Genome-wide association study of the candidate genes for grape berry shape-related traits

Chuan Zhang1, Liwen Cui2 and Jinggui Fang1*

Abstract

Background: In the breeding of new horticultural crops, fruit shape is an important selection characteristic. A variety of fruit shapes appeared during the gradual process of selection and domestication. However, few studies have been conducted on grape berry shape, especially studies related to mining candidate genes. To discover candidate genes related to grape berry shape, the present study first took the berry shape parameters analyzed by Tomato Analyzer as the target traits and used a genome-wide association analysis to analyze candidate genes.

Results: In total, 122 single-nucleotide polymorphism (SNP) loci had significant correlations with multiple berry shape traits in both years, and some candidate genes were further mined. These genes were mainly related to LRR receptor-like serine/threonine-protein kinase (At1g05700 and At1g07650), transcription factors (GATA transcription factor 23-like, transcription factor VIP1, transcription initiation factor TFIID, and MADS-box transcription factor 6), ubiquitin ligases (F-box protein SKIP19 and RING finger protein 44), and plant hormones (indole-3-acetic acid-amido synthetase GH3.6 and ethylene-responsive transcription factor ERF061). In addition, some important SNP loci were associated with multiple berry-shape traits. The study further revealed some genes that control multiple traits simultaneously, indicating that these berry shape traits are subject to the coordinated regulation of some genes in controlling berry shape.

Conclusions: In the present work, we identified interesting genetic determinants of grape berry shape-related traits. The identification of molecular markers that are closely related to these berry-shape traits is of great significance for breeding specific berry-shaped grape varieties.

Keywords: Grape berry, Berry shape, Tomato analyzer, SNP, Candidate gene
and fruit transportation [6]. The shape of the tomato fruit determines its culinary use (fresh, sliced, diced, processed, or cooked) and its market value [5]. Flat tomato fruits are popular in homes and restaurants, whereas slender and slightly blocky fruits are easier to harvest and machine process than round fruits and therefore are favored by the processing industry [3].

The classification of fruit shape is the premise of studying the genetic mechanism of fruit shape in horticultural crops. Generally, tomato varieties are correctly classified according to the fruit morphology described by the International Union for the Protection of New Plant Varieties (UPOV) and the International Plant Genetic Resources Institute (IPGRI) [9, 10]. According to the UPOV and IPGRI classification standards for tomato varieties, based on the analysis of fruit shape using analysis software, the tomato variety shapes are divided into eight categories: flat, round, heart-shaped, bull heart shape, long, rectangular, obovate and ellipse [5, 11]. In addition, the shapes of other fruits (including cherry, eggplant and watermelon) have also been classified [12–14]. However, few reports have been provided the classification of grape berry shape.

Given the importance of fruits with different shapes, researchers have conducted extensive research on fruit shapes, and a series of advances have been made [5]. These studies have mainly included morphology and genetics [15–17]. From the perspective of morphology, related studies have shown that mature fruit morphology is highly correlated with the ovary, and fruit morphology can be determined before ovary pollination [16, 17]. That is to say, the structure and morphology of the fruit are determined during flower development [15]. From the perspective of genetics, fruit shape is a complex trait controlled by multiple genes through different pathways [18]. Several genes controlling tomato fruit shape have been cloned [15, 19]. SUL and Ovate control elongated shapes, and both FASCIATED (FAS) and LOCULE NUMBER (LC) alter locule number, which has an impact on shape [2]. The allelic distribution of SUL, Ovate, LC, and FAS genes is closely related to UPOV and IPGRI fruit classification [2].

SUL encodes a protein that is a member of the IQ67-domain (IQD) protein families, and that is a positive regulator of growth, leading to elongated fruit [20]. The mutation of SUL is the result of a gene replication event mediated by the retrotransposon Rider [21]. Overexpression of SUL results in very elongated parthenocarpic fruits in addition to twisted stems and leaf axes [22]. Further study found that, SUL changed the expression of auxin-related genes, including those involved in auxin biosynthesis, homeostasis, signal transduction, and polar transport, indicating that SUL may regulate the ovary/fruit shape by regulating the expression of auxin-related genes in the early stage of ovary formation [23]. In addition, studies have shown that SUL has no significant effect on fruit weight, and it regulates tomato fruit shape by changing the cell division mode (increasing longitudinal cell division and reducing transverse fruit cell division) and re-regulating fruit quality [22]. Ovate encodes a negative regulator of growth, which may be an inhibitor of transcription, thus reducing the length of fruit [19, 24]. The fruit regulated by the Ovate allele carries an early termination codon; this allele is presumed to be an invalid allele [23]. A mutation in FAS resulted in flattened tomatoes due to an increase in the number of ventricles that affect fruit quality [25]. Further studies have shown that the underlying gene of FAS is CLAVATA3 (CLV3) [26], and the down-regulation of this gene is caused by large insertion in the first intron (estimated to be 6–8 kb), resulting in fruits with high locule numbers [25]. In addition to tomato [19–25], more and more genes related to fruit shape have been revealed in other horticultural crops, such as watermelon [27], peach [28–31] and cucumber [32–34].

With the rapid development of molecular biology, the genetic mechanism of fruit shape has gradually been revealed [35]. Through map-based cloning, protein interaction studies, and genome editing, a common genetic mechanism for morphological diversity in fruit and other plant organs has been identified [35]. Namely, the cell division pattern during ovary development is regulated by the Ovate Family Protein (OFP) and Tonneau1 Recruiting Motif (TRM) proteins, thereby changing the final fruit shape [35]. Furthermore, research suggests that OFPs and TRMs control the shapes of fruits, tubers, vegetables and grains in domesticated plants, and that the apparent universality of this OFP-TRM module may be part of the network required for coordinated multicellular growth in all plants [35].

However, compared with other horticultural crops, few studies have examined grape berry shape, and berry shape-related gene mining has not been reported. Grape (Vitis vinifera L.) is one of the most widely cultivated fruit crops [36]. Grape berries are commercially grown for fresh fruits, juices and raisins, but are used mainly for fermentation into wine [36]. Berry development is a complex process that involves profound physiological and metabolic changes [37]. At the stage of berry ripening, further metabolic changes make the fruit edible and attractive, which promotes the spread of seeds, including changes in peel color, cell swelling and an influx of water, the softening of berries, the accumulation of sugar in the pulp, the loss of organic acids and tannic acid and volatile
growth. Unless otherwise stated, we sampled berries between 08:00 and 10:00 in the morning. Berries with the same level of maturity and no defects on the berry surface were selected for testing.

**Experimental methods**

**Classification of the grape berry shape**

Based on the UPOV (UPOV, 2001) and IPGRI (IPGRI, 1996) classification systems, we divided the berry shapes of the varieties in the present study into nine different berry shapes: flat round, heart-shaped, curve-shaped, obovoid, ovoid, elliptic, round, long elliptic and long round.

**Analysis of grape berry shape-related parameters using the tomato analyzer**

We selected five berries with essentially the same size at maturity, cut them longitudinally with a surgical blade, and photographed the samples with reference to [11]. We used the Tomato Analyzer 3.0 (The Ohio State University, Columbus, OH, USA) to determine the following indicators: perimeter, area, width mid-height, maximum width, height mid-width, maximum height, curved height, fruit shape index external I, fruit shape index external II, curved fruit shape index, proximal fruit blockiness, distal fruit blockiness, fruit shape triangle, shoulder height, proximal angle micro, proximal angle macro, proximal indention area, distal angle micro, distal angle macro, width widest pos (the ratio of the height at which the maximum width occurs to the maximum height), eccentricity, proximal eccentricity, distal eccentricity, fruit shape index internal and eccentricity area index [40].

**Whole-genome resequencing and reference genome information**

The DNA of 279 grape varieties was extracted using a plant genome DNA kit (Tiangen Biotech (Beijing) Co. Ltd., Beijing, China), and the DNA of qualified samples was sequenced with an Illumina HiSeqTM 2500 (Illumina, Inc., CA, USA) [41]. The average sequencing depth of each material was expected to be 8 × for the development of single nucleotide polymorphism (SNP) markers within the population. We used the grape genome (PN40024) as the reference genome. The grape genome (PN40024) was downloaded from: ftp://ftp.ensemblgenomes.org/pub/release-23/plants/fasta/vitis_vini-fera/dna/ [42].

**Identification of SNP markers**

Sequencing reads were compared to the reference genome by BWA software (Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK) [43], and the genome-wide SNP markers were developed by GATK software.
varieties (Table S2), the LD of SNPs in all samples was
Based on the 566,129 SNPs developed from 279 grape
Annotation of genes related to grape berry-shape traits
rected using Bonferroni’s method: $\alpha \leq 0.1$ and $\alpha \leq 0.05$
(5 - 8), accounting for 1.79–2.87% of the varieties. In addition, the number of the varieties
shapes: flat round, heart-shaped, obovoid, ovoid, curve-
shapes were relatively fewer, with 21 and 10 varieties,
and variation analysis. PCA and variation analysis were
carried out using SPSS version 16.0 (IBM, Armonk, NY,
Expression analysis of candidate genes for grape berry-shape
traits
The expression values of the candidate genes for fruit
shape traits in the pericarp (including the skin and flesh),
flower, and seed were screened from the Gene Expression
Omnibus (GEO Datasets, No.GSE36128) [51]. The loga-
GWASthe process of GWAS, individual genetic rela-
tionships and population structure are the main factors
resulting in false-positive associations. The GWAS was
based on SNPs and used TASSEL software (Cornell Uni-
versity, Ithaca, NY, USA) [48] to obtain correlation val-
ues using a compressed mixed linear model (MLM). The
formula is as follows: $Y = \alpha X + \beta Q + \mu K + e$, $\alpha$, $\beta$, $\mu$, and $e$. In the equation, $Y$ is the phenotypic trait, $X$ is the indi-
cator matrix of the genotype (fixed effect), $\alpha$ is the esti-
Figure 2, the number distribution of varie-
ties with elliptic shape was the largest at 112, account-
ing for 40.14% of the varieties, followed by the round
berry shape at 110, accounting for 39.43% of the vari-
ties. The varieties with long elliptic or long round berry
shapes were relatively fewer, with 21 and 10 varieties,
respectively, accounting for 7.53 and 3.58% of the var-
ties, respectively. The curve-shaped varieties were the
least common, with only two, accounting for 0.72% of
the investigated varieties.

Population structure and attenuation analysis of linkage disequilibrium at the population level
The population structure of the 279 samples was analyzed
using admixture software (University of California, Los
Angeles, CA, USA) [46] with the following operating
parameters: the number of subgroups (K-value) rang-
ing from 2 to 20, $K$ of iterative operations starting from 2
and the number of runs and repetitions of each time set to
10,000. According to the $K$-value with the lowest error rate
in the cross validation, the optimal number of subgroups
determined. popLDdecay software (Xi’an Jiaotong Uni-
versity, Xi’an, China) [47] was used to analyze the Linkage
Disequilibrium (LD) at the population level, and the param-
eters were set at -MAF 0.05 -MaxDist 500 -Miss 0.25.

GWAS
In the process of GWAS, individual genetic rela-
tionships and population structure are the main factors
resulting in false-positive associations. The GWAS was
based on SNPs and used TASSEL software (Cornell Uni-
versity, Ithaca, NY, USA) [48] to obtain correlation val-
uers using a compressed mixed linear model (MLM). The
formula is as follows: $Y = \alpha X + \beta Q + \mu K + e$, $\alpha$, $\beta$, $\mu$, and $e$. In the equation, $Y$ is the phenotypic trait, $X$ is the indi-
cator matrix of the genotype (fixed effect), $\alpha$ is the esti-
Parameter of fixed effect, $Q$ is the indicator matrix
determined parameter of fixed effect, $Q$ is the indicator matrix
of population genetic structure, $\beta$ is the effect of SNP, $K$
is the indicator matrix of the individual genetic relation-
ship, $\mu$ is the predicted random individual, and $e$ is the random residual, obeying $e \sim (0, \delta^2)$. Among them, the sample population structure $Q$ (Fig. S1) was calculated by admixture software (University of California, Los
Angeles, CA, USA) [46], and the affinity $K$ between sam-
ples was calculated using SPAGEdi software (Uni-
versité Libre de Bruxelles, Brussels, Belgium) [49]. MLMs
use $Q + K$ information. Finally, each SNP locus can
obtain a correlation value ($P$). The $P$-values were cor-
corrected using Bonferroni's method: $\alpha \leq 0.1$ and $\alpha \leq 0.05$
($P \leq 1.77 \times 10^{-7}$ and $P \leq 8.33 \times 10^{-8}$, respectively) [50].

Annotation of genes related to grape berry-shape traits
Based on the 566,129 SNPs developed from 279 grape
varieties (Table S2), the LD of SNPs in all samples was
analyzed, denoted by $r^2$ (Fig. S2). The $r^2$ value decays
to half of the initial value of 6.15 kb. A 6-kb region was
taken from the upstream and downstream of the SNP
sites with associations, and functional genes for the
associated regions were mined. We used the Clusters
of Orthologous Genes (COG), Gene Ontology (GO),
Kyoto Encyclopedia of Genes and Genomes (KEGG),
Swiss-Prot and Non-redundant (NR) databases for gene
annotation according to the regions formed by the asso-
ciated SNPs.

Expression analysis of candidate genes for grape berry-shape
traits
The expression values of the candidate genes for fruit
shape traits in the pericarp (including the skin and flesh),
flower, and seed were screened from the Gene Expression
Omnibus (GEO Datasets, No.GSE36128) [51]. The loga-
rithm of the original value based on 10 was taken, and the
heat map was drawn with Excel 13 (Microsoft Corpora-
tion, Redmond, Washington, USA).

Statistical analysis
Due to the high correlation of most fruit shape traits in 2
years, we used the mean value of two-year data for Prin-
cipal Component Analysis (PCA), correlation analysis,
and variation analysis. PCA and variation analysis were
 carried out using SPSS version 16.0 (IBM, Armonk, NY,
USA). Linear correlation analysis was performed using
Excel 13 (Microsoft Corporation, Redmond, Washing-
ton, USA).

Results
Quantitative distribution of different berry shapes
in grapes
In the present study, there were nine different fruit
shapes: flat round, heart-shaped, obovoid, ovoid, curve-
shaped, elliptic, round, long elliptic and long round
(some representative varieties are shown in Fig. 1). As
shown in Fig. 2, the number distribution of varie-
ties with elliptic shape was the largest at 112, account-
ing for 40.14% of the varieties, followed by the round
berry shape at 110, accounting for 39.43% of the vari-
ties. The varieties with long elliptic or long round berry
shapes were relatively fewer, with 21 and 10 varieties,
respectively, accounting for 7.53 and 3.58% of the vari-
ties, respectively. The curve-shaped varieties were the
least common, with only two, accounting for 0.72% of
the varieties. In addition, the number of the varieties
with flat round; heart-shaped, ovoid and obovoid berry
shapes was between 5 and 8, accounting for 1.79–2.87%
Fig. 1 Representative grape varieties with different shapes. **A**: flat round, Yiliang; **B**: elliptical, Lady Washington; **C**: round, Lival; **D**: heart-shaped, Kamea; **E**: ovoid, Jingkejing; **F**: curve-shaped, Lünai; **G**: obovoid, Beni Fuji; **H**: long elliptical, Qichakapulie; **I**: long round, Manai

Fig. 2 Quantitative distribution of different berry-shapes in grape varieties
Table 1 The distribution of different berry shapes in the *V. vinifera* L. and *V. vinifera* × *V. labrusca*

| Grape berry shape | Distribution of different berry shapes in the *V. vinifera* L. and *V. vinifera* × *V. labrusca* |
|-------------------|--------------------------------------------------------------------------------------------------|
|                   | *V. vinifera* L. | *V. vinifera* × *V. labrusca* |
| Curve-shaped      | 1 (50.00%)       | 1 (50.00%)                   |
| Obovoid           | 1 (20.00%)       | 4 (80.00%)                   |
| Heart-shaped      | 5 (100.00%)      | 0 (0.00%)                    |
| Ovoid             | 6 (100.00%)      | 0 (0.00%)                    |
| Flat round        | 5 (62.50%)       | 3 (37.50%)                   |
| Long round        | 10 (100.00%)     | 0 (0.00%)                    |
| Long elliptic     | 21 (100.00%)     | 0 (0.00%)                    |
| Round             | 75 (68.18%)      | 35 (31.82%)                  |
| Elliptic          | 80 (71.43%)      | 32 (28.57%)                  |

The number outside the brackets indicates the number of varieties in the interval, and the number inside the brackets indicates the percentage of the varieties in the interval.

Distribution of different grape berry-shape morphologies and quantities of the *V. vinifera* L. and *V. vinifera* × *V. labrusca*

The number distribution of different grape berry shapes of the *V. vinifera* L. and *V. vinifera* × *V. labrusca* is shown in Table 1. From the perspective of berry shape types, there were nine different berry shapes in the *V. vinifera* L.: curve-shaped, flat round, heart-shaped, obovoid, ovoid, elliptic, round, long elliptic and long round. There were five different berry shapes in the *V. vinifera* × *V. labrusca*: curve-shaped, flat round, obovoid, elliptic, and round. From the distribution of the four different berry shapes common in the *V. vinifera* L. and *V. vinifera* × *V. labrusca*, 62.5% of flat round shapes were distributed in the *V. vinifera* L.; 80% of obovoid berries were distributed in the *V. vinifera* × *V. labrusca*; and 20% obovoid berries were distributed in the *V. vinifera* L. The distributions of elliptic berries in the *V. vinifera* L. and *V. vinifera* × *V. labrusca* were 71.43 and 28.57%, respectively; 68.18% of round berries were distributed in

Fig. 3  Longitudinal section of the representative grape varieties with different shapes. A: flat round, Yiliang; B: elliptical, Lady Washington; C: round, Lival; D: heart-shaped, Kamea; E: ovoid, Jingkejing; F: curve-shaped, Lünai; G: obovoid, Beni Fuji; H: long elliptical, Qichakapulie; I: long round, Manai
Variation of different grape berry-shape parameters using the tomato analyzer

The longitudinal section of representative grape varieties of different shapes is shown in Fig. 3. Box diagrams of the grape berry shape-related parameters are shown in Fig. 4 and Fig. S3; these include the variation of these berry shape parameters. Table S3 shows that the variation in different berry shape parameters was different, and the variation range was 0.18% (distal eccentricity) to 63.64% (proximal indentation area). We found that among these parameters, the traits with a coefficient of variation less than 1% included distal eccentricity and proximal eccentricity; traits with a coefficient of variation higher than 20% included the maximum width, width mid-height, perimeter, curved height, maximum height, height mid-width, area, shoulder height and proximal indentation area.

The PCA and correlation analysis of grape berry shape-related parameters

The PCA of grape berry shape-related parameters is shown in Fig. 5, Table S4, and Table S5. As shown in Table S4, the characteristic values of the first five components in this study were all greater than 1, and the cumulative contribution rate was 78.426%, indicating that the explanatory rate of these five factors to the whole population was nearly 80%, so the first five factors could be extracted. The cumulative contribution rate of the first two principal components reached 54.752%. The perimeter, area, width mid-height, maximum width, height mid-width, maximum height, curved height, fruit shape

**Fig. 4** The distribution of the basic measured characters and shape index parameters of grape berries
index external I, fruit shape index external II, curved fruit shape index, and fruit shape index internal were highly correlated with PCA1 and PCA2. These characters contained most of the variation of fruit shape characters. The correlation analysis of the same fruit shape-related parameters in 2 years (2019 and 2020) is shown in Fig. S4. For most fruit shape traits, the correlation of two-year data is high, which indicates that these traits have high heritability. In addition, the correlation analysis of the grape berry shape-related parameters is shown in Fig. 6 and Table S6. The seven traits of perimeter, area, width mid-height; maximum width, height mid-width, maximum height and curved height were positively correlated with each other, and the coefficients were 0.836–0.999. The fruit shape index external I, fruit shape index external II, fruit shape index internal and curved fruit shape index were positively correlated with each other, and the correlation coefficients were 0.854–0.997.

Results of grape berry shape traits in the GWAS

Using the 25 fruit shape traits in 2019 and 2020 as the target traits, GWAS was performed using MLM. The GWAS results showed that the four fruit shape traits analyzed in this study (curved fruit shape index, fruit shape index external I, fruit shape index external II and fruit shape index internal) were significantly correlated with multiple SNP loci within 2 years. Therefore, this study only provides the analysis results of these four traits. GWAS results with the MLM for curved fruit shape index, fruit shape index external I, fruit shape index external II, and fruit shape index internal are shown in Figs. 7, 8, 9 and 10, respectively, and the detailed results are shown in Table S7, Table S8, Table S9, and Table S10, respectively.

SNP loci associated with simultaneous control of multiple fruit shape traits

As shown in Table S11, we found that multiple berry shape traits were associated with the same SNP loci and the same berry shape trait was significantly associated with multiple SNP loci. In both years, 122 SNP loci had
significant correlations with multiple berry shape traits, which were distributed on 19 chromosomes. Among them, 18 SNPs were located on chromosome 9, 16 SNPs were distributed on chromosome 12 and one SNP was distributed on chromosomes 1, 2, 11 and 15. The number of SNPs distributed on other chromosomes ranged from 3 to 16.

The grape berry shape traits of curved fruit shape index, fruit shape index external II, and fruit shape index internal were associated with four significant SNP loci (marker names: 17_173150, 17_414350, 17_43912 and 17_584380) in 2 years; fruit shape index external I and fruit shape index external II were associated with one significant SNP locus (marker name: 17_604843) in 2 years; curved fruit shape index, fruit shape index external I, fruit shape index external II and fruit shape index internal were associated with 29 significant SNP loci (marker name: 17_103635, 17_158117, 17_165943, 17_165944, 17_165946, 17_165948, 17_165952, 17_165952).

**Fig. 6** Correlation analysis of grape berry shape-related parameters. A: perimeter; B: area; C: width mid-height; D: maximum width; E: height mid-width; F: maximum height; G: curved height; H: fruit shape index external I; I: fruit shape index external II; J: fruit shape index internal; K: curved fruit shape index; L: proximal fruit blockiness; M: distal fruit blockiness; N: fruit shape triangle; O: eccentricity; P: proximal Eccentricity; Q: distal eccentricity; R: width widest pos; S: eccentricity area index; T: proximal angle micro; U: proximal angle macro; V: distal angle micro; W: distal angle macro; X: proximal indentation area; Y: shoulder height.
17_173305, 17_176706, 17_211324, 17_230957, 17_304846, 17_305195, 17_309493, 17_312547, 17_326935, 17_335450, 17_337013, 17_337015, 17_337017, 17_337018, 17_482077, 17_491897, 17_516416, 17_528926, 17_552659, 17_572735, 17_590009, 17_88457, 17_88459, 17_88457 and 17_88459) in 2 years.

A1 and A2 are Manhattan plots of the mixed linear model (MLM). The abscissa represents the position of the chromosome, and the ordinate represents the $P$-value ($-\log_{10} P$). The negative logarithm of the base 10 and the scattered points (or lines) on the graph represent the $-\log_{10} (p)$ corresponding to each SNP locus. The red dotted line is the negative logarithm of 0.05/all SNPs, and the blue dotted line is the negative logarithm of 0.1/all SNPs. Scattered dots (or lines) above the threshold line are candidate sites. B1 and B2 are QQ-plot plots. The abscissa represents the expected value, and the ordinate represents the observed value. In the initial stage, the actual observed $P$-value was close to the expected $P$-value, indicating that the influence of population structure on the association analysis could be effectively controlled under this model, and thus false positives could also be effectively controlled. The same applies below.

**Polygenic control of grape berry shape traits**

In our analysis, we found that some candidate genes were candidates for multiple berry shape traits, and the same trait could be coordinated by multiple genes (Table 2). The relevant candidate genes unearthed in this study mainly included genes related to transcription factors, cell wall metabolism, plant hormones, ubiquitin ligases and serine/threonine protein kinases. Five transcription factor-related genes (VIT_02s0025g01360, VIT_06s0004g01280, VIT_12s0057g00880, VIT_16s0022g02330 and VIT_18s0076g00330), four cell wall metabolism-related genes (VIT_07s0001g04110, VIT_08s0007g00440, VIT_08s0007g00290 and VIT_09s0096g00850), two plant hormones-related genes (VIT_02s0025g01360 and VIT_12s0134g00230), and two LRR receptor-like serine/threonine protein kinase genes (VIT_09s0002g03030 and VIT_09s0070g00140) were associated with the fruit
Fig. 8 Genome-wide association study with the MLM for fruit shape index external I

Fig. 9 Genome-wide association study with the MLM for fruit shape index external II
shape index external II and fruit shape index internal; two ubiquitin ligases-related genes (\textit{VIT}_03s0088g01090 and \textit{VIT}_10s0003g04300) were associated with the fruit shape index external I, fruit shape index external II and fruit shape index internal. The results showed that these berry shape traits were co-regulated by some genes involved in the regulation of berry morphology.

**Tissue expression analysis of candidate genes annotated by grape fruit shape traits**

Tissue expression analysis of candidate genes annotated by grape fruit shape traits was performed by GEO Datasets (No.GSE36128) [51], as shown in Fig. 11. The results showed that two plant hormone-related genes (\textit{VIT}_02s0025g01360, ethylene-responsive transcription factor ERF061; \textit{VIT}_12s0134g00230, incole-3-acetic acid-amido synthetase GH3.6), two ubiquitin ligase-related genes (\textit{VIT}_03s0088g0109, RING finger protein 44; \textit{VIT}_10s0003g04300, F-box protein SKIP19), two LRR receptor-like serine/threonine-protein kinase-related genes (\textit{VIT}_09s0002g03030, At1g05700; \textit{VIT}_09-s0070g00850, At1g07650); four cell wall metabolism-related genes (\textit{VIT}_07s0005g04110, cellulose synthase A catalytic subunit 4; \textit{VIT}_08s0007g00290, pectin acetylerase 5; \textit{VIT}_08s0007g00440, expansin-A6; \textit{VIT}_09s0096g00850, probable polygalacturonase At3g15720), and other genes were expressed to varying degrees in grape pericarp, flowers, and seeds at different developmental stages. Thus, these genes can be used as candidate genes for grape fruit shape traits.

**Discussion**

Compared with wild varieties with round fruits, a cultivar’s fruit shape has a high degree of diversity [2]. To meet different market demands, various types of fruit shapes have gradually been produced during the genetic improvement of horticultural crops [1–3]. Compared with the fruit shape research in tomato [3, 6], watermelon [17] and sweet pepper [18], few studies have examined grape berry shapes. The mining of berry shape-related genes holds great significance for breeding new grape varieties with different fruit shapes.

Measuring the fruit morphology and color characteristics of vegetable and fruit crops in an objective and reproducible manner is important for the detailed phenotypic analysis of these traits [40]. The Tomato Analyzer is a software program that measures 37 two-dimensional shape-related attributes in a semi-automated and reproducible manner [11, 52].
| Berry shape traits                                                                 | Gene ID                          | Location                    | Nr annotation                                                                 |
|-----------------------------------------------------------------------------------|----------------------------------|------------------------------|--------------------------------------------------------------------------------|
| Fruit Shape Index External II, Fruit Shape Index Internal                         | VIT_03s0097g00140                | 3:9892772–9,893,537          | PREDICTED: *Vitis vinifera* L-type lectin-domain containing receptor kinase IX.1 (LOC100265547), mRNA |
|                                                                                  | VIT_04s0023g01970                | 4:18530909–18,534,706        | PREDICTED: *Vitis vinifera* putative glucose-6-phosphate 1-epimerase (LOC100266694), mRNA |
|                                                                                  | VIT_06s0044g00020                | 6:95994–97,288               | PREDICTED: *Vitis vinifera* NAC-domain-containing protein 68 (LOC100263939), mRNA |
|                                                                                  | VIT_06s0044g01280                | 6:1524152–1,537,550          | PREDICTED: GATA transcription factor 23-like (*Vitis vinifera*)                 |
|                                                                                  | VIT_07s0005g04080                | 7:7175304–7,176,037          | PREDICTED: *Vitis vinifera* MDIS1-interacting receptor like kinase 2-like (LOC100246300), mRNA |
|                                                                                  | VIT_07s0005g04100                | 7:7199256–7,202,858          | PREDICTED: *Vitis vinifera* MDIS1-interacting receptor like kinase 2-like (LOC100266867), mRNA |
|                                                                                  | VIT_08s0007g00290                | 8:14603038–14,609,600        | PREDICTED: *Vitis vinifera* pectin acetyl-esterase S (LOC100244164), transcript variant X1, mRNA |
|                                                                                  | VIT_08s0007g00440                | 8:14732840–14,734,936        | PREDICTED: *Vitis vinifera* expansin-A6 (LOC100245911), mRNA                  |
|                                                                                  | VIT_09s0002g03020                | 9:2728009–2,740,198          | PREDICTED: *Vitis vinifera* putative leucine-rich repeat receptor-like protein kinase At1g19210 (LOC100266874), mRNA |
|                                                                                  | VIT_09s0002g03030                | 9:2778212–2,782,247          | PREDICTED: *Vitis vinifera* probable LRR receptor-like serine/threonine-protein kinase At1g05700 (LOC100251452), mRNA |
|                                                                                  | VIT_09s0007g00850                | 9:14657091–14,663,881        | PREDICTED: *Vitis vinifera* probable LRR receptor-like serine/threonine-protein kinase At1g07650 (LOC100251452), transcript variant X1, mRNA |
|                                                                                  | VIT_10s0092g00200                | 10:11506152–11,510,357       | PREDICTED: GDP-mannose transporter GONST3 (*Vitis vinifera*)                  |
|                                                                                  | VIT_12s0134g00230                | 12:7753336–7,755,364        | PREDICTED: *Vitis vinifera* indole-3-acetic acid-amido synthetase GH3.6 (LOC100255170), mRNA |
|                                                                                  | VIT_12s0134g00240                | 12:7768803–7,769,369         | PREDICTED: *Vitis vinifera* calcium-binding protein PBPI (LOC100265313), mRNA |
|                                                                                  | VIT_16s0022g02330                | 16:14940190–14,955,349       | PREDICTED: *Vitis vinifera* MADS-box transcription factor 6 (LOC100266085), transcript variant X1, mRNA |
|                                                                                  | VIT_19s0014g00550                | 19:560043–564,043           | PREDICTED: *Vitis vinifera* TM1-like protein 2 (LOC100264036), transcript variant X1, mRNA |
|                                                                                  | VIT_19s0014g00560                | 19:572717–579,014        | PREDICTED: *Vitis vinifera* probable arabinosyltransferase ARAD1 (LOC100241759), transcript variant X2, mRNA |
| Curved Fruit Shape Index, Fruit Shape Index External II, Fruit Shape Index Internal | VIT_07s0005g04110                | 7:7204959–7,209,235          | PREDICTED: *Vitis vinifera* cellulose synthase A catalytic subunit 4 (UDP-forming) (LOC100241197), mRNA |
Analysis of the variability of grape fruit shape-related traits using the tomato analyzer

The output produced by the Tomato Analyzer can be used in many applications. In genetic research, the output has been used to detect fruit shape QTLs in several isolated populations derived from crosses between different cultivated tomato varieties, including LA1589 (Solanum lycopersicum) and wild species Solanum pimpinellifolium accessions [11, 53]. The Tomato Analyzer also has been used for shape diversity [54] and fruit color [55] studies. In this study, we used the Tomato Analyzer to analyze 25 tomato shape-related traits in 279 varieties. We found that these fruit shape-related variations ranged from 0.18 to 63.64%. Among them, the distal eccentricity and proximal eccentricity had relatively small variation, while proximal indentation area and shoulder height exhibited relatively large variation. For other fruit-shaped traits, the degree of variation was between 2.00 and 40.24%. Some studies have examined the genetic variation of fruit traits using the Tomato Analyzer. A previous study found that the broad heritability of most fruit traits was high in tomato and Capsicum annuum [56, 57]. The broad heritability of shoulder height was 0.56 [56]. In the present study, the coefficient of variation of shoulder height was found to be high, which indicated that it responded to variety characteristics.

For the studies on candidate genes related to fruit shape, in addition to tomatoes [2, 20, 25, 51], in-depth research has been conducted in horticultural crops such as peaches [28–31] and cucumbers [32, 33].

| Berry shape traits                                                                 | Gene ID           | Location          | Nr annotation                                                                 |
|----------------------------------------------------------------------------------|-------------------|-------------------|------------------------------------------------------------------------------|
| Fruit Shape Index Internal I, Fruit Shape Index External II, Fruit Shape Index Internal | VIT_02s0025g01360 | 2:1274105–1,275,209 | PREDICTED: ethylene-responsive transcription factor ERF061 [Vitis vinifera] |
|                                                                                  | VIT_07s0005g02370 | 7:4729295–4,730,035 | PREDICTED: Vitis vinifera germin-like protein 5–1 (LOC100260064), mRNA       |
|                                                                                  | VIT_07s0005g02380 | 7:4743365–4,744,223 | PREDICTED: Vitis vinifera germin-like protein subfamily 2 member 4 (LOC100267594), mRNA |
|                                                                                  | VIT_09s0002g08430 | 9:4924676–9,434,529 | PREDICTED: Vitis vinifera protein NUCLEAR FUSION DEFECTIVE 4 (LOC100262975), mRNA |
|                                                                                  | VIT_09s0096g00850 | 9:12518294–12,520,356 | PREDICTED: Vitis vinifera probable polygalacturonase At3g15720 (LOC100251699), mRNA |
|                                                                                  | VIT_09s0070g00140 | 9:13115563–13,117,457 | PREDICTED: Vitis vinifera CBL-interacting serine/threonine-protein kinase 5 (LOC100255067), mRNA |
|                                                                                  | VIT_09s0018g02060 | 9:19788292–19,790,440 | PREDICTED: Vitis vinifera sugar transport protein 8 (HT14), mRNA               |
|                                                                                  | VIT_10s0033g04300 | 10:7394269–7,397,548 | PREDICTED: Vitis vinifera F-box protein SKIP19 (LOC100265193), transcript variant X1, mRNA |
|                                                                                  | VIT_10s0116g00680 | 10:300617–304,777 | PREDICTED: Vitis vinifera homeobox-leucine zipper protein MERISTEM L1 (LOC100264009), mRNA |
|                                                                                  | VIT_12s0057g00880 | 12:9587450–9,598,955 | PREDICTED: transcription initiation factor TFIID subunit 15b [Vitis vinifera] |
|                                                                                  | VIT_13s0019g05160 | 13:6937392–6,997,802 | PREDICTED: Vitis vinifera pullulanase 1, chloroplastic (LOC100247866), transcript variant X1, mRNA |
|                                                                                  | VIT_18s0076g00330 | 18:16209694–16,218,061 | PREDICTED: Vitis vinifera transcription factor VIP1 (LOC100241011), mRNA |
| Curved Fruit Shape Index, Fruit Shape Index External II, Fruit Shape Index Internal | VIT_16s0013g00860 | 16:6368533–6,436,142 | PREDICTED: Vitis vinifera polycorm group protein EMBRYONIC FLOWER 2 (LOC100252876), transcript variant X1, mRNA |
|                                                                                  | VIT_09s0002g08370 | 9:9246440–9,256,248 | PREDICTED: Vitis vinifera protein TOPOLESS (LOC100248092), mRNA               |
|                                                                                  | VIT_13s0064g00360 | 13:21879764–21,898,935 | PREDICTED: Vitis vinifera notchless protein homolog (LOC100262007), transcript variant X1, mRNA |
|                                                                                  | VIT_03s0088g01090 | 3:9340156–9,341,562 | PREDICTED: Vitis vinifera RING finger protein 44 (LOC104878839), mRNA          |
Fig. 11 Tissue expression analysis of candidate genes annotated by grape fruit shape traits. FS: fruit set; PFS: post-fruit set; V: veraison; MR: mid-ripening; R: ripening; FB: flowering begins (10% caps off); F: flowering (50% caps off)
However, few studies have examined candidate genes related to grape shape.

**GWAS of genes related to grape berry shape traits**

In this study, we used a genome-wide association study to analyze grape berry shape as the target trait using the Tomato Analyzer and mined some of the candidate genes that control grape berry shape. The relevant candidate genes unearthed in this study included genes related to plant hormones, ubiquitin ligases, LRR receptor-like serine/threonine-protein kinase, and transcription factors.

**GWAS of plant hormone-related genes related to grape berry shape characters**

In the present study, through the mining of functional gene-associated regions, we identified two genes related to plant hormones that were related to berry shape—namely, *VIT_12s0134g00230* (indole-3-acetic acid-amido synthetase GH3.6) and *VIT_02s0025g01360* (ethylene-responsive transcription factor ERF061). These two genes were identified as candidate genes for the berry shape traits fruit shape index external II and fruit shape index internal.

Plant hormones play an important role in fruit organ morphogenesis and development [58–61]. Studies have identified auxin and gibberellin as early signs of fruit setting and fruit growth [58–60]. Both hormones have a positive effect on cell division and cell expansion [61]. *AtOFP1* is reported to be located in the nucleus and acts as an active transcriptional repressor, regulating a gene in the gibberellin biosynthetic pathway (*AtGA20ox1*). The reduction of cell elongation is partly caused by the inhibition of gibberellin biosynthesis [24]. Further research showed that *CaOvate* down-regulated the *CaGA20ox1* gene, which was similar to the tomato *GA20ox1* gene. *CaGA20ox1* regulates the effect of *CaOvate* on fruit elongation [62]. Unfortunately, the gibberellin-related gene was not linked in this study. In addition, a previous study pointed out that calcium signaling regulates cell polarity and cell elongation by regulating auxin transport in tobacco [63]. The number and size of short fruit cells were lower than those of long fruit, which may have been caused by abnormal auxin signal transduction in short fruit [64]. Studies have shown that the binding of indole-3-acetic acid (IAA)-amido synthetases with amino acids is an important aspect of auxin stability in vivo [65, 66]. In addition to auxin and gibberellin, ethylene content may also be involved in fruit morphogenesis [58]. Studies have suggested that genes involved in hormone action, such as ethylene-related
genes, are upregulated in fruits, indicating that genes related to cell cycle control and hormone action may contribute to the process of fruit development from cell division to cell expansion [58]. Based on gene chips from GEO Datasets (No.GSE36128) [51], we analyzed the expression levels of these two plant hormone-related genes (*VIT_12s0134g00230* and *VIT_02s0025g01360*) in pericarp, flower, and seed, and found that the two hormone-related genes were expressed to varying degrees in different grape organs at different growth and development stages, indicating that these two genes may be involved in the formation of grape fruit organ morphology. On the basis of findings from these related studies, supported by the results of this study, the two genes *VIT_12s0134g00230* (indole-3-acetic acid-amido synthetase GH3.6) and *VIT_02s0025g01360* (ethylene-responsive transcription factor ERF061) may affect grape berry morphology by regulating auxin and ethylene content, but the specific mechanism requires further study.

**GWAS of ubiquitin ligase-related genes related to grape berry-shape characters**

In the present study, through the mining of functional gene-associated regions, we identified two genes related to ubiquitin ligases that were related to berry shape—namely, one RING finger protein 44 -related gene (*VIT_03s0088g01090*) and one F-box protein SKIP19-related gene (*VIT_10s0003g04300*). The two genes are candidate genes for the berry characteristics fruit shape index external I, fruit shape index external II and fruit shape index internal. In addition, the gene *VIT_03s0088g01090* is also a candidate gene of the curved fruit shape index. Ubiquitination is a fine post-translational modification that is widely found in all eukaryotes [67]. Ubiquitin is a conserved protein with 76 amino acids that has a high degree of conservation and involves various aspects of cell physiology [67, 68].

RING-type E3 is one of the ubiquitin ligases, and many studies have been conducted on its regulation of plant organ morphology [67, 68], especially the regulation of seed organs [67]. Seed size is an important agronomic trait. Several regulatory pathways that determine seed size have been identified, among which RING-type E3 ligases are involved, mainly by regulating gametogenesis and cell cycle processes. RING-type E3 DA2 negatively regulates seed size by reducing cell proliferation and synergistic interactions with the ubiquitin receptor DA1 in developing seeds. The ubiquitin receptor DA1 is also a key regulator of seed size [69]. Previous studies have shown that the DA2 homolog RING-type E3 OsGW2 (Grain Width and Weight 2) in rice has a negative effect on particle size and final yield.
by mediating cell division [70]. In addition to controlling seed shape, ubiquitin ligase may also play an important role in controlling fruit shape. In the present study, the two ubiquitin ligase genes (VIT_03s0088g01090 and VIT_10s0003g04300) associated with multiple berry shape traits were normally expressed in the pericarp (as shown in Fig. 10), suggesting that they may play an important role in regulating berry morphogenesis. The in-depth mechanism of the regulation of grape berry shape traits by two ubiquitin ligase genes (VIT_03s0088g01090 and VIT_10s0003g04300) needs further study.

GWAS association analysis of LRR receptor-like serine/threonine-protein kinase genes related to grape berry-shape traits
In the present study, the functional gene mining of berry-shaped trait-associated regions was performed, and two LRR receptor-like serine/threonine-protein kinase genes (VIT_09s0002g03030, At1g05700; VIT_09s0070g00850, At1g07650) were obtained. Both genes are located on chromosome 9 and are candidate genes for fruit shape index external II and fruit shape index internal. Some research has suggested that the shape of the pit in peach can be used to distinguish traditional varieties [71]. Mapping-based cloning methods have revealed that candidate genes for this trait may be LRR-RLK protein kinases rather than MADS-box genes [72]. Although tissue expression patterns (as shown in Fig. 10) showed that two LRR receptor-like serine/threonine-protein kinase genes (VIT_09s0002g03030, At1g05700; VIT_09s0070g00850, At1g07650) were normally expressed in grape pericarp, the expression of the two genes in different fruit shapes was not analyzed. The specific mechanism of VIT_05s0002g03030 and VIT_09s0070g00850 regulating the traits fruit shape index external II and fruit shape index internal needs further study.

In addition, we examined some transcription factors (transcription factor VIP1, GATA transcription factor 23-like, transcription initiation factor TFIIID and MADS-box transcription factor 6) related to grape berry shape in this study. Tissue expression patterns showed that these transcription factor correlations were expressed to a certain extent in tissues at different stages of grape development (Fig. 10). However, few reports, are available about the relationship between these genes and fruit shape. Whether these genes regulate grape berry shape and the specific mechanism needs to be further studied.

Conclusion
To discover candidate genes related to grape berry shape, the present study first took fruit shape parameters analyzed by the Tomato Analyzer as the target traits and used genome-wide association study to analyze candidate shape related genes. The relevant candidate genes unearthed in this study included genes related to plant hormones, ubiquitin ligase, LRR receptor-like serine/threonine-protein kinase and transcription factors. The present study increased the understanding of the genetic control of grape berry shape traits. The identification of molecular markers that are closely related to these berry shape traits holds great significance for breeding specific berry shape varieties.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s12870-022-03434-x.

Additional file 1: Figure S1. Population structure of natural populations.
Additional file 2: Figure S2. Attenuation analysis of LD at the population level.
Additional file 3: Figure S3. Distribution of the other morphological traits of grape berries.
Additional file 4: Figure S4. Correlation analysis of the same berry shape-related traits across 2 years.
Additional file 5: Table S1. Experimental material used in this study.
Additional file 6: Table S2. SNP number distribution on each chromosome.
Additional file 7: Table S3. The variation in different berry shape parameters.
Additional file 8: Table S4. Total variance explained.
Additional file 9: Table S5. Component matrix of grape berry shape-related parameters.
Additional file 10: Table S6. Correlation analysis of grape berry shape-related parameters.
Additional file 11: Table S7. Details of SNP loci associated with curved fruit shape index identified via GWAS from both years.
Additional file 12: Table S8. Details of SNP loci associated with fruit shape index external I identified via GWAS from both years.
Additional file 13: Table S9. Details of SNP loci associated with fruit shape index external II identified via GWAS from both years.
Additional file 14: Table S10. Details of SNP loci associated with multiple berry-shape traits identified via GWAS from both years.
Additional file 15: Table S11. Details of SNP loci associated with curved fruit shape index identified via GWAS from both years.

Abbreviations
BWA: Burrow-Wheeler Aligner; CLAVATA3: CLV3; COG: Clusters of Orthologous Genes; DNA: Deoxyribonucleic acid; FAS: FASCIALED; GA: Gibberellic acid; GEO: Gene Expression Omnibus; GO: Gene Ontology; GW1: Width and Weight 1; GWAS: Genome-wide association study; IAA: Indole-3-acetic acid; IPGRI: International Plant Genetic Resources Institute; IQD: IQ67-domain; KEGG: Kyoto Encyclopedia of Genes and Genomes; LC: LOCULE NUMBER; LD: Linkage Disequilibrium; LRR: Leucine-rich repeat; MAF: Minor allele frequency; MLM: Mixed linear model; NR: Non-redundant; OFP: Ovate Family Protein; PCA: Principal Component Analysis; QTL: Quantitative trait loci; Ring: Really interesting gene; SNP: Single-nucleotide polymorphism; TRM: TONNEAU1 Recruiting Motif; UPOV: International Union for the Protection of New Varieties of Plants.
Acknowledgements
We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

Authors’ contributions
JGF and CZ provided the experimental ideas and designed the research; CZ specifically implemented the research and performed data analysis; LWC made preliminary revisions to the paper; JGF, LWC and CZ made the final decision on the paper. All authors have read and approved the manuscript.

Funding
This research was funded by the National Key Research and Development Program (No. 2019YFD1000101). The funders were not involved in the experimental design of the study, data collection, analysis and interpretation, and in writing the manuscript.

Availability of data and materials
The datasets generated during and/or analysed during the current study are included in its supplementary information files. The raw Illumina sequencing data from this study have been submitted to NCBI Sequence Read Archive (SRA) under the accession number PRJNA782678.

Declarations
Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1 College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China. 2 College of Agro-grassland Science, Nanjing Agricultural University, Nanjing 210095, China.

Received: 24 May 2021 Accepted: 10 January 2022
Published online: 20 January 2022

References
1. Paran I, van der Knaap E. Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. J Exp Bot. 2007;58:3841–52.
2. Rodríguez GR, Muñoz S, Anderson C, Sim SC, Michel A, Causse M, et al. 2011. Distribution of SUN, OVATE, LC, and FAS in the tomato germplasm and the relationship to fruit shape diversity. Plant Physiol. 2011;156:275–85.
3. van der Knaap E, Chakrabarti M, Chu YH, Clevenger JP, Illa-Berenguer E, et al. What lies beyond the eye: the molecular mechanisms regulating tomato fruit weight and shape. Front Plant Sci. 2014;5:1–13.
4. Wang S, Chang Y, Ellis B. Overview of OVATE FAMILY PROTEINS, a novel class of plant-specific growth regulators. Front Plant Sci. 2016;7:417.
5. Visa S, Cao C, Gardener BM, van der Knaap E. Modeling of tomato fruits into nine shape categories using elliptic fourier shape modeling and Bayesian classification of contour morphometric data. Euphytica. 2014;200:429–39.
6. Ziaf K, Amjad M, Shakeel A, Azhar M, Saeed A. Assessment of genetic diversity in tomato for fruit morphology, composition and yield. Pak J Bot. 2016;48:2477–83.
7. Tümbilen Y, Fray A, Mutlu S, Doganlar S. Genetic diversity in Turkish eggplant (Solanum melongena) varieties as determined by morphological and molecular analyses. Int Rev J Biotechnol. 2011;2:16–25.
8. White AG, Alsipach PA, Weskett RH, Brewer LR. Harvesting tomato fruit in pear. Euphytica. 2000;112:1–7.
9. IPGRI. Descriptors for Tomato (Lycopersicon spp.). Rome: International Plant Genetic Resources Institute; 1996.
10. UPOV. Guidelines for the conduct of tests for distinctness, uniformity and stability (tomato). Geneva, Switzerland; 2001.
11. Brewer MT, Lang L, Fujimura K, Dujmovic N, Gray S, van der Knaap E. Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species. Plant Physiol. 2006;141:15–25.
12. Beyer M, Hahn R, Peschel S, Harz M, Knoche M. Analyzing fruit shape in sweet cherry (Prunus avium L.). Sci Hortic. 2002;96:139–50.
13. Tümbilen Y. Determination of genetic diversity between eggplant and its wild relatives. Master’s Thesis, IZMIR: Izmir Institute of Technology; 2007.
14. Guo SG, Zhang JG, Sun HH, Salsé J, Lucas WU, Zhang HY, et al. The draft genome of watermelon (Citrullus lanatus) and resequencing of 20 diverse accessions. Nat Genet. 2013;45:51.
15. Munos S, Ranc N, Botton E, Berard A, Rolland S, Dufte P, et al. 2011. Increase intomato locule number is controlled by two single-nucleotide polymorphisms located near WUSCHEL. Plant Physiol. 2011;156:2246–54.
16. Wei QZ, Fu WY, Wang YZ, Qin XD, Wang J, Li J, et al. Rapid identification of fruit length loci in cucumber (Cucumis sativus L.) using next-generation sequencing (NGS)-based QTL analysis. Sci Rep. 2016;6:1–11.
17. Dou J, Zhao S, Lu X, He N, Zhang L, Ali A, et al. Genetic mapping reveals a candidate gene (CIFS1) for fruit shape in watermelon (Citrullus lanatus L.). Theor Appl Genet. 2018;131:947–58.
18. Chunthawodtiporn J, Hill T, Stoffel K, Van Deynze A. Quantitative trait loci controlling fruit size and other horticultural traits in bell pepper (Capsicum annuum). Plant Genome. 2018;11:160125.
19. Liu J, van Eck J, Cong B, Tanksley SD. A new class of regulatory genes underlying the cause of pear-shaped tomato fruit. P Natl Acad Sci USA. 2002;99:13302–6.
20. Xiao H, Jiang N, Schaffner EK, Stockinger EJ, van der Knaap E. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science. 2008;319:1527–30.
21. Jiang N, Gao D, Xiao H, van der Knaap E. Genome organization of the tomato sun locus and characterization of the unusual retrotransposon rider. Plant J. 2009;60:181–93.
22. Wu S, Xiao H, Cabrera A, Meulia T, van der Knaap E. SUN regulates vegetative and reproductive organ shape by changing cell division patterns. Plant Physiol. 2011;157:1175–86.
23. Wang Y, Clevenger JP, Illa-Berenguer E, Meulia T, van der Knaap E, Sun L. A comparison of sun, ovate, f8.1 and auxin application on tomato fruit shape and gene expression. Plant Cell Physiol. 2019;60:1067–81.
24. Wang SY, Chang Y, Guo JJ, Chen G. Arabidopsis ovate Family protein 1 is a transcriptional repressor that suppresses cell elongation. Plant J. 2007;50:885–72.
25. Cong B, Barrero LS, Tanksley SD. Regulatory change in YABBY-like transcription factor led to evolution of extreme fruit size during tomato domestication. Nat Genet. 2008;40:800–4.
26. Xu C, Liberatore KL, MacAlister CA, Huang Z, Chu YH, Jiang K, et al. A cascade of arabinosyntransferases controls shoot meristem size in tomato. Nat Genet. 2015;47:784–92.
27. Lou L. Inheritance of fruit characteristics in watermelon (Citrullus lanatus (Thunb.) Matsum. & Nakai), Master’s Thesis, Raleigh, North Carolina: North Carolina State University; 2009.
28. Dirlewanger E, Pronier V, Parvency C, Rothan C, Guye A, Monet R. Genetic linkage map of peach (Prunus persica (L.) Batsch) using morphological and molecular markers. Theor Appl Genet. 1996;92:888–95.
29. Dirlewanger E, Cosson P, Boudieri K, Renaud C, Capdeville G, Tausin Y, et al. Development of a second-generation genetic linkage map for peach (Prunus persica (L.) Batsch) and characterization of morphological traits affecting flower and fruit. Tree Genet Genomes. 2006;3:1–13.
30. Tan Q, Liu X, Xiao W, Gao H, Chen X, Fu X, et al. Comparison between flat and round peaches, genomic evidences of heterozygosity events. Front Plant Sci. 2019;10:592.
31. Zhou H, Ma R, Gao L, Zhang JY, Zhang AD, Zhang XJ, et al. A 1.7-Mb transposable element-mediated gene duplication underlies morphological variation of tomato fruit. Plant Physiol. 2019;171:1662–74.
33. Zhang TT, Li XZ, Yang YT, Guo X, Feng Q, Dong XG, et al. Genetic analysis and QTL mapping of fruit length and diameter in a cucumber (Cucumis sativus L.) recombinant inbred line (RIL) population. Sci Hortic. 2019;250:142–22.

34. Wang X, Li XH, Gao Z, Wang L, Ren Z. Localization of quantitative trait loci for cucumber fruit shape by a population of chromosome segment substitution lines. Sci Rep. 2020;10:11030.

35. Wu S, Zhang B, Keyhaninejad N, Rodriguez GR, Kim HJ, Chakrabarti M, et al. Ripening transcriptomic program in red and white grapevine varieties correlates with berry skin anthocyanin accumulation. Plant Physiol. 2017;174:2376–96.

36. Conde C, Silva P, Fontes N, Dias ACP, Tavares RM, Sousa MI, et al. 2007. Biochemical changes throughout grape berry development and fruit and wine quality. Food. 2007;1:1–22.

37. Du HS, Yang J, Chen B, Zhang X, Zhang J, Yang K, et al. Target sequencing reveals genetic diversity, population structure, core-SNP markers, and fruit shape-associated loci in pepper varieties. BMC Plant Biol. 2019;19:1–16.

38. Perin C, Hagen L, Giovanni N, Besontes D, Dogimont C, Pitrat M. Genetic control of fruit shape acts prior to anthesis in melon (Cucumis melo L.). MolGenGenomes. 2002;266:813–41.

39. Rodriguez GR, Moyseenko JB, Robbins MD, Morejón NH, Francis DM, van der Knaap E, Dujmovic N, Gray S. Application of high-throughput phenotyping tool tomato analyzer to characterize Balkan Cucumis fruit diversity. Sci Hortic. 2020;260:108862.

40. Darrigues A, Hall J, van der Knaap E, Francis DM, Dujmovic N, Gray S. Analysis and QTL mapping of fruit length and diameter in a cucumber (Cucumis sativus L.) recombinant inbred line (RIL) population. Sci Hortic. 2011;141:105–17.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.