genetic linkage map on peanut chromosome 12 was constructed using JoinMap 5.0 software (https://www.kyazma.nl/index.php/JoinMap/). The recombinant ratio was converted into genetic distances (centimorgans, cM) through the function of Kosambi map. The linkage groups were calculated at a minimum logarithm of odds (LOD) score of 5. MapChart 2.3 software was used for drawing the linkage maps (Voorrips 2002). For QTL analysis, inclusive composite interval mapping of additive (ICIM-ADD) was performed using software QTL IciMapping V4.1 (Li et al. 2007; Meng et al. 2015). The trait of pink testa set to 1 and the red testa set to 0. The LOD threshold was set at 2.5 to determine the presence of a putative QTL associated with a target trait.
Prediction of candidate genes

The sequences of gene information in the candidate interval were obtained according to the cultivated peanut reference genome sequences (Version 1, https://peanutbase.org). The functions of candidate genes were annotated through Blastx program in databases of Nr (NCBI, http://www.ncbi.nlm.nih.gov), GO (https://www.geneontology.org/), KOG (http://www.ncbi.nlm.nih.gov/KOG), and KEGG (http://www.genome.jp/kegg). For sequence alignment and phylogenetic analysis of the anthocyanidin reductase (ANR) genes, the amino acid sequences of ANR proteins were downloaded from NCBI. Multiple sequence alignments were performed by clustalw (https://www.genome.jp/tools-bin/clustalw) and BoxShade online program (https://embnet.vital-it.ch/software/BOX_form.html). The phylogenetic analysis was carried out by using MEGA7 (https://www.megasoftware.net/) with the neighbor-joining statistical method. The GenBank accession numbers of these proteins are provided in Supplementary Table S2.

Quantitative real-time PCR (qRT-PCR)

We selected genes from the anthocyanin metabolic pathway, transcription factors that may regulate flavonoid biosynthesis, and genes that may play some roles in other secondary metabolite pathways based on the functional annotation of genes. A total of 13 genes were subjected to qRT-PCR analysis. After harvesting from the field, the seed coat was collected immediately before drying. The samples used in qRT-PCR analysis were the same as used for BSA-seq, including two parents and RNA pools, and three biological replicates were prepared for every sample. Total RNA was extracted using Trizol Reagent kit (TaKaRa, Inc., Dalian, China) according to the instructions of manufacturer. The reverse transcriptions were performed with PrimeScript II 1st Strand cDNA Synthesis Kit (TaKaRa, Inc., Dalian, China). The primers for qRT-PCR were designed using Primer Premier 5.0 (http://www.premierbiosoft.com/primerdesign/) and listed in Supplementary Table S1. The qRT-PCR reactions were performed on ABI7500 Real-Time System (USA) as described by previous studies (Wang et al. 2017). The relative expression levels of genes were calculated by $2^{-\Delta\Delta CT}$ method (Livak KJ et al. 2001). The $C_T$ (cycle threshold) was defined as the number of cycles required for the fluorescent signal to cross the threshold (i.e., exceeds background level).

Results

Phenotyping and genetic analysis of red testa in peanut

The phenotypic analyses showed significant differences in testa color between the two parents. For YZ9102, the testa color is traditionally pink, while ZH12 is red in testa color (Fig. 1A). The red testa has more anthocyanin than that in pink one (Fig. 1B). The total content of anthocyanin in the red testa lines is about 2–7 times higher than that in the pink lines. This is because the testa was developed from integument cells which was female somatic cells, the testa color was in consistence with the color of female parents. All F1 and F2 seeds showed pink testa, and the F2:3 seeds displayed different testa color corresponding to the coloration of either YZ9102 or ZH12/ZH2. Therefore, the pink testa to red testa was a completely dominant trait in the two populations. According to the Law of Segregation, dominant homozygous to heterozygous and to recessive homozygous was 3 to 2 to 3, and dominant trait individual to recessive trait individual was 5:3 in F2:4 populations. Among the 220 F2:4 individuals of YZZH12 population, 143 exhibited pink testa, and 77 showed red testa, corresponding to a segregation ratio of 5:3 by the Chi-square test ($\chi^2 = 0.59 < 3.84$, $p$ value > 0.05 $\chi^2 < 3.84$ indicates that the theoretical segregation ratio is

![Phenotypes and anthocyanin content](https://example.com/figure1.png)
significantly correlated with the actual). Similarly, YZZH2 population. \( F_{2.4} \) individuals also showed a 5:3 segregation ratio of pink and red \( (\chi^2 = 0.69 < 3.84, p \text{ value} > 0.05) \) (Table 1). These results demonstrated that the red testa of peanut should be controlled by a single recessive gene.

**BSA-seq analysis and mapping of gene AhRt2**

In total, 31.57, 41.22, 113.25 and 120.38 Gb raw data were generated for the YZ9102, ZH12, Pink-pool and Red-pool, representing approximately 12.34x, 16.12x, 47.08x and 44.29x genome coverage, respectively (Table 2). The filtered clean reads of each sample were mapped to the reference genome of the cultivar Tifrunner and a total of 412,874 SNPs/InDels were identified. To obtain the genomic region associated with the red testa, two approaches including, \( \Delta \text{SNP-index} \) and ED algorithms were performed to calculate the allele segregation of the SNPs and InDels between the two extreme DNA pools. \( \Delta \text{SNP-index} \) and ED algorithms showed that 54.56% and 70.11% candidate SNPs/InDels enriched on chromosome 12 (Chr.12), respectively (Fig. 2A, B). On Chr.12, the 7.8 Mb region (109.9 Mb-117.7 Mb) exhibiting significant linkage disequilibrium was identified as the candidate region for red testa (Fig. 2C).

**Table 2** The quality, sequencing depth and coverage of re-sequencing data of BSA-seq

| Samples    | Clean reads | Clean bases     | Number of reads mapped | Mapping ratio (%) | Depth(x)  | Properly paired read | Properly paired ratio (%) |
|------------|-------------|-----------------|------------------------|-------------------|-----------|----------------------|--------------------------|
| YZ9102     | 210,495,768 | 31,574,365,200  | 210,632,335            | 99.97             | 11.28     | 209,080,834          | 99.33                    |
| ZH12       | 274,833,602 | 41,225,040,300  | 275,077,939            | 99.99             | 14.72     | 272,559,740          | 99.17                    |
| Pink-pool  | 755,040,492 | 113,256,073,800 | 803,084,248            | 99.95             | 40.45     | 795,201,596          | 99.08                    |
| Red-pool   | 802,571,036 | 120,385,655,400 | 754,686,636            | 99.83             | 42.99     | 747,421,150          | 98.99                    |

**Fig. 2** Distribution of candidate SNPs and InDels per chromosome. **a** Candidate SNPs and InDels using \( \Delta \text{SNP-index} \) algorithm with a cut-off of \( \Delta \text{SNP-index} > 0.5 \). **b** The top 1% SNPs and InDels using ED algorithm. **c** Distribution of candidate SNPs/InDels on Chr.12. The significant region identified for red testa phenotype is shaded (109.9–117.7 Mb)
Fine mapping of the *AhRt2* Gene

According to the BSA-seq results, 21 InDel markers were developed in the candidate region of Chr.12, and 11 of them displayed stable polymorphisms between the parental line and the F2 individuals. These markers were used for constructing the genetic map of candidate region and QTL mapping using the 220 F2,4 individuals of YZZH12 population. The total length of the linkage map is 59.27 cM, which is corresponding to 110.05 Mb to 177.62 Mb region of chr.12. A QTL was detected in interval InDel_16 to InDel_18, and its LOD score was 46.01, it could explain 59.35% of the phenotypic variation. The other QTL was detected in interval InDel_18 to InDel_20, its LOD score was 50.12, and it could explain 63.82% of the phenotypic variation. The major gene was mapped in a 3.64 cM region between Indel marker InDel_16 and InDel_20 (Fig. 3). The physical location of *AhRt2* was narrowed in a 0.53 Mb region (Chr12: 117.03 Mb-117.56 Mb) (Fig. 3).

Detection of the SNPs of candidate gene in different peanut germplasm resources

To further analyze the function of the candidate gene, we detected the two SNPs of *Arahy.IK60LM* in peanut germplasm resources with different testa colors, including five pink testa, eight red testa, one white testa, two black and two black stripe peanuts. For the SNP at the upstream of the CDS, most of the germplasms were with the same genotype as the red parent, only two peanuts with black stripe with the same genotype as YZ9102. For the SNP in the third exon of the candidate gene, the “G to A” was specific only in two red parents ZH12 and ZH2 (Fig. 4). It implied that there was more than one gene controlling the red testa of cultivated peanut. The molecular mechanism of the other six red testa varieties was different from that of the ZH12 and ZH2.

Development of diagnostic marker for screening the peanut with red testa

We developed a KASP marker *KASP_AhRt2* according to the sequence alignment in the candidate interval. A total of 220 lines of YZZH12 populations and 57 lines of YZZH2

![Fig. 3](image-url)
populations were used for this site detection, and found that the locus of almost all red testa lines was “T:T,” and the pink testa lines were “C:C” or “T:C,” except YZZH12-127 and YZZH12-181 (Fig. 5). The two lines showed red phenotype with “T:C” genotype. We found that the “SNP_117190528” of the two lines was consistent with the red parent by

Fig. 4  Gene structure and expression analysis of candidate gene. a Gene structure of candidate gene and the locations of the SNPs. b Sequence alignment of the ANR in different species. c Expression of the candidate gene in the parental lines and pools with the red and pink testa, respectively (t-test; **P < 0.01). Error bars represent ± SE (n = 3)

Fig. 5  Detecting of SNPs of candidate gene of AhRt2 in different peanut germplasms. a Detecting of SNPs at the −312 bp of the initiation codon of the candidate gene. b Detecting of SNPs at the third exon of AhRt2
sequencing PCR products. Therefore, the two lines should be the linkage exchange between the KASP_AhRt2 and SNP_117190528. In addition, our results suggested that the KASP marker can be used as the diagnostic genotyping marker to predict red testa peanut through MAS. However, the diagnostic marker is only used in the cross event using the ZH12 or ZH2 as donor parent for the red testa and it needs further validation in different genetic backgrounds.

Identification of the candidate genes related to red testa

There are 52 genes in the candidate 0.53 Mb region of Chr.12. Among them, two genes (Arahy.JFV18T and Arahy.X3FWT9) were annotated as bHLH transcriptional factor encoding-genes, one of the core members to form the MYB-bHLH-WD40 (MBW) complex, and the latter is regarded as an important regulatory gene in plant anthocyanin biosynthesis. In addition, the candidate interval contains one anthocyanidin reductase (ANR) gene (Arahy.IK60LM), one of the structural genes in anthocyanin metabolism pathway, and named as AhANR1 (Table 3). Re-sequencing results showed that there are 36 SNPs and 7 InDels in the candidate interval. Most of these SNPs/InDels were located in the intergenic region, and 9 SNPs were located in the gene region including two SNPs in the AhANR1 gene region (Table 4). One SNP “SNP_11719149” was located in the -312 bp of the upstream of the coding DNA sequence (CDS). The other SNP “SNP_117190528” was located in the third exon of the ANR gene with a C/T variation, which leads to a transition from Threonine (A_CT) in YZ9102 to Isoleucine (A_TT) in ZH12 (Fig. 6A).

Sequence alignment and expression analysis of the candidate genes

The sequence of AhANR1 of YZ9102 was identical to that of A. duranensis, A. ipanesis, Shitouqi and Fuhuasheng, which have been released with the whole genome sequences. Sequence alignment showed that threonine is conserved both in A. duranensis and A. ipanesis and other plant species, such as rice (Oryza sativa), common bean (Phaseolus vulgaris), soybean (Glycine max) and tobacco (Nicotiana tabacum) (Fig. 6B). To further analyze this candidate region, 13 genes were selected for qRT-PCR assay. Interestingly, we found that the expression of the ANR gene was up-regulated in both the red testa ZH12 and the red testa pools (Fig. 6C). The results of qRT-PCR showed that the C_T mean values of the two bHLH genes (Arahy.JFV18T and Arahy.X3FWT9) exceeded to 35, indicating that the expression level of these two bHLH genes is very low in both red and pink testa. We also detected the expression level of the candidate gene of AhRt1 in YZ9102 and ZH12, and found that the expression of AhRt1 is very low with the C_T mean > 35. Hence, these two bHLH genes and the candidate gene of AhRt1 were not likely the genes responsible for the red testa color in population used in this study. Taken together, we predicted that the AhANR1 might be the key gene controlling the red testa.

Discussion

BSA-seq is an effective strategy for gene fine mapping

BSA-seq approach is a rapid method for identifying markers linked to the traits through constructing the DNA pools with extreme traits, and identifying the resistance loci in lettuce (Michelmore et al. 1991). With the development of next-generation sequencing (NGS) technologies, the high throughput SNPs and InDels can be rapidly detected from the genome. Recently, a series of NGS-based BSA + approaches have been developed, including BSA-seq (or QTL-seq), BSR-seq, MutMup and MutMup + (Abe et al. 2012; Fekih et al. 2013; Steuernagel et al. 2016; Takagi et al. 2013). In comparison to the food crops, such as rice and wheat, the reproduction coefficient of peanut is very low. It is difficult to fine mapping the genes and QTLs through traditional mapping methods. For BSA-seq, huge segregation population is not required, and it has been successfully used in peanut for identification of the candidate interval and development linked molecular marker for many traits, such as disease resistance, shelling percentage, dormancy, black testa and red testa (Clevenger et al. 2018a; Luo et al. 2018; Pandey et al. 2017; Zhao et al. 2020; Chen et al. 2021). In this study, the candidate interval associated with the red testa was identified through the BSA-seq method. Then, the traditional mapping method was used to construct the genetic linkage map and finally narrow down to the candidate gene in a 530-kb interval. The strategy of combined BSA-seq and traditional mapping approach were used to accelerate the identification of the candidate gene AhRt2, which could be used to identify other genes in peanut and other crops.

AhRt1 and AhRt2 were mapped in homologous fragment in different chromosomes

Previous studies have suggested that at least three pairs of genes were responsible for the red testa of peanut including one dominant gene and two recessive genes (Branch 2011). Recently, AhRt1, a major dominant gene controlling red testa, was identified in a 580-kb interval of chromosome A03 (Chr.03). One of the bHLH transcriptional factors was suggested as the possible candidate gene for AhRt1 (Chen et al. 2021). In the present study, a recessive gene controlling red testa color was successfully fine-mapped to a
Table 3  Gene information in the candidate region on Chr.12

| Gene           | Start    | End       | Function                                                                 |
|----------------|----------|-----------|---------------------------------------------------------------------------|
| Arahy.1KM4NQ   | 117,005,569 | 117,007,230 | MYB transcription factor MYB48                                           |
| Arahy.IZM7T4   | 117,019,886 | 117,024,789 | Protein kinase superfamily protein                                        |
| Arahy.BL8SW3W  | 117,025,116  | 117,031,218 | Transcription factor jumonji (jmjC) domain-containing protein              |
| Arahy.VDG15V   | 117,045,486  | 117,047,405 | Octicosapeptide/Phox/Bem1p family protein                                |
| Arahy.B3SD72   | 117,048,799  | 117,051,828 | Unknown protein                                                           |
| Arahy.J2UCUC   | 117,052,969  | 117,055,610 | Cellulase (glycosyl hydrolase family 5)                                   |
| Arahy.6S25R2   | 117,069,253  | 117,081,682 | receptor-like serine/threonine kinase                                      |
| Arahy.WE5WNX   | 117,081,781  | 117,085,905 | receptor-like serine/threonine kinase                                      |
| Arahy.KK3MU5   | 117,089,125  | 117,096,591 | Unknown protein                                                           |
| Arahy.D43PH9   | 117,097,423  | 117,101,091 | receptor-like serine/threonine kinase                                      |
| Arahy.4HL1MH   | 117,115,390  | 117,118,208 | Quinone reductase family protein                                          |
| Arahy.X3FWT9   | 117,131,960  | 117,135,240 | transcription factor bHLH123-like                                         |
| Arahy.D0FPDW   | 117,169,647  | 117,170,588 | transcription factor DYSFUNCTIONAL TAPETUM 1-like                         |
| Arahy.EYN48E   | 117,173,896  | 117,177,659 | terpene synthase                                                          |
| Arahy.7M140W   | 117,182,257  | 117,188,503 | FYVE zinc finger protein                                                  |
| Arahy.IK60LM   | 117,188,946  | 117,191,576 | anthocyanidin reductase ((2S)-flavan-3-ol-forming)                        |
| Arahy.ME5340   | 117,192,337  | 117,194,676 | Unknown protein                                                           |
| Arahy.XPM13Z   | 117,195,285  | 117,195,741 | Unknown protein                                                           |
| Arahy.6345FV   | 117,195,818  | 117,200,790 | (2R)-phospho-3-sulfolactate synthase ComA                                 |
| Arahy.R4UZ9W   | 117,212,854  | 117,218,115 | ATP binding microtubule motor family protein                              |
| Arahy.OBBN6W   | 117,218,341  | 117,220,729 | Sulfite exporter TauE/SafE family protein                                |
| Arahy.92LHRM   | 117,231,502  | 117,234,533 | receptor-like kinase 1                                                    |
| Arahy.77GZB8   | 117,272,366  | 117,277,939 | branched-chain amino acid transaminase 2                                  |
| Arahy.U1BCOV   | 117,280,589  | 117,283,465 | branched-chain amino acid transaminase 2                                  |
| Arahy.BHC7EV   | 117,284,545  | 117,287,751 | Unknown protein                                                           |
| Arahy.FCW1MU   | 117,311,541  | 117,316,007 | CBS domain-containing protein with a domain of unknown function (DUF21) |
| Arahy.K6XKIL   | 117,316,426  | 117,318,345 | Unknown protein                                                           |
| Arahy.F10236   | 117,318,361  | 117,320,009 | Unknown protein                                                           |
| Arahy.VXEK44   | 117,322,449  | 117,330,848 | LL-diaminopimelate aminotransferase                                       |
| Arahy.3DBG0K   | 117,343,268  | 117,346,310 | malate dehydrogenase                                                     |
| Arahy.JFY18T   | 117,345,716  | 117,349,561 | basic helix-loop-helix (bHLH) DNA-binding superfamily protein             |
| Arahy.WHNG6D   | 117,365,459  | 117,375,941 | DHHC-type zinc finger family protein                                      |
| Arahy.UW1SN1   | 117,376,910  | 117,380,446 | inorganic pyrophosphatase                                                 |
| Arahy.T11W9N   | 117,377,016  | 117,379,835 | Unknown protein                                                           |
| Arahy.WTJ0L8   | 117,380,766  | 117,380,972 | Unknown protein                                                           |
| Arahy.JM8NV3   | 117,382,006  | 117,386,697 | Major facilitator superfamily protein                                     |
| Arahy.7CPH5R   | 117,414,427  | 117,417,456 | Glutamyl-tRNA reductase family protein                                    |
| Arahy.Z3ANCQ   | 117,417,878  | 117,423,386 | Unknown protein                                                           |
| Arahy.52ZQRU   | 117,419,135  | 117,420,277 | LisH/CRA/RING-U-box domains-containing protein                           |
| Arahy.SN4DPF   | 117,423,355  | 117,426,869 | Unknown protein                                                           |
| Arahy.TBP304   | 117,428,488  | 117,430,556 | ubiquitin carboxyl-terminal hydrolase-like protein                        |
| Arahy.LXS01Y   | 117,442,385  | 117,450,867 | LRR and NB-ARC domain resistance protein                                 |
| Arahy.YEAJ9J   | 117,460,320  | 117,470,837 | LRR and NB-ARC domain resistance protein                                 |
| Arahy.R7V7OJ   | 117,482,548  | 117,487,344 | LRR and NB-ARC domain resistance protein                                 |
| Arahy.FTDU9L   | 117,483,618  | 117,494,316 | Unknown protein                                                           |
| Arahy.U1EPQF   | 117,490,366  | 117,494,234 | LRR and NB-ARC domain resistance protein                                 |
| Arahy.Y1XBLJ   | 117,495,399  | 117,498,854 | Unknown protein                                                           |
| Arahy.1L8Q6L   | 117,501,455  | 117,505,674 | disease resistance protein (TIR-NBS-LRR class)                           |
| Arahy.P56WT2   | 117,510,793  | 117,512,637 | Unknown protein                                                           |
physical interval of 530 kb of Chr.12, and named as AhRt2. Functional annotation and sequence analysis suggested that AhRt2 could be the anthocyanidin reductase (ANR) coding gene. Cultivated peanut is an allotetraploid with AA and BB sub genomes, and Chr.12 (B02) is corresponding to Chr.02 (A02). However, sequence alignment showed that there is no allele of Arahy.IK60LM in Chr.02. Instead, Arahy.IK60LM was mapped with the ANR gene Arahy.W8TDEC of Chr.03 with identity of 97.38%. Interestingly, the physical distance between Arahy.W8TDEC and Arahy.07ZIFT (AhRt1) is only 262 Kb. Further sequence alignment showed that most of the genes in the candidate interval (Chr12: 117.03–117.56 Mb) could be mapped in Chr.03, suggesting that there is a homologous chromosome fragment between Chr.03 and Chr.12. Therefore, it is interesting that AhRt1 and AhRt2 were mapped in homologous fragment of different sub genome of peanut. In addition, the change of the expression level of ANR gene lead to premature growth of strawberry (Fischer et al. 2014) and early flowering of tobacco (Vinay et al. 2013). Besides, the increased expression level of ANR gene could improve the plants resistance to biotic stress tolerance (Vinay et al. 2013; Xin et al. 2020). However, the ANR gene has rarely been studied in peanuts. Here, we predicted that the AhANR1 might be the key gene responsible for the regulation of red testa. The AhANR1 showed not only the difference of expression level but also the sequence variation among the pink and red testa parents. Maybe, there is a cause-and-effect relationship between the difference of sequence variation and expression level, or they are independent of each other. The function of the AhANR1 genes needs to be further verified in future studies.

**Roles of anthocyanin reductase in peanut**

Anthocyanin reductase (ANR) is an important gene in the pathway of anthocyanin metabolism that confers the accumulation of anthocyanin in the tissue or organ of plants. For example, the mutant of ANR gene was associated with the red-brown grain color of soybean (Nik et al 2012) and “black” trait in pomegranate (Trainin et al. 2021). In addition, the change of the expression level of ANR gene lead to premature growth of strawberry (Fischer et al. 2014) and early flowering of tobacco (Vinay et al. 2013). Besides, the increased expression level of ANR gene could improve the plants resistance to biotic stress tolerance (Vinay et al. 2013; Xin et al. 2020). However, the ANR gene has rarely been studied in peanuts. Here, we predicted that the AhANR1 might be the key gene responsible for the regulation of red testa. The AhANR1 showed not only the difference of expression level but also the sequence variation among the pink and red testa parents. Maybe, there is a cause-and-effect relationship between the difference of sequence variation and expression level, or they are independent of each other. The function of the AhANR1 genes needs to be further verified in future studies.

**Conclusion**

In our study, one gene regulating red testa of cultivated peanut namely AhRt2 was mapped in a 530 kb region on chromosome 12 using BSA-seq and linkage mapping approaches. Both functional annotation and sequence analysis suggested that it is an anthocyanin reductase gene that is most likely involved in the regulation of red testa color in peanut. In addition, a diagnostic marker (KASP_AhRt2) was developed for MAS in peanut. This work lays the foundation for the further understanding of the regulation mechanisms...
of peanut red testa formation and molecular breeding of new varieties with red testa.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-022-04051-w.

Acknowledgments The authors thank Naveed Ahmad of Shandong Academy of Agricultural Sciences for editing the English text of the manuscript, and thank Dr Zhengzheng and Dr Ziqi Sun of Henan Academy of Agricultural Sciences for assisting in the KASP analysis.

Authors’ contributions statement CZ and XW conceived and designed the experiments. MY, JM, MW, HZ and JP developed the populations. KZ, HX, LH, LH, SZ, PL, RT, GL, MT and CM performed the experiments. CZ and KZ wrote the manuscript. XW and CZ revised the manuscript.

Funding This research is supported by National Natural Science Foundation of China (32072090, 31861143009), Key Research and Development Project of Shandong Province (2020LZGC001, 2021LZGC025), Famous subjects construction project of Shandong Agricultural and Engineering University, Agricultural scientific and technological innovation project of Shandong Academy of Agricultural Sciences, High level Foreign Experts Introduction Program (GL20200123001), and Taishan Scholar Project of Shandong Province (ts20190964).

Availability of data and materials All information is specified in the manuscript or included as Additional Files.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Abe A, Kosugi S, Yoshida K et al (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. Nat Biotechnol 30:174–178

Agarwal G, Clevenger J, Pandey MK et al (2018) High-density genetic map using whole-genome resequencing for fine mapping and candidate gene discovery for disease resistance in peanut. Plant Biotechnol J 16:1954–1967

Bertioli DJ, Jenkins J, Clevenger J et al (2019) The genome sequence of segmental allotetraploid peanut Arachis hypogaea. Nat Genet 51:877–884

Branch W (2011) First 100 years-inheritance of testa color in peanut (Arachis hypogaea L.). Crop Sci 51:1–4

Chen H, Chen X, Xu R et al (2021) Fine-mapping and gene candidate analysis for AhRt1, a major dominant locus responsible for testa color in cultivated peanut. Theor Appl Genet 134(11):3721–3730

Chen X, Lu Q, Liu H et al (2019) Sequencing of cultivated peanut, Arachis hypogaea, yields insights into genome evolution and improvement. Mol Plant 12:920–934

Chen Y, Chen Y, Shi C et al (2018) SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. Gigascience 7:1–6

Clevenger J, Chu Y, Chavarro C et al (2017) Genome-wide SNP genotyping resolves signatures of selection and tetrasomic recombination in peanut. Mol Plant 10:309–322

Clevenger J, Chu Y, Chavarro C et al (2018a) Mapping late leaf spot resistance in peanut (Arachis hypogaea) using QTL-seq reveals markers for marker-assisted selection. Front in Plant Sci 9:83

Clevenger JP, Korani W, Ozias-Akins P, Jackson S (2018b) Haplotype-based genotyping in polyploids. Front Plant Sci 9:564

Fekih R, Takagi H, Tamiru M et al (2013) MutMap+: genetic map and mutant identification without crossing in rice. PLoS One 8:e68529

Fischer TC, Mirbeth B, Rentsch J, Sutter C, Ring L, Flachowsky H, Habegger R, Hoffmann T, Hanke MV, Schwab W (2014) Premature and ectopic anthocyanin formation by silencing of anthocyanidin reductase in strawberry (Fragaria ananassa). New Phytol 201:440–451

Table 4 SNPs in putative candidate genes in the mapping genomic region for testa color on Chr.12

| SNP Position  | aΔSNP -index | bED | cRef | dAlt | eAl   | Genes     | Gene annotation                  |
|--------------|--------------|-----|------|------|------|-----------|----------------------------------|
| 117,049,317  | 0.94         | 1.33| C    | G    |      | Exon      | Arahy.B3SD72 Uncharacterized protein |
| 117,191,491  | –            | 1.28| G    | A    | 5’UTR| Arahy.IK60LM Anthocyanidin reductase |
| 117,190,528  | 0.95         | 1.35| G    | A    | Exon | Arahy.IK60LM Anthocyanidin reductase |
| 117,328,756  | 0.89         | 1.27| A    | G    | 3’UTR| Arahy.VXEX44 Diaminopimelate aminotransferase |
| 117,420,658  | 0.92         | 1.29| A    | G    | 3’UTR| Arahy.Z3ANCQ Protein trichome birefringence-like |
| 117,444,731  | 0.9          | 1.27| A    | G    | Intron| Arahy.LX501Y NBS-LRR disease resistance protein |
| 117,444,733  | 0.9          | 1.27| A    | C    | Intron| Arahy.LX501Y NBS-LRR disease resistance protein |
| 117,504,480  | 0.92         | 1.30| T    | C    | Exon | Arahy.IL886L NBS-LRR disease resistance protein |

aΔSNP-index: the subtraction of the SNP-index of Pink-pool and Red-pool
bED: Euclidean Distance
cRef: the SNP genotype in the reference genome
dAlt: the genotype in the non-reference genome
eAl: the functional location of identified SNP in annotated candidate gene
