Pepper variome reveals the history and key loci associated with fruit domestication and diversification

Yacong Cao, Kang Zhang, Hailong Yu, Shumin Chen, Donghui Xu, Hong Zhao, Zhenghai Zhang, Yinqing Yang, Xiaozhen Gu, Xinyan Liu, et al.

CellPress
Partner Journal

Molecular Plant
Research Article

Pepper variome reveals the history and key loci associated with fruit domestication and diversification

Yacong Cao[1,7], Kang Zhang[1,7], Hailong Yu[1,7], Shumin Chen[1,7], Donghui Xu[1,7], Hong Zhao[1], Zhenghai Zhang[1], Yinqing Yang[1], Xiaozhen Gu[1], Xinyan Liu[1], Haiping Wang[1], Yaxin Jing[1], Yajie Mei[1], Xiang Wang[1], Véronique Lefebvre[2], Weili Zhang[1], Yuan Jin[1], Dongliang An[1], Risheng Wang[3], Paul Bosland[4], Xixiang Li[1], Ilan Paran[5], Baoxi Zhang[1], Giovanni Giuliano[6,*], Lihao Wang[1,*] and Feng Cheng[1,*]

[1]Key Laboratory of Vegetables, Genetics, and Physiology of China Ministry of Agriculture and Rural Affairs, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops of the Ministry of Agriculture and Rural Affairs, Sino-Dutch Joint Laboratory of Horticultural Genomics, Institute of Vegetables and Flowers, CAAS (Chinese Academy of Agricultural Sciences), 12 Zhongguancun South Street, Beijing 100081, P. R. China

[2]INRAE, GAFL, Unité de Génétique et Amélioration des Fruits et Légumes, 84140 Montfavet, France

[3]Institute of Vegetables, Academy of Agricultural Sciences of Guangxi, 174 Daxue East Road, Nanning 53007, P. R. China

[4]Department of Plant and Environmental Sciences, NMSU, Las Cruces, NM 88003, USA

[5]Institute of Plant Sciences, Agricultural Research Organization, The Volcani Center, Rishon LeZion, Israel

[6]Biotechnology and Agroindustry Division, ENEA, Italian National Agency for New Technologies, Energy and Sustainable Development, Via Anguillarese, 301–00123 Roma, Italy

[7]These authors contributed equally to this article.

***Correspondence: Feng Cheng (chengfeng@caas.cn), Lihao Wang (wanglihao@caas.cn), Giovanni Giuliano (giovanni.giuliano@enea.it)**

https://doi.org/10.1016/j.molp.2022.09.021

ABSTRACT

Pepper (*Capsicum* spp.) is an important vegetable crop that provides a unique pungent sensation when eaten. Through construction of a pepper variome map, we examined the main groups that emerged during domestication and breeding of *C. annuum*, their relationships and temporal succession, and the molecular events underlying the main transitions. The results showed that the initial differentiation in fruit shape and pungency, increase in fruit weight, and transition from erect to pendent fruits, as well as the recent appearance of large, blocky, sweet fruits (bell peppers), were accompanied by strong selection/fixation of key alleles and introgressions in two large genomic regions. Furthermore, we identified *Up*, which encodes a BIG GRAIN protein involved in auxin transport, as a key domestication gene that controls erect vs pendent fruit orientation. The *up* mutation gained increased expression especially in the fruit pedicel through a 579-bp sequence deletion in its 5′ upstream region, resulting in the phenotype of pendent fruit. The function of *Up* was confirmed by virus-induced gene silencing. Taken together, these findings constitute a cornerstone for understanding the domestication and differentiation of a key horticultural crop.

Key words: *Capsicum*, variome, domestication, fruit orientation, pungency, fruit shape, sweet pepper, blocky fruit pepper, narrow fruit pepper

Cao Y., Zhang K., Yu H., Chen S., Xu D., Zhao H., Zhang Z., Yang Y., Gu X., Liu X., Wang H., Jing Y., Mei Y., Wang X., Lefebvre V., Zhang W., Jin Y., An D., Wang R., Bosland P., Li X., Paran I., Zhang B., Giuliano G., Wang L., and Cheng F. (2022). Pepper variome reveals the history and key loci associated with fruit domestication and diversification. Mol. Plant. **15**, 1744–1758.

INTRODUCTION

With $10.73 billion of gross production value in 2019 (http://www.fao.org/faostat/), pepper (*Capsicum* spp.) is the third most produced vegetable crop and a major component of spicy food, highly appreciated in the Americas, the Mediterranean area, east Asia, south Asia, and southeast Asia. Its pungency is conferred by capsaicinoids, primarily capsaicin and

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

dihydrocapsaicin (Srinivasan, 2016), and is sensed by a vanilloid receptor also involved in heat and pain perception (Jordt and Julius, 2002). The fruits of wild peppers are mainly pungent, small, nearly round, brightly colored, and erect, discouraging mammalian herbivores, which are sensitive to pungency, and favoring seed dispersal by birds, which have good color vision and an impaired capacity to sense pungency (Tewksbury and Nabhan, 2001; Jordt and Julius, 2002).

About 42 species have been described in the *Capsicum* genus, including five domesticated species: *Capsicum annuum* L., *Capsicum frutescens* L., *Capsicum chinense* Jacq., *Capsicum baccatum* L., and *Capsicum pubescens* Ruiz & Pavon (Carrizo García et al., 2016; Barboza et al., 2020). *C. annuum* is the most widely cultivated of the domesticated species. Archeological microfossil evidence (Perry et al., 2007) indicates that cultivated pepper species underwent distinct domestication events as early as 6000 years ago in primary diversity centers in South and Meso-America (Andrews, 1995; Perry et al., 2007; Aguilar-Melendez et al., 2009; Bosland and Votava, 2012; Carrizo García et al., 2016). Pepper was introduced from the West Indies into Europe in the late 15th and early 16th centuries, and it was then rapidly distributed to Africa and Asia, including China, where the earliest written record of pepper dates back to 1591 (Ming Period) (Andrews, 1995; Bosland and Votava, 2012; Zou et al., 2020; Tripodi et al., 2021). During domestication and breeding, non-deciduous peppers emerged with diverse fruit shapes, sizes, weights, pendent fruit orientation, and a range of pungency levels (Paran and van der Knaap, 2007). The change in fruit position from erect to pendent was selected during early domestication and provides an adaptation to increased fruit size and better protection from sun exposure and bird predation. It is thus a key agronomic trait in different fruit-bearing crops (Paran and van der Knaap, 2007). A more recent product of selection was the emergence of very large, blocky, non-pungent fruits (sweet bell peppers), whose earliest record dates to the 1700s (Bosland and Votava, 2012).

Studies exploiting the natural variability of pepper enabled the identification of several quantitative trait loci (QTLs) and candidate genes controlling capsaicinoid levels, such as *Pun1*, *pAMT*, *CaKR1*, and *Pun3* (Stewart et al., 2005; Han et al., 2018; Arce-Rodríguez and Ochoa-Alejo, 2019; Zhu et al., 2019), and fruit shape/size (*longifolia 1-like*) (Chaim et al., 2001; Colonna et al., 2019; Lee et al., 2020). By contrast, the molecular basis of other key fruit traits, such as narrow versus blocky pepper types, has not yet been described.

The large variations in fruit size, shape, weight, orientation, and pungency found in pepper germplasm offer an opportunity to explore the genomic events that underlie the diversification of these important agronomic traits and the temporal sequence in which they appeared. In spite of the availability of high-quality genomic sequences of several pepper species and accessions (Kim et al., 2014; Qin et al., 2014; Ou et al., 2018), the understanding of the molecular evolution of this crop is lagging behind that of its close relative, tomato. To fill this gap, we resequenced 347 accessions of 12 *Capsicum* species, characterized the major fruit traits in these accessions, and uncovered the genomic variations associated with these traits. Our findings reveal the major genomic events and key genes or loci that have shaped the present-day diversity of this important horticultural species.

RESULTS AND DISCUSSION

The main trajectories of *C. annuum* domestication and improvement

A total of 347 accessions from 12 species of *Capsicum*, collected from gene banks in Asia, the Americas, Africa, and Europe, 311 of which were *C. annuum*, were resequenced to an average depth of ~9×, generating 10.11 Tb of sequencing data (Supplemental Table 1). Because a reference genome with good contiguity was critical for efficient identification of the genetic variations in the analyzed samples, we updated the Zunla-1 reference genome by closing assembly gaps with the assistance of approximately 10× PacBio HiFi reads. The total size of the gaps was reduced from 83.37 Mb to 14.14 Mb, and 92.27% (97 082) of the gaps in 12 chromosomes were successfully closed. A total of 40 155 protein-coding genes were then predicted in the updated genome. Benchmarking Universal Single-Copy Orthologs (BUSCO) evaluation of the protein predictions achieved a score of 96.10%, indicating a significant improvement compared with the previous genome version (84.08%). Based on the updated genome, a variome map was obtained, including 18 372 022 single nucleotide polymorphisms (SNPs) and 802 875 insertions/deletions (InDels), with an accuracy of >95%, verified by Kompetitive Allele-Specific PCR (KASP) (Supplemental Table 2). Variants were uniformly distributed along the 12 chromosomes, with the exception of a genomic region of chromosome 9 that contained substantially more variants (Figure 1A), and they were about twice as abundant in intergenic regions as in gene bodies (Supplemental Figure 1A). The median heterozygosity of the accessions was 1.11% (Supplemental Figure 1B), and 56 182 SNPs and 3080 InDels caused changes in the protein sequences of coding genes (Supplemental Table 3). We also identified presence/absence variations (PAVs) and copy number variations (CNVs) (methods). A total of 20 563 PAVs and 229 CNVs were obtained, with median lengths of 7.30 kb and 4.20 kb, respectively (Supplemental Figure 2A and 2B). On average, 487 genes and ~84.14 Mb of sequences were affected by PAV/CNVs in each accession (Supplemental Figure 2C and 2D). Owing to limitations of the short-read mapping, these PAVs and CNVs may not represent the full spectrum of such variations in the pepper population.

We used 165 864 synonymous SNPs located in genes to investigate the phylogenetic relationships among the accessions. Different *Capsicum* species formed distinct branches (Figure 1B), and the 311 *annuum* accessions formed nine groups (Figure 1C): (I) the wild/ancestral group, which included two wild *C. annuum* var. *glabriusculum* and 10 ancestral accessions and was located immediately next to non-*annuum* species; (II) a group composed mainly of old landraces; (III) cultivars with diverse geographic origins; (IV) and (VI) blocky fruit peppers; (V) cultivars with diverse fruit types and origins; (VII) accessions from the northwest and north of China; (VIII) accessions from central China; and (IX) accessions from southwest China collected from high-altitude areas in Yunnan, Guizhou, Sichuan, and Tibet (Supplemental Figure 3).

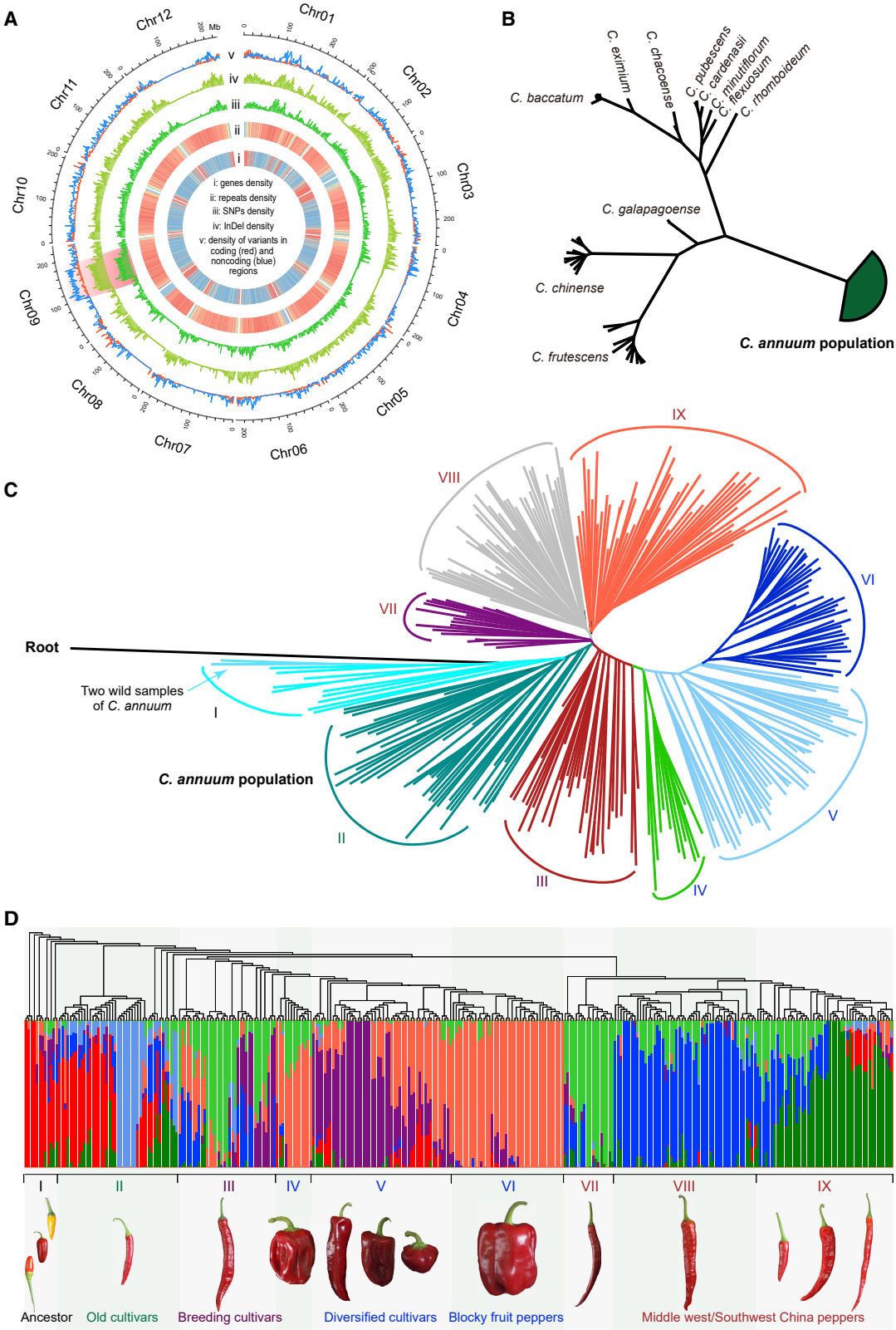

Figure 1. The pepper variome.

(A) Distribution of the variants in the pepper genome; red shading marks a region on chromosome 9 with an extremely high variant density.

(B) Phylogenetic tree of 347 resequenced accessions from 12 *Capsicum* species.

(C) Phylogenetic tree of the 311 *C. annuum* accessions. Different colored branches indicate the nine groups discussed in the text.

(D) Genetic admixture analysis of the nine *C. annuum* groups. Representative fruit types are shown below each group.

	Groups	I	II	III	IV	V	VI	VII	VIII	IX
F_{ST}	I		0.1546	0.1581	0.2952	0.2175	0.4086	0.2566	0.2567	0.1893
	II			0.1184	0.2874	0.2101	0.3612	0.1463	0.1530	0.1015
	III				0.1699	0.1223	0.2662	0.1093	0.1215	0.0953
	IV					0.2050	0.1935	0.3755	0.3353	0.2811
	V						0.2068	0.2847	0.2611	0.2263
	VI							0.4811	0.4010	0.3653
	VII								0.1401	0.0858
	VIII									0.0941
	π	0.2939	0.2860	0.2935	0.1643	0.2327	0.1418	0.2202	0.2371	0.2831

Table 1. Genetic differentiation (estimated by F_{ST}) between pairs of the nine pepper groups and genetic diversity (estimated by π) within each pepper group

Dark and medium orange colors denote the main origins (first column) of each pepper group (first row); light orange indicates partially mixed origins.

Groups I to IX represent the main domestication and breeding trajectories of pepper worldwide. The evolutionary relationships (Figure 1C and 1D) and genetic diversity (π) within each group and the genetic differentiation (F_{ST}) between groups (Table 1) both suggest the following scenario: group I is the ancestral group containing the early domesticates, as suggested by its position near the root of the tree and its high genetic diversity (π = 0.2939); group II represents old landraces, being closest to group I (F_{ST} = 0.1546), whereas group III represents later cultivars, being among the closest to group II (F_{ST} = 0.1184); both groups II and III exhibit a high genetic diversity (π = 0.2860 and 0.2935, respectively), suggesting the existence of either minor genetic bottlenecks or diversifying selection during the early steps of *C. annuum* domestication. The evolutionary relationships of groups I, II, and III were further supported by the genotypic compositions, with more ancestral alleles present in group I and more derived alleles from selection in group III (Supplemental Figure 4A). Groups II and III gave rise, directly or indirectly, to all other groups (Figure 1C and 1D and Table 1): directly to groups IV (blocky), V, VII, and VIII; and indirectly to groups VI (large fruited, blocky, derived from group IV) and IX (high-altitude Chinese peppers, derived from group VII).

Compared with groups IV and VI, all other groups exhibit relatively large π values, indicating the inheritance of a large variety of different alleles, or the action of diversifying selection, during their formation. Group V exhibits large variations in fruit shape and likely represents a transition group between traditional and blocky fruit peppers (Figure 1D). All groups present relatively high levels of genetic admixture (Figure 1D), confirming the absence of major genetic bottlenecks during domestication and subsequent breeding, with the exception of groups VI (large fruited, blocky peppers) and VIII (central China). Among the three Chinese groups (VII to IX), group IX exhibited the highest genetic diversity (π = 0.2831) (Table 1) and a predominant genetic component (represented by dark green in Figure 1D) present at significant levels in ancestral groups I and II, which may have been reintroduced in group IX to favor adaptation to high altitudes. The large genetic variation in group IX resulted in large fruit length variations, including a specific slim fruit type (Figure 1D).

Genomic signatures of the evolution of narrow fruit peppers

The two wild accessions (*C. annuum* var. *glabriusculum*) have short, very small, waterdrop-shaped, erect fruits with high capsaicinoid content (839–1146 mg/kg dry weight [DW]). Compared with the wild accessions, the early domesticates of group I exhibit a significant increase in fruit size, a large variation in fruit shape (olivary, short, conical), the appearance of pendent fruits (8 of 10), and very large variation in capsaicinoid content (0–1972 mg/kg DW), indicating strong diversifying selection exerted on these traits during early domestication (Supplemental Table 1).

The average fruit length increased further, from around 5.0 cm in group I to 8.0 cm in group II and 11.0 cm in group III, without a corresponding increase in fruit diameter, resulting in increasingly elongated fruit types (Figure 2A). The Chinese peppers in groups VII, VIII, and IX showed fruit lengths comparable to those of group III. The increase in length, which resulted in increased surface-to-volume ratio, probably served a dual purpose: making the early domesticates distinguishable from their wild ancestors, and facilitating air drying, a common technique still used to this day to conserve chili peppers. By contrast, capsaicinoid levels, after initial diversifying selection in the early domesticates, showed a multi-phasic trend with a slight increase in group II and a clear reduction in group III (Figure 2B). Pungency increased again later in groups VII, VIII, and IX, consistent with secondary selection for increased capsaicinoid levels in China, where spicy food is popular. Finally, pendent fruit types, which were already prevalent in groups I and II, became almost exclusive in groups III to IX (Supplemental Table 1).

Selective pressure generates genomic selection signals, measured as a reduction of nucleotide diversity (ROD) (Xu et al., 2012). Several genomic selection signals were detected in the pepper genome during early domestication (group I to group II), in particular on chromosomes 4, 8, 9, and 11 (Figure 2C and Supplemental Table 4). Three previously

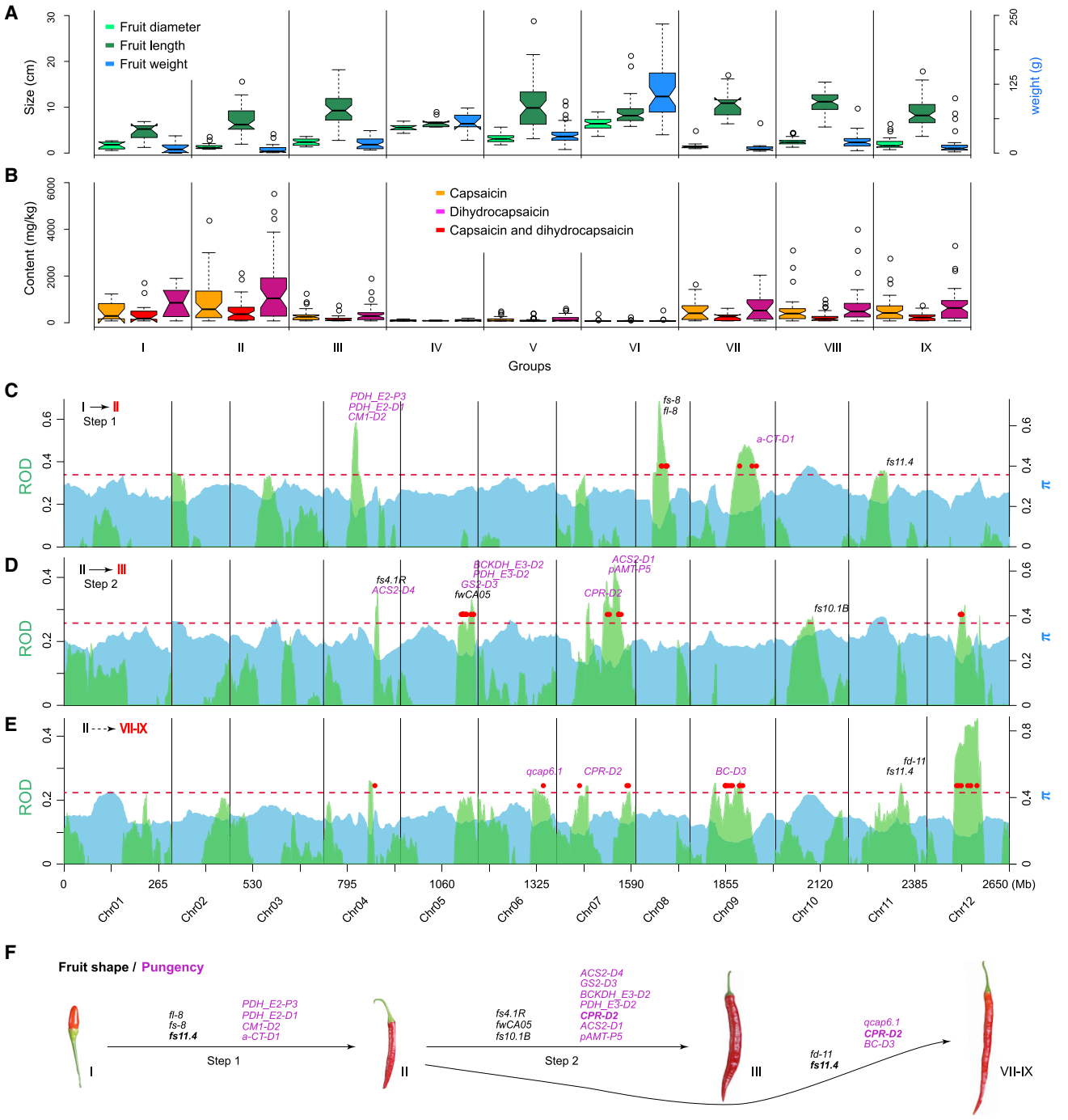

Figure 2. Selection for fruit shape and pungency in narrow fruit peppers.

Box plots of fruit diameter, length, and weight **(A)** and capsaicinoid content **(B)** in the nine *C. annuum* groups. Genomic selection signals detected by π (colored in blue) and ROD (colored in green) in the group I→II **(C)**, group II→III **(D)**, and group II→VII–IX **(E)** transitions.
(F) Genetic loci controlling fruit shape and pungency under selection during the main evolutionary transitions in narrow fruit peppers.

reported QTLs for fruit shape and length (Yarnes et al., 2012; Han et al., 2016) and four capsaicinoid biosynthesis genes (*PDH_E2-P3*, *PDH_E2-D1*, *CM1-D2*, and *a-CT-D1*) (Qin et al., 2014) are located in these genomic regions. There were 348 genes under selection in the transition from group I to group II, including the A-class flower homeotic gene *AP2-A* (*Capana04g002188*) (Supplemental Tables 5 and 6).

Genomic regions of five chromosomes were found to be under selection in the second transition (group II to group III) (Figure 2D and Supplemental Table 4). Two previously reported fruit shape QTLs (*fs4.1R* and *fs10.1B*), one fruit weight-related gene (*fw/CA05g10770*), and seven capsaicinoid biosynthesis genes (*BCKDH_E3-D2*, *PDH_E3-D2*, *GS2-D3*, *ACS2-D4*, *ACS2-D1*, *CPR-D2*, and *pAMT-P5*) are located in these regions,

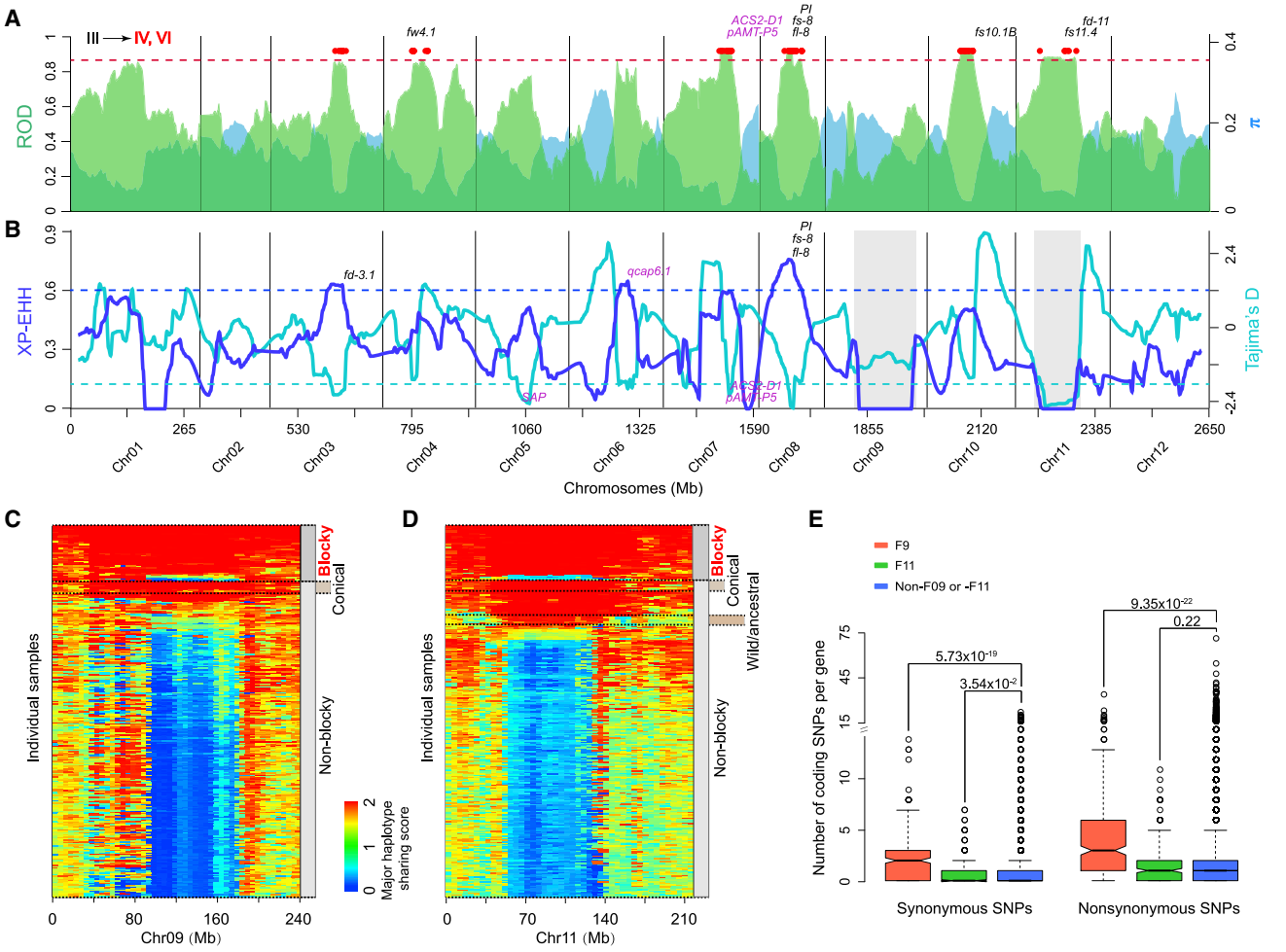

Figure 3. Genomic selection and introgression in blocky fruit peppers.

Selection signals detected by π and ROD **(A)** or by Tajima's D and XP-EHH **(B)** in blocky fruit pepper groups IV and VI compared with group III. Dashed lines indicate genomic regions under selection, identified as the top 5% outliers of ROD and XP-EHH or the bottom 5% outliers of Tajima's D. Genetic loci for fruit shape (colored in black) and pungency (purple) that overlap with genomic regions under selection are labeled at the corresponding positions. Major haplotype sharing (MHS) scores on chromosomes 9 **(C)** and 11 **(D)** of blocky and non-blocky fruit peppers, using blocky haplotypes as the reference.

(E) Differences in the numbers of coding SNPs in genes located in the F9 and F11 introgressed regions compared with all other genomic regions.

including a second A-class gene, *AP2-A* (*Capana02g000700*) (Supplemental Tables 5 and 6).

Thus, it appears that early evolutionary transitions in narrow fruit peppers (group I to II and group II to III) involved selection on large groups of candidate genes for fruit pungency and/or shape, probably relying on the vast genetic diversity for these traits that is present in these groups and on the absence of genetic bottlenecks. By contrast, transition from group II to the Chinese groups VII–IX involved selection on a narrower group of genes (Figure 2E and 2F, Supplemental Tables 5 and 6), consistent with the hypothesis that a genetic bottleneck was active during this transition, probably due to transport of a subset of the group II gene pool via sea or land (the silk road) to east Asia.

Recent emergence of sweet, blocky fruit peppers

Blocky fruit peppers (groups IV and VI) exhibit distinctive phenotypes, such as large increases in fruit diameter and weight, decreased variation in fruit shape, reduction of capsaicinoid levels to almost zero, and pendent fruit orientation, which is necessary to support the large fruit (Figure 2A and Supplemental Table 1). As mentioned above, they also exhibit very low genetic diversity (Table 1), consistent with their recent emergence (Bosland and Votava, 2012), and a higher fraction of fixed alleles, either ancestral or derived, compared with the other groups (Supplemental Figure 4A). Of the two groups, group VI was probably selected later, as suggested by its higher F_{ST} value with respect to group III, lower π value, higher proportion of fixed alleles, and also larger fruits. The linkage disequilibrium (LD) values of groups IV and VI are the highest in the whole *C. annuum* population, further confirming their recent emergence (Supplemental Figure 4B).

By comparing groups IV and VI with group III, several genomic selection signals were identified using the ROD measure (Figure 3A and Supplemental Table 4), overlapping with previously described QTLs for fruit shape, length, or weight (*fs-8*, *fs10.1B*,

fs11.4, *fl-8*, *fd-11*, and *fw4.1*) (Zygier et al., 2005; Borovsky and Paran, 2011; Yarnes et al., 2012; Han et al., 2016) and with two capsaicinoid biosynthesis genes (*ACS2-D1* and *pAMT-P5*) (Qin et al., 2014). Given the recent emergence of blocky fruit peppers, cross population extended haplotype homozygosity (XP-EHH) (Sabeti et al., 2007) and Tajima's D (Tajima, 1989) were used to find additional genomic selection signals (Figure 3B), which overlapped with the QTLs *fd-3.1* for fruit diameter, *SAP* for flower and ovule development, and *qcap6.1* for pungency. *Capana07g001005*, an *Agamous* family gene that regulates ovule development, *Capana10g000984* and *Capana10g001014*, which encode cyclin-dependent protein kinase regulators of the cell cycle, and *Capana05g000060*, a member of the IQD family that includes *SUN*, which regulates fruit shape in tomato (Xiao et al., 2008), were located in these genomic regions and found to be under strong selection (Supplemental Tables 5 and 6).

Two genomic regions, named F9 and F11, on chromosomes 9 and 11 showed very low XP-EHH values (Figure 3B). The two regions exhibited clear differences between blocky and non-blocky types in the depth of reads mapped to the reference genome, which is derived from a non-blocky pepper (Supplemental Figure 5A and 5B), suggesting that these two regions may derive from distant introgressions. To confirm this hypothesis, we determined the major haplotypes in the genomes of blocky fruit peppers and estimated their similarity to the total *C. annuum* population by calculating the major haplotype sharing score (MHS). Two regions with consistently lower MHS scores co-localized with F9 and F11 (Supplemental Figure 5C). In F9, all blocky types except three, as well as five conical fruit accessions from group V, were highly homologous to each other, whereas most (92.46%) of the other non-blocky types diverged (Figure 3C). In F11, all but four blocky types showed high homologies to each other, as well as to four conical fruit peppers from group V and two wild and five ancestral peppers from group I, whereas most (91.06%) of the other non-blocky types diverged (Figure 3D). These data, taken together, suggest that F11 probably originated from an introgression from a wild *C. annuum* that occurred in ancestral peppers of group I, persisted at low frequency in groups II and III, and was almost fixed in blocky fruit peppers. F9 is more divergent from the reference fragment than is F11, as the former has a higher frequency of coding SNPs compared with the rest of the pepper genome (Figures 1A and 3E). We further compared genotypes of loci in F9 with the released sequences of *C. annuum* var. *glabriusculum* (Qin et al., 2014) and built a phylogenetic tree to trace their evolutionary relationships. These results support the conclusion that F9 was introgressed from this wild *C. annuum* (Supplemental Figure 6). We also checked gene contents in the two introgressed fragments, and no enrichment of disease resistance gene clusters was found, indicating that the introgressions were not driven by gain of disease resistance.

Selection at a few key loci controls main transitions in pepper fruit evolution

We performed a genome-wide association study (GWAS) to locate genomic loci that were responsible for major traits during domestication and breeding selection of *C. annuum*, including fruit orientation, fruit shape, and pungency, using a linear mixed model (LMM) implemented in GEMMA with the genotype and phenotype datasets of the 311 *C. annuum* accessions. The LD values of the *C. annuum* population are greater than 500 kb (Supplemental Figure 4B), which makes it difficult to perform fine mapping of genes using population data alone.

Fruit orientation is an important agricultural trait in both vegetable crops and fruit trees, but its molecular basis is unknown. As mentioned above, fruit orientation transitioned from erect (up) in wild peppers to pendent (down) in domesticated large-fruited peppers. The *up* locus controls fruit orientation in pepper (Figure 4A), but the gene underlying this variation is unknown, although recent studies have further narrowed down the gene candidates (Hu et al., 2021; Solomon et al., 2021). We performed GWAS analysis for this trait and found a strong association signal on chromosome 12, where the *up* locus resides (Lefebvre et al., 1995) (Figure 4B). The most significant signal was in the promoter region of the gene *Capana12g000954*, which is expressed in the flower pedicel and the placenta of pepper fruit (Supplemental Figure 7A). *Capana12g000954* encodes a BIG GRAIN 1-like (BGL) protein whose rice ortholog is expressed in vascular tissues and mediates auxin transport (Liu et al., 2015). This gene was one of two genes previously considered as candidates for the control of pepper fruit orientation (Lee et al., 2020). A 579-bp deletion was detected in the 5′ upstream region of this gene in the pendent accessions; it showed an extremely significant association with the fruit orientation trait ($p = 6.00 \times 10^{-175}$) and was confirmed in a test population of 241 samples (Figure 4C, Supplemental Figure 7B and 7C). RNA sequencing (RNA-seq) and quantitative Real Time PCR (qRT-PCR) analyses found that this deletion was associated with high expression levels of the gene in pedicels of pendent fruits, but accessions with erect fruits exhibited low expression levels of the gene (Figure 4D). *BG1-like* genes have been implicated primarily in the control of organ size and yield in rice, Arabidopsis, and maize (Liu et al., 2015; Simmons et al., 2020). Additional growth-related traits such as plant height, tiller angle, and gravitropism, as well as stress tolerance, are affected by down- or upregulation of these genes. The function of *BG1-like* genes has not yet been determined in fruit crops. The novel putative role of *BG1-like* in control of fruit orientation in pepper may be mediated by differential distribution of auxin and by the level of gravitropism response in the pedicel.

We crossed a wild pepper accession (erect) with a blocky pepper accession (pendent) and obtained an F_2 population of ~360 individual plants. Bulked segregant analysis with whole genome resequencing (BSA-seq) identified a single significant signal on chromosome 12 (methods). Inspection of the genomic position of the peak ΔSNP-index signal found that it overlapped with the GWAS signal associated with *Capana12g000954* (Figure 4E). We further verified the function of the *BG1-like* gene through virus-induced gene silencing (VIGS) in an accession that produces pendent fruits (methods). Plants infected with the TRV2::*up* vector showed erect fruits compared with the pendent fruits of the wild-type accession and of the same accession infected with the empty TRV2 vector (Figure 4F and Supplemental Figure 8). The results of qRT-PCR showed that

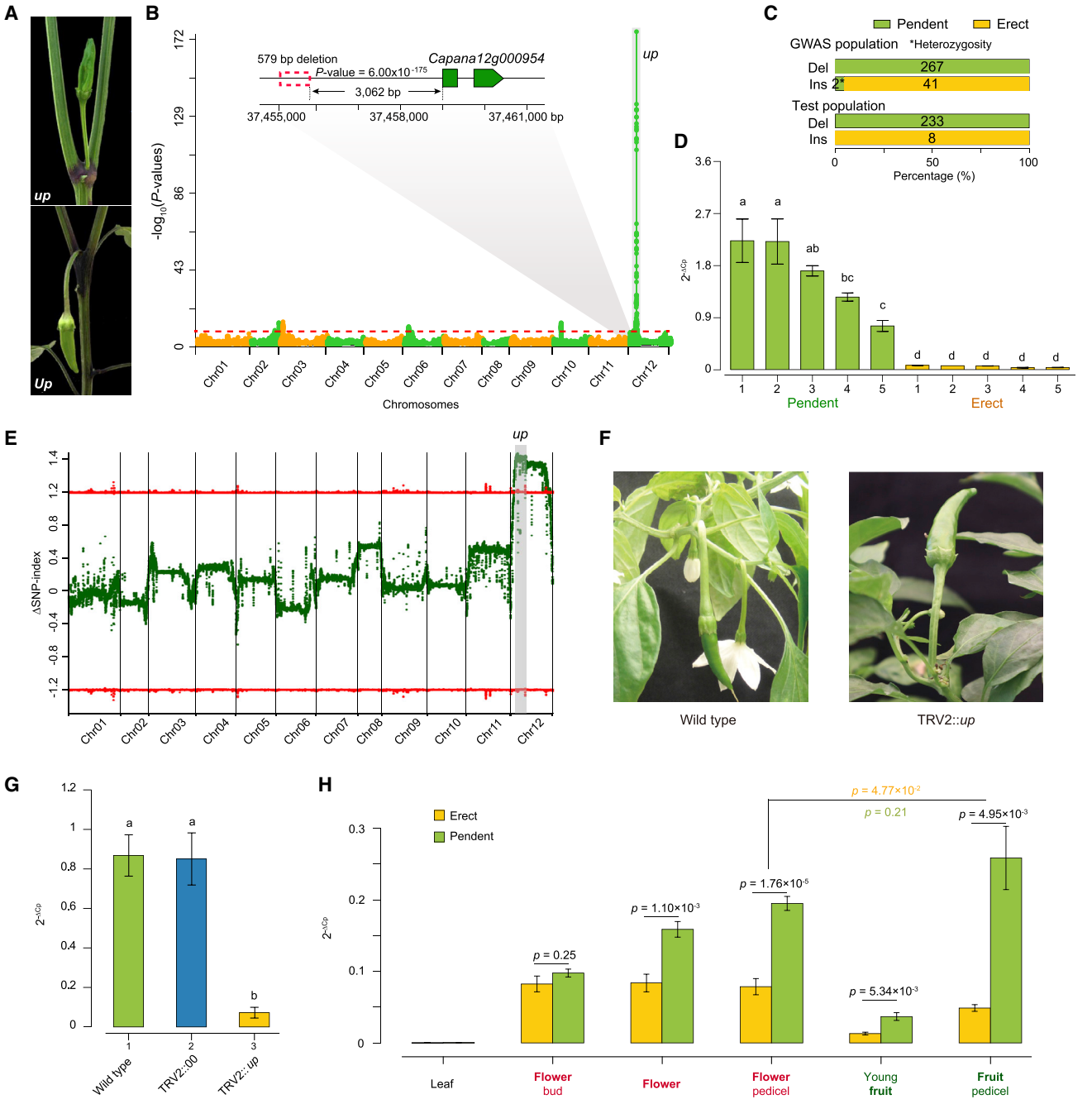

Figure 4. Identification of the *Up* gene that controls pepper fruit orientation.

(A) *C. annuum* plants segregating for the *up* mutation, which controls fruit orientation.

(B) GWAS association signal of the 579-bp deletion in the promoter region of *Capana12g000954* (*up*); **(C)** frequencies of the deletion in the GWAS and test populations.

(D) qRT-PCR analysis of *Up* expression in young fruit pedicels with different fruit orientations (pendent versus erect). Vertical bars: standard error.

(E) BSA-seq ΔSNP-index signal based on the F_2 population of pendent and erect accessions.

(F) The wild-type pendent fruit of the pepper accession "Changyang chili" and the erect fruit of "Changyang chili" after VIGS treatment with TRV2:*up*; **(G)** qRT-PCR analysis of *Up* expression in fruit pedicels of wild-type and VIGS samples showing pendent and erect fruits. Vertical bars: standard error. TRV2::00 and TRV2::*up* denote plants of "Changyang chili" subjected to VIGS with the empty vector and *up* gene (*Capana12g000954*), respectively.

(H) The expression of *Capana12g000954* (qRT-PCR) in different tissues of pepper samples with pendent or erect fruits. The gene is expressed in flower-related organs in both samples, but it is significantly upregulated in the pendent fruit sample, especially in the fruit pedicel.

the expression of the *BG1-like* gene was suppressed in pedicels of erect fruits infected with the TRV::*up* vector but not in pedicels of uninfected pendent fruits or those infected with the empty TRV2 vector (Figure 4G). We performed RNA-seq of multiple organs, including leaf, stem, root, flower, and fruit, from both the erect and pendent fruit accessions carrying the *Up* and *up*

genes, respectively, in an attempt to determine why the wild type of *Up* showed no/low expression and likely no function, whereas a deletion mutation in the 5′ upstream region of *up* was associated with activated expression and probably promoted the development of pendent fruit. We found that *Up* exhibited a high expression level only in the flowers of the erect fruit accession (Supplemental Table 7), indicating that the wild-type gene functions primarily in flowers. Because all peppers develop downward- or lateral-facing flowers, this result also suggests that the wild-type *Up* may play a role in regulating the development of downward-facing flowers, which facilitates self-crossing in pepper plants. The mutation-type *up* was highly expressed in both the flowers and fruits of the pendent fruit accession (Supplemental Table 7), implying that the mutated *up* had gained additional function through increased expression in fruit, thereby resulting in pendent fruits. Moreover, qRT-PCR analyses on leaf, flower bud, flower, flower pedicel, young fruit, and fruit pedicel (Figure 4H) showed that wild-type *Up* was stably expressed in the flower-related organs but decreased significantly in expression after the flowers developed into fruits, consistent with downward-facing flowers eventually growing into erect fruits after pollination. However, the expression of the mutated *up* was higher and increased continuously as the flowers developed into fruits, reaching its highest level in the fruit pedicel, corresponding to the maintenance or increase of downward orientation as flowers developed into pendent fruits. Interestingly, the 579-bp structure variation (SV) is located at a LINE-type retrotransposon (Supplemental Figure 9A). Transposable elements are known to be silenced through methylation to suppress their transposition activity, a process that will also suppress the expression of their adjacent genes—the so-called position effect (Hollister and Gaut, 2009; Hirsch and Springer, 2017). It is likely that the SV (sequence deletion) in pendent accessions attenuated this suppression through partial deletion of the methylated LINE sequence, thereby promoting the expression of *up* in flowers and fruit (Figure 4H). We thus assayed the methylation status of the SV sequence (methods) and found that cytosines, especially at CpG and CHG sites, were indeed heavily methylated (Supplemental Figure 9B), supporting the scenario in which deletion mutation of the SV upregulates the expression of the *up* gene by removing a large portion of the methylated retrotransposon sequence. These results indicate that the *up* gene has gained functions, such as increased expression, especially in the fruit pedicel, which have led to the novel phenotype of pendent fruits in domesticated peppers.

Fruit shape is an important agronomic trait and is controlled by a conserved network of interacting gene products in distantly related plants (van der Knaap and Østergaard, 2018; Wu et al., 2018). In pepper, fruit shape is extremely varied and serves the dual purpose of distinguishing different cultivars from each other and facilitating air drying for long-term storage of elongated pepper types. We found strong, overlapping GWAS signals for fruit shape index, length, and diameter on chromosome 3 and named the associated locus *fsi*. The most significant SNP of *fsi* overlapped with previously mapped QTLs for fruit shape and length (*fs-3.1*, *fl-3.2*) and is consistent with a previously mapped SNP (Colonna et al., 2019). This SNP caused a nonsynonymous Ile340Thr mutation in the *Capana03g002426* (*TRM25*) gene (Figure 5A and 5B and Supplemental Figure 10A), which encodes a TONNEAU 1 Recruiting Motif (TRM) protein. TRM proteins are part of a protein complex that interacts with microtubule arrays and controls cell division patterns, and they are well-known regulators of fruit shape in tomato and cucumber (Wu et al., 2018). The candidate gene *TRM25* was expressed in the early stage of pepper fruit development (Figure 5C and Supplemental Figure 10B). Examination of *TRM25* expression in accessions with different fruit types revealed that its expression was not correlated with specific fruit types. This result indicates that *TRM25* may regulate fruit shape through functional mutation rather than transcriptional alteration (Supplemental Figure 10C), consistent with the fact that the most strongly associated SNP caused an amino acid change rather than a regulatory mutation in the promoter region. An additional gene, *Capana09g001401*, located in the chromosome 9 introgression in blocky fruit types, was highly expressed in the pericarp of non-blocky fruit peppers but not of blocky fruit peppers (Supplemental Figure 11). *Capana09g001401* encodes a glycine-rich cell wall structural protein (*GRP*) that is associated with cell elongation/expansion and differentiation in various tissues in rice (Xu et al., 1995). The gene was found to be under selection in blocky fruit types (Supplemental Table 6) and is thus a strong candidate for the control of blocky fruit peppers.

Several genes that control pungency in pepper have been identified, encoding either structural genes in the capsaicin biosynthesis pathway or, in one case, a transcriptional regulator (Stewart et al., 2005; Han et al., 2018; Arce-Rodríguez and Ochoa-Alejo, 2019; Zhu et al., 2019). A GWAS analysis based on a binary classification of capsaicinoid presence or absence in all the *annuum* pepper accessions identified a strong association signal on chromosome 2 ($p = 2.45 \times 10^{-72}$) (Figure 5D and 5E) in the vicinity of *Pun1* (*Capana02g002340*), which was previously reported as a major heat regulation gene that mediates the last step in capsaicin biosynthesis (Stewart et al., 2005). The expression of *Pun1* was positively correlated with capsaicinoid content (Supplemental Figure 12A). We surveyed the nucleotide deletion reported to be associated with *Pun1* using read mapping information (Supplemental Figure 12B), and a loss-of-function deletion associated with the recessive allele *pun1* was found. This structural variation has the most significant association ($p = 2.23 \times 10^{-308}$) with the heat trait (Figure 5D), thus supporting the conclusion that *Pun1* is the causal gene of these local association signals. In addition, GWAS analysis of narrow-fruited peppers showed a strong association signal for pungency variation (named *punv*) on chromosome 6, where the candidate gene *Capana06g001204* is located (Supplemental Figure 13A and 13B). *Capana06g001204* encodes a phospholipid-flipping ATPase (flippase) and is highly expressed in the pericarp and placenta during the middle and late developmental stages of pepper fruit (Supplemental Figure 13C). Flippases translocate lipids (mainly phospholipids) across biological membranes using the hydrolysis of ATP, and they are involved in a series of physiological responses (Nintemann et al., 2019). The role of this protein in the control of pungency may be directly related to capsaicinoid transport across membranes, generation of blisters to accumulate capsaicinoids in the placental tissue of pungent peppers, or membrane protection against the

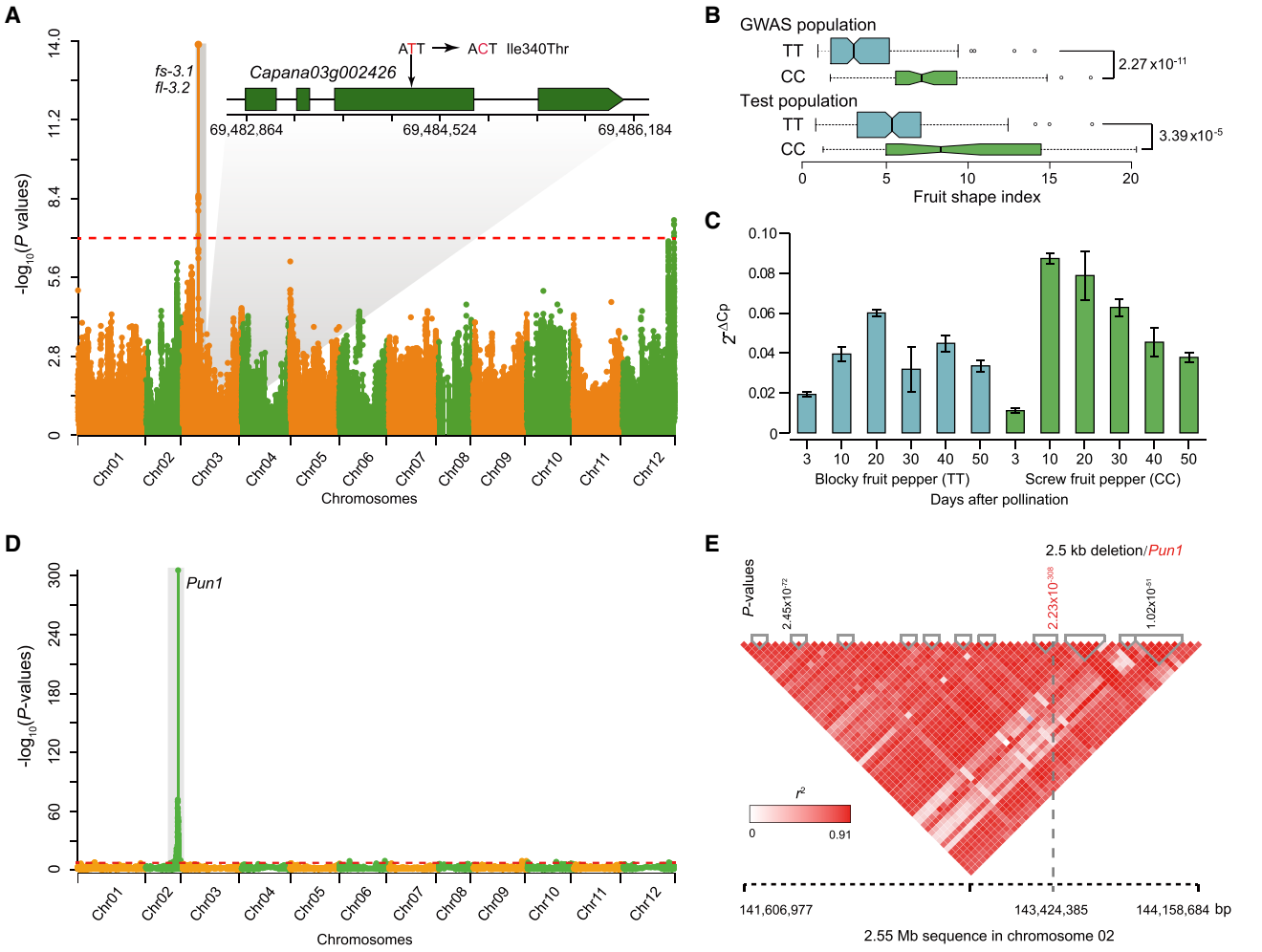

Figure 5. Genomic loci and candidate genes associated with elongated fruit shape and the pungency trait.

(A) GWAS identified the locus *fsi* that showed an association signal with fruit shape index, and there was an Ile340Thr mutation in the candidate gene *Capana03g002426* (*TRM25*); a red dashed line denotes the threshold at 1×10^{-7}.

(B) Differences in fruit shape index in genotypes carrying the nonsynonymous mutation in GWAS and test populations.

(C) The expression of *Capana03g002426* (qRT-PCR) in the pericarp of blocky fruit pepper (TT genotype) and screw fruit pepper (CC genotype) in different stages of fruit development.

(D) GWAS identified the previously reported *Pun1* gene that showed an association signal with pungency (binomial data).

(E) Linkage disequilibrium among SNP loci that had extremely significant association signals (p-value < 1.00×10^{-30}) with the pungent trait.

destabilizing effects of high capsaicin concentrations (Aranda et al., 1995).

The key temporal sequence of pepper fruit domestication and diversification

Analysis of the pepper variome enables temporal reconstruction of the key events that have shaped the high diversity of today's peppers. Starting from fruit orientation, the 579-bp deletion in the *up* promoter associated with pendent fruits was already present in a high proportion of ancestral group I, increased in groups III to IX, and reached complete fixation in blocky groups IV and VI (Figure 6A). Interestingly, the genotype of the *punv* locus that controls fruit pungency shows a very similar trend to *up*, reaching 100% frequency in group III and remaining high thereafter. In the analyzed population, the key variants of the two genes show very high association ($p = 2.32 \times 10^{-11}$), which is not due to physical linkage, as the two genes map to chromosomes 6 and 12, respectively. Similarly, the F9 and F11 introgressions associated with the blocky fruit type were found at different frequencies (8.33% and 58.33%, respectively) in group I, but thereafter they showed a very high association in all groups ($p = 1.12 \times 10^{-21}$).

Strong associations between unlinked loci can be explained by a number of different scenarios: (1) reduced gene flow of the populations that contain the associated regions with respect to the general gene pool; this hypothesis is unlikely in the present case, as it would influence the LD of additional unlinked loci, which does not seem to be the case; (2) simultaneous selection for two different traits encoded by the associated loci; this seems to be the case for *up* and *punv* during early domestication; and (3) cooperative action of the associated unlinked loci in determining a single phenotype; this seems to be the case for the F9 and F11

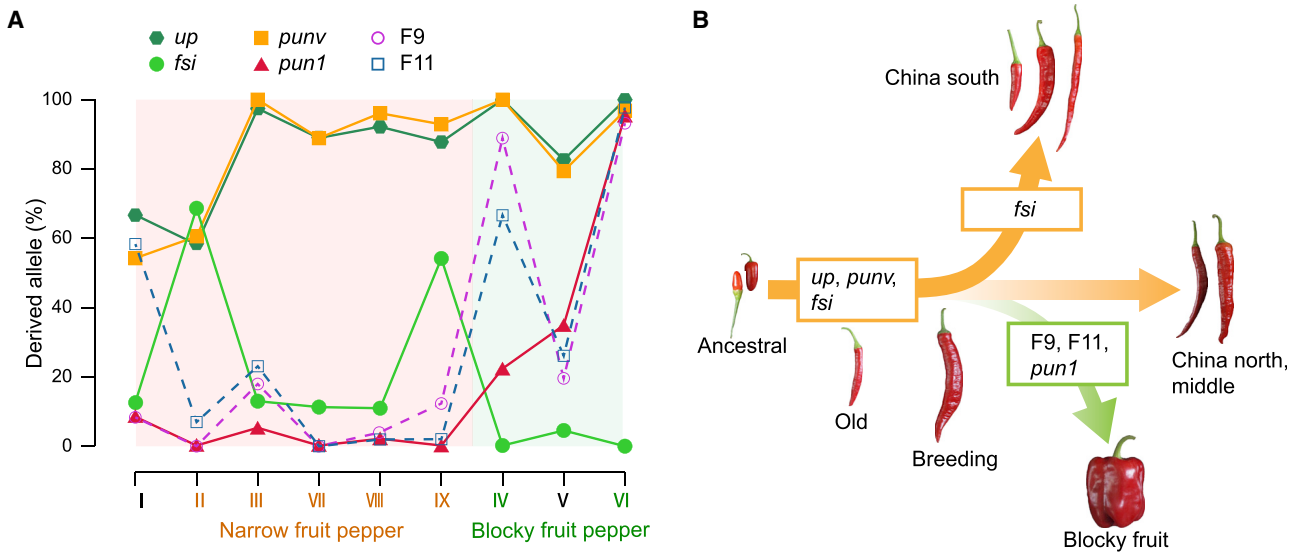

Figure 6. Key events that have shaped pepper fruit domestication and diversification.
(A) Frequency distributions of the genotypes at the *up* and *fsi* loci, which control fruit orientation and shape, respectively, the *punv* and *pun1* loci, which control fruit pungency, and the F9 and F11 introgressions, which are associated with blocky fruit shape, in the nine *C. annuum* groups.
(B) Chronodiagram of the key genetic events that have affected fruit characteristics during pepper domestication and differentiation.

introgressions, which are almost always found together in blocky fruit types.

By contrast, the knockout *pun1* allele that controls fruit pungency was extremely rare in narrow pepper groups I to IX, and its frequency increased progressively in groups IV and VI (blocky) (Figure 6A). The most likely explanation is that early selection for blocky fruits accidentally co-opted the *pun1* sweet pepper allele in a subset of group IV accessions and that the associated "sweet" phenotype was subsequently selected and reached complete fixation in group VI, which is the most recent blocky fruit group and exhibits the largest fruits. This selection probably accompanied a switch in the culinary uses of pepper, from a spice in which small, elongated, easy-to-air-dry fruits prevailed, to a large-fruited, fresh-market vegetable for consumption in raw or cooked form.

On the basis of the above data, we present the following model for pepper fruit domestication and diversification (Figure 6B): (1) all alleles found in one or more later groups were pre-existing in the ancestral group; (2) during early domestication (groups I→III), the *up* allele frequency increased to almost complete fixation, mediating the conversion from erect to pendent fruits; (3) the F9 and F11 introgressions were co-opted, leading to the appearance of blocky fruit peppers (groups IV and VI), which also became sweet through the increase and fixation of *pun1*. By contrast, the genetic circuits controlling fruit elongation and pungency in narrow fruit peppers appear to be more complex: (4) fruit elongation between groups in groups II and IX was primarily mediated by the *fsi* genotype, whereas in other groups the primary contribution appears to be from additional genes (Figure 2F); (5) similarly, in spite of the low frequency of *pun1* in group III, this group has lower capsaicinoid content than group II, which is associated with complete fixation of *punv* and also probably accompanied by selection at other loci that control capsaicinoid content (Figure 2F).

In conclusion, the variome map of pepper described here revealed the main genomic events underlying the initial transition from small, almost round, erect, pungent fruits, to larger, more elongated fruits with a greater variation in capsaicinoid content, followed by further diversification in fruit shape, pungency and the recent appearance of sweet, blocky peppers. These findings greatly expand our understanding of pepper fruit domestication and diversification and will contribute to further breeding and improvement of this important crop.

METHODS

Pepper sample collection

A total of 347 pepper accessions were used in this study, including the core collection of the China Gene Bank (Gu et al., 2019; Wang et al., 2018), core breeding inbred lines of the Institute of Vegetables and Flowers, and international typical local varieties and wild species. In total, 12 species were collected, including 309 *C. annuum* (cultivated species), two *C. annuum* var. *glabriusculum* (wild species), 10 *C. frutescens* (cultivated species), 10 *C. chinense* (cultivated species), five *C. baccatum* (cultivated species), one *C. baccatum* var. *baccatum* (wild species), two *C. pubescens* (cultivated species), one *C. cardenasii* (wild species), two *C. chacoense* (wild species), one *C. eximium* (wild species), one *C. flexuosum* (wild species), one *C. galapagoense* (wild species), one *C. minutiflorum* (wild species), and one *C. rhomboideum* (wild species). Samples originated from the Americas, Europe, Asia, and Africa and were obtained from several gene banks.

Investigation of agricultural characteristics

The experimental materials were planted five times: in a greenhouse in Langfang during spring 2016; in a greenhouse in Beijing, a plastic greenhouse in Dehong, and an open field in Urumqi during spring 2017; and in an open field in Urumqi during spring 2018. Three replicates were planted in each experiment. In the Langfang and Beijing experiments, two to three plants were planted, and two fruits were measured in each replicate; in the Dehong and Urumqi experiments, 10 plants were planted, and 10 fruits were measured in each replicate.

The following mature fruit traits were measured: (1) fruit length: the distance from the pedicel attachment to the fruit apex; (2) fruit diameter: the maximum width; (3) fruit shape index: the ratio of fruit length to fruit diameter; (4) fruit orientation: if the fruit pointed upward, it was recorded as erect, otherwise it was recorded as pendent; and (5) the content of capsaicin and dihydrocapsaicin in dried fruits without seeds and pedicels, as measured by ultra-performance liquid chromatography (UPLC).

For measurements of capsaicin and dihydrocapsaicin, mature fruits without seeds and pedicels were dried at 55 to 60°C for about 2 days, then ground and processed according to an improved agricultural industry standard (NY/T 1381-2007, Determination of Capsaicin by High-Performance Liquid Chromatography [HPLC]). For each determination, 0.2 g (accurate to 0.0001 g) of sample powder was extracted with 25 ml of methanol–tetrahydrofuran (1:1) solution (HPLC grade; Sigma, USA), shaken, extracted in an ultrasonic extractor for 30 min in a 60°C water bath, and then filtered. The filter residue and filter paper were re-extracted with 25 ml methanol–tetrahydrofuran solution in an ultrasonic extractor for 10 min and filtered, and then the above operation was repeated. The three collected filtrates were combined, concentrated on a rotary evaporator (70–75°C water bath) to about 30 ml, and then transferred to a 50-ml volumetric flask, to which methanol was added to produce the final volume. After filtration through a 0.22-μm organic phase filter membrane, chromatographic analysis was performed. The capsaicin content in the test solution was diluted to 0.13–160 mg/l, and the dihydrocapsaicin content was diluted to 0.04–160 mg/l. The filtrate was injected into a UPLC system (Waters UPLC, USA) with an ACQUITY UPLC BEH C18 column (1.7 μm, 2.1×100 mm; Waters, USA) and a Photo Diode Array Detector. A mixture of methanol and distilled water (70:30) was used as the mobile phase. The flow rate was 0.20 ml/min. The injection volume was 2 μl, and the column temperature was 30°C. The detection wavelength was 280 nm, and the retention times were 3.276 min for capsaicin and 4.395 min for dihydrocapsaicin. Capsaicin and dihydrocapsaicin external standards (Sigma, USA) were prepared as 200 mg/l stocks in absolute methanol. The stock solution was diluted with methanol to obtain a series of standard working solutions of 100 mg/l, 50 mg/l, 20 mg/l, 10 mg/l, 1 mg/l, and 0.2 mg/l, then measured under the liquid chromatography conditions above. With the concentrations of capsaicin and dihydrocapsaicin as the ordinates and the corresponding peak area integrals as the abscissa, the standard curves and linear regression equations were calculated. After the prepared test solution was measured, it was calibrated at multiple points and quantified by the peak area integral values. Finally, the capsaicin and dihydrocapsaicin contents were converted into mg/kg DW.

Sequencing, variant calling, and annotation

DNA libraries for all pepper accessions were prepared according to the manufacturer's instructions of the NEBNext Ultra DNA Library Prep Kit for Illumina. All the libraries (300-bp insert size) were sequenced on the Illumina Solexa platform, and 150-bp paired-end reads were generated. Trimmomatic v0.33 was used to trim the raw Illumina fastq reads and remove adapters based on the manufacturer's adapter sequences to obtain clean reads. Clean reads were mapped to the Zunla-1 2.0 genome (Qin et al., 2014) using BWA 0.7.17 (Li and Durbin, 2009) with the following parameters: mem -M -A 1 -B 4 -O 6 -E 1. SAMtools was used to sort the sam files and convert them into bam files. The function MarkDuplicates integrated into GATK (version 4.1.7) was used to deduplicate the bam files. Variants were then called using the HaplotypeCaller function of GATK with default parameters, and raw VCF files were generated. Variant loci that were not genotyped in more than 40% of the 311 *C. annuum* accessions, as well as loci with a heterozygosity ratio >10% and a minor allele frequency <5% in the *C. annuum* population, were further filtered using an in-house Perl script. The genotype information for these remaining reliable variant loci in the 311 *C. annuum* accessions and 37 accessions of other *Capsicum* species were then obtained for further data analysis. By doing this, we reduced the bulk and biased variants resulting from the large genetic differences between the genomes of *C. annuum* and the other *Capsicum* species. ANNOVAR (Wang et al., 2010) was used to annotate variant functions.

Phylogenetic and population analysis

To select neutral and reliable SNP variants for investigation of the phylogenetic relationships among the resequenced pepper accessions, genic SNPs that were annotated as synonymous variations by ANNOVAR were extracted (Wang et al., 2010). It was determined that 165 864 such SNPs were present in the 347 resequenced pepper accessions. Distance matrices with genotype data for all 347 pepper accessions were generated. Within the 311 *C. annuum* accessions, the pairwise allele sharing distance was calculated for each pair of accessions, and the data in the distance matrix were used to construct a neighbor-joining phylogenetic tree with PHYLIP 3.69 (Retief, 2000).

ADMIXTURE software (version 1.3.0) (Pritchard et al., 2000) was used to perform genetic structure analyses using 160 000 randomly selected SNPs from the 311 *C. annuum* accessions. ADMIXTURE was run with k = 2–15 clusters and default parameters. Finally, k = 7 (seven major genetic components, Figure 1D), which showed the smallest CV error, was selected to represent the population genetic features of *C. annuum*.

Genome diversity and selection sweeps

Two measures were used to estimate the genomic diversity of the *C. annuum* population: π and ROD. π measures genomic diversity (Tajima, 1983) by computing the average difference per locus for each pair of accessions, whereas ROD is a measure based on π that estimates the diversity reduction of a sample group with respect to a control sample group (Xu et al., 2012). In this study, ROD detected genomic regions under group-specific selection compared with the parental group.

Akey's F_{ST} was used to calculate the pairwise genomic differentiation between two groups of samples (Akey et al., 2002), thereby evaluating the strength of genomic divergence between the two compared groups. Values ranged from 0 (no differentiation) to 1 (complete or fixed differentiation).

XP-EHH was used to detect signals of recent positive selection (group-specific selection signatures) (Sabeti et al., 2007) in the blocky fruit group of *C. annuum* using a combination of Beagle (version 5.1) (Browning et al., 2018) and selscan (version v1.3.0) (Szpiech and Hernandez, 2014) software with default parameters. Large XP-EHH values indicate unusually long haplotypes in the derived group under selection compared with the control group. Tajima's D was used to estimate genomic regions under negative selection/purifying selection, and outliers were detected based on a threshold of less than −2 (Tajima, 1989).

To obtain the strongest selection signals from the large genome of pepper (3.6 Gb), which is larger than most sequenced crop genomes, the calculated values of these population genomic measures were further averaged using a 10-Mb window with a 1-Mb increment sliding across the whole pepper genome. An exception was Tajima's D, which was calculated directly in each sliding window. The top 5% of ROD and XP-EHH values were considered to be outlier signals of genomic regions under selection. If two outlier windows overlapped, they were then merged into one region.

MHS analysis

The major haplotypes in the blocky fruit groups were first determined using a 10-Mb window with a 1-Mb increment sliding across the 12 chromosomes of the pepper reference genome. In total, there were 548 such windows. In each window, the major allele frequency of each SNP in the blocky fruit pepper was calculated, and the SNP loci with a value ≥0.8 were counted and collected. If more than 80% of the SNPs in the window were counted, then these major alleles comprised the major haplotype of

the blocky fruit groups in the analyzed window. These major haplotypes were kept and compared with all of the resequenced *C. annuum* accessions. The average score of allele sharing between the major haplotype and each of the 311 accessions in the analyzed window was calculated and considered to be the MHS score. The MHS then estimated the relationship on the origin of the blocky fruit pepper-specific haplotype in the pepper population.

GWAS analysis

GEMMA (Zhou and Stephens, 2012) was used to perform GWAS. The relatedness (kinship) matrix was first calculated by GEMMA and used to correct the pepper population stratification in the 311 accessions of *C. annuum*. An LMM analyzed the correlation of the population, and p-values of all SNPs to each trait were generated by GEMMA. The GWAS signals of traits analyzed in this study were replicable in the five phenotypic trials, and GWAS results on the averaged values of the five trials were used to construct Manhattan plots.

Virus-induced gene silencing

VIGS was used to verify the function of the *up* gene *Capana12g000954*. We selected a specific sequence region (4–303 bp) of 5′ *Capana12g000954* to construct the VIGS vector. We confirmed the sequence specificity by BLAST searching of the Zunla genome databases. A 300-bp fragment of *Capana12g000954* was cloned from a pepper cDNA template using the gene-specific primers F-TGAGTAAGGTTACCGA ATTCCAAGCTGGCCAAGAGAAGAA and R-GGAGGCCTTCTAGAGAAT TCCCTCCAGGAGAAATCTGCTTA, containing an *EcoRI* site. The resulting product was inserted into the *EcoRI* site of pTRV2 to construct TRV2::*up*. A fragment of about 100 bp of the *CaPDS* gene was inserted into pTRV2 to construct the control vector TRV2::*CaPDS*. The tobacco rattle virus (TRV)-based vectors TRV1 and TRV2::*rec* were transformed into *Agrobacterium tumefaciens* (strain GV3101) for infiltration. Bacteria were grown and resuspended in infiltration solution at an OD_{600} of about 4.0. After standing at room temperature for 3 h, the *Agrobacterium* suspensions containing the TRV1 vector and one of the TRV2 vectors (TRV1+TRV2; TRV1+TRV2::*CaPDS*; TRV1+TRV2::*up*) were mixed at a 1:1 ratio and infiltrated. Two-month-old plants of the accession "Changyang chili" (pendent fruit) were used for the VIGS experiments. The plant growth environment was 23 ± 2°C, the relative humidity was 60%–80%, and the light regime was 16 h/8 h (light/dark). After infiltration, the plants were placed in the dark, cultured for 24 h, and then returned to a 16 h/8 h (light/dark) regime. Five plants were used for each of the three groups (non-infiltrated, TRV2::00, and TRV2::*up*). Only plants infiltrated with TRV2::*up* showed fruits with an erect orientation 2 months after infiltration. To investigate the expression of the *up* gene, total RNA was isolated from pendent fruit pedicels of wild-type and TRV2::00 plants and from erect pedicels of TRV2::*up* plants. Gene expression analysis was performed by qRT-PCR as described above. Four biological replicates (pedicels) and four technical replicates were performed.

Data availability

The raw resequencing data from this study have been deposited in the Genome Sequence Archive at the National Genomics Data Center, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number CRA003831 and are publicly accessible at https://bigd.big.ac.cn/gsa.

A complete version of the Materials and Methods is available in the supplemental information.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Molecular Plant Online*.

FUNDING

This work was supported by the China National Key Technology Research and Development Program (2016YFD0100200, 2016YFD0101700, 2018YFD1000800, and 2020YFD1001100), the National Natural Science Foundation of China (NSFC grants 31722048, 31972411, 3170110879, 32102382, and 32102386), Central Public-interest Scientific Institution Basal Research Fund (Y2022GH04, IVF-BRF2021015), the Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-IVFCAAS), the Earmarked Fund for the China Agriculture Research System (CARS-25), the CAAS-GAAS Coordinated Innovation Project of the Chinese Academy of Agricultural Sciences (2019), the Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture and Rural Affairs, P.R. China, and the Horizon 2020 G2P-SOL project (grant no. 677379) from the EU.

AUTHOR CONTRIBUTIONS

L.W. and F.C. conceived and designed the study. F.C., K.Z., and Y.C. analyzed the data. Y.Y., Y. Jing, Y.M., and X.W. participated in bioinformatics and GWAS analyses. Y.C., X.G., H.W., B.Z., V.L., and X. Li collected and analyzed the samples. Y.C., H.Y., and H.Z. investigated the agricultural characteristics. H.Y., S.C., W.Z., Y. Jin, and D.A. performed molecular experiments. D.X. and X. Liu performed the measurements of capsaicin and dihydrocapsaicin. Z.Z. and B.Z. helped with the data interpretation. R.W., P.W.B., I.P., and V.L. participated in analysis and made important intellectual contributions to the draft. F.C., L.W., G.G., Y.C., K.Z., and H.Y. mainly wrote the manuscript. All authors reviewed and revised the manuscript.

ACKNOWLEDGMENTS

We thank Prof. Xiaowu Wang for his helpful suggestions. The authors declare no competing interests.

Received: August 18, 2022
Revised: September 22, 2022
Accepted: September 26, 2022
Published: September 28, 2022

REFERENCES

Aguilar-Melendez, A., Morrell, P.L., Roose, M.L., and Kim, S.C. (2009). Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum annuum*; Solanaceae) from Mexico. Am. J. Bot. **96**:1190–1202.

Akey, J.M., Zhang, G., Zhang, K., Jin, L., and Shriver, M.D. (2002). Interrogating a high-density SNP map for signatures of natural selection. Genome Res **12**:1805–1814.

Andrews, J. (1995). Peppers: The Domesticated Capsicums (University of Texas Press). https://books.google.com.hk/books?id=SsjvX31EMekC.

Aranda, F.J., Villalaín, J., and Gómez-Fernández, J.C. (1995). Capsaicin affects the structure and phase organization of phospholipid membranes. Biochim. Biophys. Acta **1234**:225–234. https://doi.org/10.1016/0005-2736(94)00293-X.

Arce-Rodríguez, M.L., and Ochoa-Alejo, N. (2019). Biochemistry and molecular biology of capsaicinoid biosynthesis: recent advances and perspectives. Plant Cell Rep. **38**:1017–1030. https://doi.org/10.1007/s00299-019-02406-0.

Barboza, G.E., García, C.C., Scaldaferro, M., and Bohs, L. (2020). An amazing new Capsicum (Solanaceae) species from the andean-amazonian piedmont. PhytoKeys **167**:13–29. https://doi.org/10.3897/phytokeys.167.57751.

Browning, B.L., Zhou, Y., and Browning, S.R. (2018). A one-penny imputed genome from next-generation reference panels. Am J Hum Genet **103**:338–348. https://doi.org/10.1016/j.ajhg.2018.07.015.

Borovsky, Y., and Paran, I. (2011). Characterization of fs10.1, a major QTL controlling fruit elongation in *Capsicum*. Theor. Appl. Genet. **123**:657–665.

Bosland, P.W., and Votava, E.J. (2012). Peppers: Vegetable and Spice Capsicums (CABI). https://books.google.com.hk/books?id=5AWTPZeFL8QC.

Carrizo García, C., Barfuss, M.H.J., Sehr, E.M., Barboza, G.E., Samuel, R., Moscone, E.A., and Ehrendorfer, F. (2016). Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae). Ann. Bot. **118**:35–51. https://doi.org/10.1093/aob/mcw079.

Chaim, A.B., Paran, I., Grube, R.C., Jahn, M., van Wijk, R., and Peleman, J. (2001). QTL mapping of fruit-related traits in pepper (*Capsicum annuum*). Theor. Appl. Genet. **102**:1016–1028. https://doi.org/10.1007/s001220000461.

Colonna, V., D'Agostino, N., Garrison, E., Albrechtsen, A., Meisner, J., Facchiano, A., Cardi, T., and Tripodi, P. (2019). Genomic diversity and novel genome-wide association with fruit morphology in *Capsicum*, from 746k polymorphic sites. Sci. Rep. **9**:10067. https://doi.org/10.1038/s41598-019-46136-5.

Gu, X.-z., Cao, Y.-c., Zhang, Z.-h., Zhang, B.-x., Zhao, H., Zhang, X.-m., Wang, H.-p., Li, X.-x., and Wang, L.-h. (2019). Genetic diversity and population structure analysis of Capsicum germplasm accessions. J Integr Agric **18**:1312–1320.

Han, K., Jeong, H.J., Yang, H.B., Kang, S.M., Kwon, J.K., Kim, S., Choi, D., and Kang, B.C. (2016). An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (*Capsicum annuum*). DNA Res. **23**:81–91. https://doi.org/10.1093/dnares/dsv038.

Han, K., Lee, H.Y., Ro, N.Y., et al. (2018). QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in *Capsicum*. Plant Biotechnol. J. **16**:1546–1558. https://doi.org/10.1111/pbi.12894.

Hirsch, C.D., and Springer, N.M. (2017). Transposable element influences on gene expression in plants. Biochim. Biophys. Acta. Gene Regul. Mech. **1860**:157–165. https://doi.org/10.1016/j.bbagrm.2016.05.010.

Hollister, J.D., and Gaut, B.S. (2009). Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res. **19**:1419–1428. https://doi.org/10.1101/gr.091678.109.

Hu, F., Cheng, J., Dong, J., Zhong, J., Zhou, Z., and Hu, K. (2021). Fine mapping and candidate gene analysis of the up locus determining fruit orientation in pepper (Capsicum spp.). Theor. Appl. Genet. **134**:2901–2911. https://doi.org/10.1007/s00122-021-03867-2.

Jordt, S.-E., and Julius, D. (2002). Molecular basis for species-specific sensitivity to "hot" chili peppers. Cell **108**:421–430. https://doi.org/10.1016/S0092-8674(02)00637-2.

Kim, S., Park, M., Yeom, S.-I., Kim, Y.-M., Lee, J.M., Lee, H.-A., Seo, E., Choi, J., Cheong, K., Kim, K.-T., et al. (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. Nat. Genet. **46**:270–278. https://doi.org/10.1038/ng.2877.

Lee, H.Y., Ro, N.Y., Patil, A., Lee, J.H., Kwon, J.K., and Kang, B.C. (2020). Uncovering candidate genes controlling major fruit-related traits in pepper *via* genotype-by-sequencing based QTL mapping and genome-wide association study. Front. Plant Sci. **11**:1100. https://doi.org/10.3389/fpls.2020.01100.

Lefebvre, V., Palloix, A., Caranta, C., Pochard, E., Pochard, and Edmond. (1995). Construction of an intraspecific integrated linkage map of pepper using molecular markers and doubled-haploid progenies. Genome **38**:112–121.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows– Wheeler transform. Bioinformatics **25**:1754–1760.

Liu, L., Tong, H., Xiao, Y., Che, R., Xu, F., Hu, B., Liang, C., Chu, J., Li, J., and Chu, C. (2015). Activation of *Big Grain1* significantly improves grain size by regulating auxin transport in rice. Proc. Natl. Acad. Sci. USA **112**:11102–11107. https://doi.org/10.1073/pnas.1512748112.

Nintemann, S.J., Palmgren, M., and López-Marqués, R.L. (2019). Catch you on the flip side: a critical review of Flippase mutant phenotypes. Trends Plant Sci. **24**:468–478. https://doi.org/10.1016/j.tplants.2019.02.002.

Ou, L., Li, D., Lv, J., Chen, W., Zhang, Z., Li, X., Yang, B., Zhou, S., Yang, S., Li, W., et al. (2018). Pan-genome of cultivated pepper (*Capsicum*) and its use in gene presence–absence variation analyses. New Phytol. **220**:360–363. https://doi.org/10.1111/nph.15413.

Paran, I., and van der Knaap, E. (2007). Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. J. Exp. Bot. **58**:3841–3852.

Perry, L., Dickau, R., Zarrillo, S., Holst, I., Pearsall, D.M., Piperno, D.R., Berman, M.J., Cooke, R.G., Rademaker, K., Ranere, A.J., et al. (2007). Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. Science **315**:986–988.

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

Qin, C., Yu, C., Shen, Y., Fang, X., Chen, L., Min, J., Cheng, J., Zhao, S., Xu, M., Luo, Y., et al. (2014). Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. Proc. Natl. Acad. Sci. USA **111**:5135–5140.

Sabeti, P.C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E.H., McCarroll, S.A., Gaudet, R., et al. (2007). Genome-wide detection and characterization of positive selection in human populations. Nature **449**:913–918. https://doi.org/10.1038/nature06250.

Simmons, C.R., Weers, B.P., Reimann, K.S., Abbitt, S.E., Frank, M.J., Wang, W., Wu, J., Shen, B., and Habben, J.E. (2020). Maize BIG GRAIN1 homolog overexpression increases maize grain yield. Plant Biotechnol. J. **18**:2304–2315.

Solomon, A.M., Kim, T.G., Han, K., Lee, H.Y., Patil, A., Siddique, M.I., Ahn, J., and Kang, B.C. (2021). Fine mapping and candidate gene identification for the CapUp locus controlling fruit orientation in pepper (Capsicum spp.). Front. Plant Sci. **12**:675474. https://doi.org/10.3389/fpls.2021.675474.

Srinivasan, K. (2016). Biological activities of red pepper (*Capsicum annuum*) and its pungent principle Capsaicin: a review. Crit. Rev. Food Sci. Nutr. **56**:1488–1500. https://doi.org/10.1080/10408398.2013.772090.

Stewart, C., Kang, B.C., Liu, K., Mazourek, M., Moore, S.L., Yoo, E.Y., Kim, B.D., Paran, I., and Jahn, M.M. (2005). The *Pun1* gene for pungency in pepper encodes a putative acyltransferase. Plant J. **42**:675–688.

Szpiech, Z.A., and Hernandez, R.D. (2014). selscan: an efficient multithreaded program to perform EHH-based scans for positive selection. Mol Biol Evol **31**:2824–2827. https://doi.org/10.1093/molbev/msu211.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**:585–595.

Tewksbury, J.J., and Nabhan, G.P. (2001). Directed deterrence by capsaicin in chillies. Nature **412**:403–404. https://doi.org/10.1038/35086653.

Tripodi, P., Rabanus-Wallace, M.T., Barchi, L., Kale, S., Esposito, S., Acquadro, A., Schafleitner, R., van Zonneveld, M., Prohens, J., Diez, M.J., et al. (2021). Global range expansion history of pepper (Capsicum spp.) revealed by over 10, 000 genebank accessions. Proc. Natl. Acad. Sci. USA **118**, e2104315118. https://doi.org/10.1073/pnas.2104315118.

van der Knaap, E., and Østergaard, L. (2018). Shaping a fruit: developmental pathways that impact growth patterns. Semin. Cell Dev. Biol. **79**:27–36. https://doi.org/10.1016/j.semcdb.2017.10.028.

Wang, H.P., Li, X.X., and Song, J.P. (2018). Vegetable Genetic Resources in China. Hortic Plant J **4**:83–88.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res **38**:e164. e164.

Wu, S., Zhang, B., Keyhaninejad, N., Rodríguez, G.R., Kim, H.J., Chakrabarti, M., Illa-Berenguer, E., Taitano, N.K., Gonzalo, M.J., Díaz, A., et al. (2018). A common genetic mechanism underlies morphological diversity in fruits and other plant organs. Nat. Commun. **9**:4734. https://doi.org/10.1038/s41467-018-07216-8.

Xiao, H., Jiang, N., Schaffner, E., Stockinger, E.J., and van der Knaap, E. (2008). A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science **319**:1527–1530. https://doi.org/10.1126/science.1153040.

Xu, D., Lei, M., and Wu, R. (1995). Expression of the rice *Osgrp1* promoter-*Gus* reporter gene is specifically associated with cell elongation/expansion and differentiation. Plant Mol. Biol. **28**:455–471. https://doi.org/10.1007/BF00020394.

Xu, X., Liu, X., Ge, S., Jensen, J.D., Hu, F., Li, X., Dong, Y., Gutenkunst, R.N., Fang, L., Huang, L., et al. (2012). Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. Nat. Biotechnol. **30**:105–111. https://doi.org/10.1038/nbt.2050.

Zou, X., Ma, Y., Dai, X., Li, X., and Yang, S. (2020). Spread and industry development of pepper in China. Acta Hortic. Sin. **47**:1715–1716. https://doi.org/10.16420/j.issn.0513-353x.2020-0103 %.

Yarnes, S.C., Ashrafi, H., Reyes-Chin-Wo, S., Hill, T.A., Stoffel, K.M., Van Deynze, A., and Gulick, P. (2012). Identification of QTLs for capsaicinoids, fruit quality, and plant architecture-related traits in an interspecific *Capsicum* RIL population. Genome **56**:61–74.

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. Nat Genet **44**:821–824. https://doi.org/10.1038/ng.2310.

Zhu, Z., Sun, B., Cai, W., Zhou, X., Mao, Y., Chen, C., Wei, J., Cao, B., Chen, C., Chen, G., et al. (2019). Natural variations in the MYB transcription factor *MYB31* determine the evolution of extremely pungent peppers. New Phytol. **223**:922–938. https://doi.org/10.1111/nph.15853.

Zygier, S., Chaim, A.B., Efrati, A., Kaluzky, G., Borovsky, Y., and Paran, I. (2005). QTLs mapping for fruit size and shape in chromosomes 2 and 4 in pepper and a comparison of the pepper QTL map with that of tomato. Theor. Appl. Genet. **111**:437–445. https://doi.org/10.1007/s00122-005-2015-7.