Genome-Wide Association Study of *Actinidia Chinensis*

Yu Zhang (✉ yuzhan20160315@outlook.com)

Shaanxi University of Technology

Yuexing Wang

Xinjiang Agricultural University

Shimao Zheng

Shaanxi University of Technology

Wanying Zhou

Shaanxi University of Technology

Wenhui Zhang

Ankang municipality agricultural science research institute

Shengbao Wang

Hanzhong institute of agricultural science

Dong Qu

Shaanxi University of Technology

Research article

Keywords: Actinidia chinensis, Genotyping-by-Sequencing, SNPs, GWAS, post-GWAS, GO, KEGG

Posted Date: December 10th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-120590/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: A total of 74,936 SNPs were employed to carry out GWAS and post-GWAS of the fifty-six accessions, representing the two most valuable varieties of *Actinidia chinensis*, namely, *A. chinensis* var. *chinensis* and *A. chinensis* var. *deliciosa*, are mainly distributed in China. The percentage of heterozygous sites of *A. chinensis* var. *deliciosa* is higher than that of *A. chinensis* var. *chinensis*, which could be one of the reasons for *A. chinensis* var. *deliciosa* high disease resistance. LD decay distance of male plants is shorter than that of female plants.

Results: Fifty-six accessions were divided into two subgroups. Analysis of molecular variance shows that the frequency of genetic variations within the population is 83.53% and 16.47% between populations. F_{st} between the two populations is 0.14, and N_m is 1.60. Set at $\alpha \leq 0.05$, a total of 327 SNPs and 260 haplotypes were related to the hispidus. A total of 246 proteins were annotated using GO and KEGG analyses, which indicated the membrane-related genes and stress-resistant metabolic pathways are related to the hispidus condition of leaves, stems, and peels of kiwifruit, which is also the result of the adaptation of *A. chinensis* var. *deliciosa* to its growth environment.

Conclusions: Haplotype analysis showed that the evolution of is *A. chinensis* var. *deliciosa* later than that of *A. chinensis* var. *chinensis*.

Background

Kiwifruit belongs to the Actinidiaceae *Actinidia* Lindl. species. There are no reproductive differences between multiple subspecies under *Actinidia* Lindl. that provides a physiological basis for gene penetration between these subspecies, especially a cross between different varieties of *A. chinensis*, which produce a series of intermediate types, coupled with complex chromosome ploidy, and making the classification of some species and variants unclear, ultimately resulting in multiple classification methods (Peng and Zhang 1995; Zhao and Liu 1996; Huang 2013; Wang et al. 2018). It is difficult to distinguish kiwifruit varieties on the mere basis of morphological characteristics, particularly at the seedling stage in kiwifruit.

China is the original center and predominant kiwifruit-producing area (Huang 2009). Except for kiwifruit from Nepal, Vietnam, and Japan, the rest are concentrated in the Qinba region and east of the Hengduan Mountains in China. The two most valuable varieties of *Actinidia chinensis*, namely, *A. chinensis* var. *chinensis* and *A. chinensis* var. *deliciosa*, are mainly distributed in the Qinba region. Soft or hard, as well as bearing more or less hispidus on the leaves, stems, and peels of the kiwifruit, is the main morphological character that distinguishes the two varieties. The *A. chinensis* var. *chinensis* is often distributed in areas below 1,000 m above sea level, the *A. chinensis* var. *deliciosa* is highly resistant to cold, is easy to store, and has a wider range of planting.

Gene flow caused by crosses is one of the important routes in generating varieties and in adapting to the environment. Before the formation of varieties, individuals within the population can randomly mate, so there are numerous shared haplotypes among individuals within the population. If gene flow occurs during the process of population differentiation or secondary contact, then this will also be reflected in the length of shared haplotypes. With the occurrence of recombination, the length of haplotypes shared between populations gradually decreases. Shared haplotype analysis can provide a reference for the evolution of *A. chinensis*.

Currently, basic research on kiwifruit remains limited, particularly at the molecular basis as most of the widely cultivated kiwifruit come from conventional breeding technology such as the seedling of wild kiwifruit, as is the case with the “Hayward”, the world's highest selling variety in the 1980s. Our research also shows that the most cultivated kiwifruit varieties have the highest homogeneity. So far, research studies have focused on the genetic diversity of kiwifruit germplasm resources by AFLP, SSR, and SCoT markers in different regions. These research studies show that wild germplasm resources possess an abundant genetic information (Zhang et al. 2018; Wang et al. 2018; Chen et al. 2018; Shen et al. 2017; Zhang et al. 2017; Zhong et al. 2018; Liu et al. 2006; Tang et al. 2014; Zhong et al. 2019). SSR markers are the most commonly used molecular markers. The profound study of kiwifruit in relation to biological problems is hindered by the fact that SSR markers have an imbalanced distribution in the genome, a weak electrophoretic resolution, and is time-consuming and labor-intensive. Germplasm resource identification, phylogeny, gene cloning, and molecular marker-assisted selection of kiwifruit have not been reported to date. Hale et al. (2018) identified sex-linked molecular markers of natural populations of *A. arguta* and *A. kolomikta* in the USA by Genome-wide association study (GWAS), and we verified that these were not suitable for gender determination in *A. chinensis* var. *chinensis* and *A. chinensis* var. *deliciosa*. Takashi et al. (2018, 2019) identified possible sex-chromosome and male-specific fragments of *A. chinensis* var. *chinensis* by GWAS. Based on 74,936 single nucleotide polymorphism (SNPs) generated by genotyping-by-sequencing (GBS) in 56 samples (36 *A. chinensis* var. *deliciosa*, 18 *A. chinensis* var. *chinensis*, and 2 *A. polygama* in the light of morphological observations), GWAS and post-GWAS were conducted in this study to provide theoretical support for kiwifruit-

related research, particularly in terms of a multigene pyramiding breeding technology centering on molecular marker-assisted selection that can select genotypes of quality production traits at the DNA level. Such a method circumvents the disadvantage of low accuracy of conventional breeding technology improves breeding efficiency.

Results

Mapping rate and heterozygous proportion

Kiwifruit is a functional dioecious plant. Currently, only the genome of the female plant, named, red 5, has been released in 2018 (ncbi.nlm.nih.gov/genome/genomes/16401). Therefore, the sequence of kiwifruit can only be aligned with red 5. The mapping rate (number of reads on the reference genome divided by the number of valid sequencing data) and the heterozygous proportion can roughly reflect the classification of the sample. The mapping rate of 56 kiwifruits ranged from 68.08% to 97.30% (average: 90.49%). Among these, the mapping rate of the *A. chinensis* var. *chinensis* increased from 91.31% to 97.30%, with an average of 95.57%. The mapping rate of *A. chinensis* var. *deliciosa* increased from 68.08% to 89.85%, with an average of 87.96%, which was lower than that of *A. chinensis* var. *chinensis*. The mapping rates of the two plants that were assessed morphologically observations, *A. polygama*, were 89.94% and 90.51%, respectively. *A. polygama* may be an intermediate type between *A. chinensis* var. *chinensis* and *A. chinensis* var. *deliciosa*. The average mapping rate of the female plant is 91.60%, and the mapping rate of the male plant is 89.39%. Therefore, the difference at the genome level between male and female plant is less than that between subspecies of *A. chinensis*.

The heterozygous proportion ranged from 45.02% to 12.03%. The mapping rate was negatively correlated with the heterozygous proportion. *A. chinensis* var. *chinensis* showed a high mapping rate and a low heterozygous promotion. However, *A. chinensis* var. *deliciosa* showed a low mapping rate and a high heterozygous proportion (Fig. 1).

LD and haplotype blocks analysis

Approximately 74,936 SNPs were unevenly distributed among 29 chromosomes (Fig.2a); chromosome 23 contained the highest number of markers (4,645), whereas chromosome 17 had the lowest (1,824). SNPs are infrequently inherited independently and thus have a propensity to be inherited in clusters, thereby resulting in linkage disequilibrium (LD) as well as haplotypes. Haplotype blocks are expected in *A. chinensis* genomes to be non-randomly distributed (Fig.2b). The 29 chromosomes yielded a total of 4,379 predicted haplotype blocks; chromosome 23 possessed the highest number of haplotype blocks (303), whereas chromosome 17 had the least (108). The largest haplotype blocks were composed of 18 SNPs, and the average number of SNPs within all haplotype blocks was 3. Whether at a single chromosome (Fig. 3a) or genome-wide level (Fig. 3b), LD shows that as the physical distance increases, the smaller becomes r^2 between two SNP sites or even null. For different samples, LD also decreases with increasing distance between the two SNP sites (Table 1). LD in the male plant had significantly higher decay than that of the female plant, and the LD of *A. chinensis* var. *deliciosa* had significantly higher decay than *A. chinensis* var. *chinensis*. The longest haplotype blocks span 194.74 kb, the average length on haplotype blocks was 3.07 kb. *A. chinensis* var. *chinensis* is shorter than *A. chinensis* var. *deliciosa* in terms of average span of haplotypes, this shows that the evolution of *A. chinensis* var. *chinensis* is later than that of *A. chinensis* var. *deliciosa*.

Population structure

Clustering of population structure was performed by principal component analysis (PCA), the unweighted pair-group method with arithmetic means (UPGMA), and Bayesian clustering algorithm.

PC analysis

In PCA, the first three principal components accounted for 18.45%, 11.35%, and 5.02% of the total variation, respectively. The results of clustering based on the first three PCs were initially separated by varieties and intermediate types of material. *A. polygama* (No. 12 and 38) and *A. chinensis* var. *chinensis* were clustered together, and the yellow pulp kiwifruit (No. 47), a variant of *A. chinensis* var. *chinensis*, was separated from the other samples (Fig. 4).

UPGMA

The phylogenetic tree can be divided into two subgroups consisting of 45 (subgroup I) and 11 (subgroup II) accessions. The results of primary clustering and morphological observations are not completely consistent (red represents *A. chinensis* var. *chinensis*, green represents *A. chinensis* var. *deliciosa*, and yellow-green represents *A. polygama*) (Fig. 5). In subgroup I, the 45 accessions, including *A. chinensis* var. *Chinensis* and *A. chinensis* var. *deliciosa*, can be divided into several sub-classes at different similarity coefficients levels,

in which 20 accessions that were genetically related were gathered together [including two *A. polygama* plants and one yellow pulp kiwifruit (asterisks)]. In subgroup II, all accessions belong to *A. chinensis* var. *deliciosa*. Based on morphological observations, those may have occurred earlier in evolution than *A. chinensis* var. *deliciosa* in subgroup I.

Bayesian clustering

We set K from 1 to 5 using previous population information. Delta K reached a maximum value at K=2 (Fig. 6a), suggesting that the 56 accessions were divided into two subgroups (consisting of 19 and 37 accessions) (Fig. 6b). In particular, the results of No. 47 sample are inconsistent among the three clustering methods. In the population structure analysis, the results from K = 2 to K = 5 revealed the occurrence of gene introgression between *A. chinensis* var. *Chinensis* and *A. chinensis* var. *deliciosa*, accounting for approximately 57.14% of the observed variations (calculated by K = 2).

Introgression analysis

The present study determined that gene flow (Nm) was 1.60. The Nm in animal-pollinated plants is 0.801 and 5.380 in wind-pollinated plants (Chen 1996). Kiwifruit is a typical insect-pollinated species under natural conditions. However, artificial pollination is often used in cultivation to improve quality and production. If a certain region of the genome is introduced into a population by gene flow, then it will continuously exist in the population, indicating that the region may be subject to positive selection or balanced selection; the degree of LD in these regions increases, and the length of the shared haplotype will be shorter. The average haplotype length of the 56 samples is 3.07 kb. The average proportion identity-by-descent (IBD) was 0.49, which ranged from 0.17 to 0.99. Otherwise, this area will be removed from this population by negative selection or genetic drift. Gene flow within *A. chinensis* species is one of the important ways to meet the market demand for this crop and to adapt to the environment.

Analysis of molecular variance

AMOVA indicated that 16.47% of the total variation was due to differences among populations, and 83.53% of the variation was due to a difference within populations (Table 2).

Genotype-phenotype association analysis

No value can be correctly associated with soft or hard hispidus as well as to more or less hispidus on the leaves, stems, and peels of kiwifruit, but a phenotypic value of 1 or 2 was assigned to the hispidus associated study between *A. chinensis* var. *deliciosa* and *A. chinensis* var. *chinensis*. The threshold of association test was set at $\alpha \leq 0.05$, that is, $P \leq 0.05 / 74936 = 6.67E-7$.

a) Single SNP association test

327 loci were associated with hispidus condition (Table S1). The strongest signal came from the SNP at position 15,347,013 on Chr. 16, followed by those from markers on Chrs. 8 and 16 (Fig.7).

b) Haplotype association test

260 haplotype blocks were associated with hispidus condition. One signal from Chr. 16 represented the most significant association. Followed by those from markers on Chrs. 8 and 16. (Table S2). The results of gene mapping for hispidus in *Actinidia chinensis* based on haplotype association test are shown in Figure 8.

GO-enrichment and KEGG pathway

GO function analysis showed (Fig.9) that differential genes related to cell components significantly enriched in integral components of the membrane, nucleus, membrane, and chloroplast. Differential genes involved in molecular functions significantly enriched the categories of ATP binding, metal ion binding, DNA binding, and protein kinase activity.

To further understand the biological pathways in which the differential genes are mainly involved, KEGG pathway enrichment analysis of the identified related genes was performed (Fig.10). The results showed that the related genes belong to 105 metabolic pathways. The first three types of pathways include metabolic, biosynthesis of secondary metabolites, and biosynthesis of antibiotics.

Discussion

In the past few decades, SSR technology has been widely used in related fields of bioscience. Since the first report of the application of genome-wide association study based on SNPs named “complement factor H polymorphism in age-related macular degeneration” (Robert 2005). with the development of next- generation sequencing technology, SNPs has many advantages in the fields of genotyping, phyletic evolution, diversity, linkage mapping, QTL mapping and so on. In order to improve the cost-effectiveness of sequencing, the simplified genomic sequencing technology based on the restriction site-associated DNA sequencing (RAD tags) has been widely used (Zhang 2020a). Among these, at present, GBS technology has become a main technology to simplify sequencing for simple process and using methylation sensitive enzymes avoiding the genomic repeat region. It is especially suitable for large sample size research, which can quickly identify high-density mutations and accelerate biological research.

GWAS is based on population structure analysis. Many studies have shown that complex population structure will lead to false positive between genotype and phenotype (Price 2010). There are numerous intermediate types in nature within *A. chinensis* due to gene flow, and if classification is based on morphological characteristics alone, then specific limitations have to be considered. In this study, different clustering methods by PC, UPGMA, and Bayesian algorithm generated different results. Basically, the 56 accessions were divided into two groups; however, the genetic distance between the two groups was not significant, indicating that the samples have a low level of genetic stratification. It should be noted that the two samples identified as *A. polygama* by phenotypic observation were similar to those of *A. chinensis* var. *chinensis*, whether it's mapping rate, heterozygous proportion or clustering results, indicating that the hispidus on the leaves, stems, and peels of the kiwifruit plays an important role in the classification of kiwifruit.

Haplotype analysis has been widely used in the genetic evolution of biological resources (Wang 2017; Xu 2018). Haplotype is likely to play a role in the process of genetic evolution as a functional unit. In this study, haplotype analysis showed that the evolution of *A. chinensis* var. *chinensis* is later than that of *A. chinensis* var. *deliciosa*, which is different from Huang (Huang 2013). The loss of the hispidus of *A. chinensis* var. *chinensis* may be the result of low altitude cultivation, which also leads to the decrease of resistance to low temperature and disease. Haplotypes also play an important role in genes mapping related to complex diseases as well as important traits in animals and plants, and provide important systematic information for selection, matching and breeding of quantitative traits. In the associated analysis based on single SNP test and haplotype test, the peak SNP is often appeared within haplotype. According to the previous haplotype test of genome-wide association study, we can see that the haplotype implies gene clusters related to traits (Zhang 2020b).

Statistical analysis of association may lead to some wrong decisions when strategies that reduce data analysis or scientific reasoning to mechanical standards are used alone; for example, $\alpha < 0.05$. So far, no report has been found about the combination of GWAS and post-GWAS analysis. It is necessary to perform post-GWAS such as GO and KEGG to better understand the pathways related to the development of traits. Thus, we merged the results of the SNP and the haplotype tests, a total of 327 SNPs and 260 haplotypes were related to the hispidus. A total of 246 proteins were annotated using GO, the related genes are mainly enriched in molecular functions and cell components, but less in biological processes. KEGG analyses showed that the membrane-related genes and stress-resistant metabolic pathways are related to the hispidus condition of leaves, stems, and peels of kiwifruit, which is also the result of the adaptation of *A. chinensis* var. *deliciosa* to its growth environment. Meanwhile, hispidus is simultaneously accompanied by changes in leaf color, i.e., the more hispidus on the leaves, the lighter the shade of green on the leaves. The results provide a genetic basis for the evolution of *A. chinensis* and the formation of hispidus.

Materials And Methods

Plant Materials

A total of 56 accessions, representing most of the *A. chinensis* germplasm resource of the Qinba area in China, were collected from the kiwifruit experimental farm of different regions during the 2019 growing season (Table 3) and comprised 18 accessions of *A. chinensis* var. *chinensis*, 36 accessions of *A. chinensis* var. *deliciosa*, and 2 accession of *A. polygama*, which were established based on morphological observations.

Genotype data

Genotyping-by-sequencing (*Mse*I digestion) was used to obtain SNP data for this panel of 56 accessions using the Illumina HiSeq2000 sequencing platform. A total of 74,936 SNPs passed the MAF lower limit of 0.05.

LD and haplotype analysis

Genotype data were then used to calculate LD between SNPs and to construct haplotypes using the EM algorithm implemented by PLINK1.90 (<https://www.cog-genomics.org/plink2>). The command line “-r²” and “-blocks” were used to calculate LD and assign SNPs to their respective haplotypes by calculating inter-maker LD within a 200kb window, respectively.

Population structure

NTSYS-pc2.10e (Rohlf 2000) was used for principal components analysis, Identity-by-state (IBS) distance matrix generated by TASSEL5.0 (<http://www.maizegenetics.net/tassel>), building a UPGMA tree, MEGA5.0 (<http://http://www.megasoftware.net/>) editing and visualizing. The Bayesian clustering algorithm implemented in STRUCTURE 2.3.4 (<http://taylor0.biology.ucla.edu/structureHarvesteroybase.org/tools.php>) was used to simulate the population's genetic structure. To obtain an estimate of the most probable number of population (K), K values from 1 to 5 were simulated with 10 iterations for each K, using 10 000 burn-in periods followed by 10 000 Markov Chain Monte Carlo iterations. Delta K was plotted against K values and the best number of clusters was determined following the method proposed by Evanno et al (Evanno et al. 2005), and the population structure diagram was obtained via the Structure Harvester platform (Earl et al. 2012).

AMOVA and introgression analysis

On the basis of mapping rate and population structure analysis, for analysis of molecular variance (AMOVA), 56 accessions were classified into two groups, namely, *A. chinensis* var. *chinensis* and *A. chinensis* var. *deliciosa*. The components of variance attributable to different varieties were estimated from the genetic distance matrix. The Tajima & Nei method was used to estimate molecular distance. The distance method used for locus-by-locus analysis was pairwise difference. As specified in the AMOVA procedure, ARLEQUIN 3.1 (Excoffier et al. 2005) with 10,000 permutations was used to assess the significance of variance in components associated with the different possible levels of genetic structure in this study. The pairwise F_{st} value, a value of F statistic analogs computed from AMOVA, was used to compare genetic distances between the two groups. The matrix of significant F_{st} is 0.05, and N_m was calculated using the following equation: $N_m = 1 - F_{st}/4F_{st}$.

Association analysis

Single SNP association with the allelic association tests were run in Plink1.90, Haplotype association tests were performed with haplotype-based C/C tests using Plink1.07.

GO-enrichment and KEGG analysis

We integrated the results of the commonly detected loci/regions, including the single SNP association tests and haplotype association tests. A total of 246 genes were assigned, and functional annotation was performed using KEGG pathway analysis in KEGG Mapper (<https://www.genome.jp/kegg/mapper.html>).

Declarations

Ethics approval: Not applicable.

Consent to participate: All authors declare that we agreed to participate in the publication.

Consent for publication: All authors declare that we agreed to publish the paper in [BMC Plant Biology](#).

Availability of data and material (data transparency)

Conflicts of interest/Competing interests: The authors declare that they have no conflict of interest.

Funding: 1.Collection and Evaluation of Wild Germplasm Resources and Breeding of New Varieties for Kiwifruit in Qinba 2018SZZ-27-0322. 2. Project of Education Department of Shaanxi: Effects of Potassium on Kiwifruit canker resistance 2018SZZ-27-0322. 3. Project of Education Department of Shaanxi: Province "Mechanisms of Exogenous Salicylic Acid on inducing Kiwifruit canker resistance" 2017JS0222.

Authors' contributions: Yu Zhang analyzed the data, and wrote the manuscript. Wanying Zhou, Yuexing Wang plot the data after analysis. Wenhui Zhang, Shengbao Wang and Dong Qu edited and revised the manuscript. All authors read and approved the paper.

Acknowledgments: This study was supported by Project of Shaanxi Provincial Department of Science and Technology "Collection and Evaluation of Wild Germplasm Resources and Breeding of New Varieties for Kiwifruit in Qinba" [K3370219025], "Effects of Potassium on Kiwifruit canker resistance" [2018SZS-27-03] and Project of Education Department of Shaanxi Province "Mechanisms of Exogenous Salicylic Acid on inducing Kiwifruit canker resistance" [17JS022].

References

1. Chen BL, Zhang J, Huang JK, Huang CM, Zhou QJ, Tao W, Tang J. Application of SCoT markers on genetic diversity analysis and variation identification of Actinidia. *Journal of Agricultural Biotechnology*. 2018; 26: 77-86
2. Chen XY. Gene flow of plants and its role in the conservation of endangered plants. *Biodiversity Science*. 1996; 2: 97-102
3. Earl DA, Vonholdt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour*. 2012; 4:359-361
4. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *MOL ECOL*. 2005; 14:2611-2620
5. Excoffier L, Laval G, Schneider S. Arlequin version 3.0: An integrated software package for population genetics data analysis. *Evol Bioinform online*. 2005; 1:47-50
6. Hale I, Melo ATQ, Gustafson H. Sex-linked molecular markers for two cold-hardy kiwifruit species, *Actinidia arguta* and *A. kolomikta*. *Eur J Hortic Sci*. 2018; 4:236–246
7. Huang HW. History of 100 years of domestication and improvement of kiwifruit and gene discovery from genetic introgressed populations in the wild. *Chinese Bulletin of Botany*. 2009; 2:127-142
8. Huang HW. *Actinidia germplasm resources in China*. China: China Forestry Press; 2013
9. Liu YL. Spatial genetic structure in natural populations of two closely related *Actinidia* species (Actinidiaceae) as revealed by SSR analysis. *Biodiversity Science*. 2006; 5:421-434
10. Peng YH, Zhang WC. Studies on the Selection of Varieties and Region in Kiwifruit Culture in Changjiang River Area. *Scientia Agricultura Sinica*. 1995; 3:14-20
11. Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. *Nat Rev Genet*. 2010; 7:459-463
12. Racimo F, Sankararaman S, Nielsen R, Huerta SE. Evidence for archaic adaptive introgression in humans. *Nat Rev Genet*. 2015; 6:359-371
13. Robert JK, Caroline Z, Emily YC, Jen-Yue T, Richard SS, Chad H, Alice KH, John PS, Shrikant MM, Susan TM, Michael BB, Frederick LF, Jurg O, Colin B, Josephine H. Complement Factor H Polymorphism in Age-Related Macular Degeneration. *Science*. 2005; 5720:385–389
14. Rohlf F. NTSYS-pc - Numerical Taxonomy and Multivariate Analysis System, 2000; version 2.1. New York: Exeter Software
15. Shen GZ. SSR analysis of kiwifruit of Hongyang and some wild kiwifruit in Hangzhou. *Journal of Zhejiang Agricultural Sciences*. 2017; 11:1906-1909
16. Akagi T, Henry IM, Ohtani H, Morimoto T, Beppu K, Kataoka I, Tao R. A Y-Encoded Suppressor of Feminization Arose via Lineage-specific Duplication of a Cytokinin Response Regulator in Kiwifruit. *Plant Cell Advance Publication*, 2016; published on April 6
17. Akagi T. Two Y chromosome-encoded genes determine sex in kiwifruit. *Nature Plants*; 2019
18. Tang JL, Wu H, Lang BB, Qu XY, Huang CH, Xu XB. Association Analysis on Leaf and Fruit ASA Content and SSR Markers of Wild *Actinidia eriantha*. *Acta Horticulturae Sinica*. 2014; 5:833–840
19. Wang SY, Zhong M, Huang CH, Tao JJ, Qu XY, Wu H, Xu XB. Fingerprints Construction of SSR Markers of Thirty-three Core Male Plants in *Actinidia*. *Molecular Plant Breeding*. 2018; 5:1598-1606
20. Wang YG, Yang J, Chen JK. Potential value, status analysis, and protection strategy of wild kiwifruit genetic resources in the Yangtze River Basin. *Biodiversity Science*. 2018; 4: 373–383
21. Wang YJ. Study on Genetic Evolution of Kam Sweet Rice and Influence from Dong's Traditional Culture. *Scientia Agricultura Sinica*. 2017
22. Xu DD. Haplotypes and Evolution Analysis Ear Trait Domestication Genes. *Scientia Agricultura Sinica*. 2018
23. Zhang AS, Zhang ZH, Qi XJ, Fang JB, Lin MM, Luo Y. Genetic Diversity Analysis and Fingerprints Construction of SC-SSR Markers in Kiwifruit (*Actinidia* Lind L.). *Molecular Plant Breeding*. 2017; 11:4706-4714

24. Zhang H, ZHANG SX, WU SH, TIAN WM, PENG XL, LIU SB. Genetic diversity of 33 kiwifruit germplasms based on AFLP markers. *Journal of Biology*.2018; 2:29-33
25. Zhang XN, He X, Liu XZ, Wei Z, Zhang Y, Zhang HY. AFLP Analysis of Open-Pollinated Progenies of 'Hort 16A' Kiwi Fruit. *Molecular Plant Breeding*.2018; 11:3615-3621
26. Zhang Y, François B. Genome-wide Association Study for Sclerotinia sclerotiorum Resistance of Soybean. *Acta Agriculturae Boreali-Sinica*.2020a; 1:205-213
27. Zhang Y, Zhou WY, Sun W. The Two Major Technologies of Sequencing Based on Simplified Genome By Restriction Enzyme Digestion. *Molecular Plant Breeding*.2020b; 11:3562-3570
28. Zhao YL, Liu XP. Study on the Geographical Distribution and Floristic Characteristics of the Genus *Actinidia* Lind L. From Hunan. *Guihaia*.1996; 1:9-16
29. Zhong M, LIAO GL, LI ZY, ZOU LF, HUANG Q, CHEN L, HUANG CH, TAO JJ, ZHU B, XU XB. Genetic diversity of wild male kiwifruit (*Actinidia eriantha* Benth) germplasms based on SSR and morphological markers. *Journal of Fruit Science*.2018; 6:658-667
30. Zhong M, Tao JJ, Huang CH, Huang Q, Zou LF, Liao GL, Chen L, Xu XB. Analysis of genetic diversity of populations in *Actinidia eriantha* Benth based on Simple Sequence Repeat (SSR) markers. *Journal of Nuclear Agricultural Sciences*.2019; 5:0863-0869

Tables

Table 1. LD and haplotype analysis

Sample	Average distance between SNP(Kb)			Haplotype			
	$R^2 \geq 0.2$	$R^2 \geq 0.8$	$R^2 = 1$	Number of haplotypes	Average span(Kb)	Longest span(Kb)	Number of SNPs for the longest haplotype
mixed(56)	30.29	8.61	5.29	4379	3.07	194.742	18
<i>A. chinensis</i> var. <i>chinensis</i> (20)	73.21	46.49	45.09	506	3.99	125.932	13
<i>A. chinensis</i> var. <i>deliciosa</i> (36)	34.01	6.78	5.14	3667	7.42	198.199	28
♀(25)	32.68	18.52	17.98	20	2.43	24.915	9
♂(26)	31.63	15.13	14.31	1898	4.15	195.319	25

Table 2. AMOVA design and results (average over 74936 loci)

Source of variation	Sum of squares	Variance components	Percentage of variation
Among populations	73740.05	1691.85	16.47
Within populations	744697.04	8583.32	83.53
Total	818437.09	10275.17	

Note: Fixation Indices F_{st} : 0.16; Significance Level = 0.05; Significance tests (1023 permutations)

Table 3. List of 56 accessions used in this study

Sample	Taxa	gender	Mapping rate%	Proportion Heterozygous%	sample	Taxa	gender	Mapping rate%	Proportion Heterozygous%
1	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	87.88	34.38	29	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.65	42.62
2	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	97	12.11	30	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.55	45.02
3	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	93.23	18.19	31	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.85	34.73
4	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	97.04	12.03	32	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.43	38.52
5	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	93.25	20.15	33	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.52	37.72
6	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.51	41.69	34	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.23	37.60
7	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.28	42.59	35	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.61	38.80
8	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.94	43.69	36	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.3	33.91
9	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.47	44.40	37	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.96	34.61
10	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.73	34.56	38	<i>A.</i> <i>Polygama</i>	♀	89.94	18.37
11	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.79	38.26	39	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.93	43.44
12	<i>A.</i> <i>Polygama</i>	♀	90.51	18.89	40	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	68.08	37.47
13	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.21	34.64	41	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	87.2	34.83
14	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.3	35.13	42	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	83.96	34.41
15	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.17	34.27	43	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	85.07	38.54
16	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	87.8	33.66	44	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	97.09	12.17
17	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	91.31	22.75	45	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	97.12	16.00
18	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.32	36.82	46	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.66	37.83
19	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	97.1	15.83	47	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	92.02	38.17
20	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.25	34.13	48	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	87.94	36.76
21	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	89.19	34.72	49	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	88.89	37.99
22	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	♀	96.39	16.31	50	<i>A.chinensis</i> <i>var.deliciosa</i>	♀	86.93	32.95

23	<i>A.chinensis</i> <i>var.deliciosa</i>	□	88.75	43.47	51	<i>A.chinensis</i> <i>var.deliciosa</i>	□	88.88	42.44
24	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	□	96.58	16.49	52	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	□	97.3	14.53
25	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	□	96.75	12.57	53	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	□	93.56	19.55
26	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	□	93.39	19.26	54	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	□	96.94	14.64
27	<i>A.chinensis</i> <i>var.deliciosa</i>	□	88.93	42.52	55	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	□	97.29	15.57
28	<i>A.chinensis</i> <i>var.deliciosa</i>	□	88.6	43.07	56	<i>A.chinensis</i> <i>var.</i> <i>chinensis</i>	□	96.97	15.31

Figures

Figure 1

The relationship among mapping rate, heterozygous proportion and samples. Note:The red box indicates the sample of *A. chinensis* var. *Chinensis*.The green box indicates the sample of *A. chinensis* var. *deliciosa*.The yellow green box indicates the sample of *A. polygama*.

Figure 2

SNPs and Haplotype distribution on chromosome a. SNPs distribution on chromosome; b. Haplotype distribution on chromosome.

Figure 3

A single chromosome and genome-wide LD pattern. a. LD decay plot on chromosome 1; b. Genome-wide LD decay plot.

Figure 4

PCA plots of the first three components.

Figure 5

UPGMA tree built from IBS distance matrix.

Figure 6

Structure estimation of populations and Bayesian clustering. a. Structure estimation of populations for K ranging from 1 to 5 by delta K. b. Bayesian clustering based on 74936 SNPs for 56 plants. Note: Red: group I; Green: group II. Vertical lines on the X-axis refer to each sample. The proportion of each color represents probability rate with which a given genotype belongs to each group. The cluster of yellow pulp kiwifruit gives different results using various clustering methods. PA analysis revealed that yellow pulp is a distinct characteristic of kiwifruit.

Figure 7

Manhattan plot of the association using the allelic association tests. Note: Q-values (adjusted p-value) for the association of each single nucleotide polymorphism with trait are shown on the y-axis. The SNPs are plotted on the x-axis according to their chromosomal location. Blue line: $\alpha=0.05$; red line: $\alpha=0.01$

Figure 8

Prioritizing candidate haplotype

Figure 9

GO functional enrichment analysis.

Figure 10

KEGG pathways enrichment analysis.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [9.20kiwifruitSupplementarydata.doc](#)