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

Giovanni Giuliano (giovanni.giuliano@enea.it)
National Agency For New Technologies, Energy and Sustainable Economic Development
https://orcid.org/0000-0002-2486-0510

Yacong Cao
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Kang Zhang
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Hailong Yu
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Donghui Xu
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Shumin Chen
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Hong Zhao
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Zhenghai Zhang
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Yinqing Yang
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Xiaozhen Gu
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Xinyan Liu
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Haiping Wang
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Yaxin Jing
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences https://orcid.org/0000-0002-1657-6931

Yajie Mei
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences

Véronique Lefebvre

INRAE https://orcid.org/0000-0001-9916-8433
Weili Zhang
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences
Yuan Jin
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences
Dongliang An
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences
Risheng Wang
Institute of vegetables, Academy of Agricultural Sciences of Guangxi
Paul Bosland
Dept. of Plant & Env. Sciences, NMSU
Xixiang Li
Institute of Vegetables and Flowers of Chinese Academy of Agricultural Sciences
Ilan Paran
Institute of Plant Sciences, Agricultural Research Organization, The Volcani Center
Baoxi Zhang
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences
Lihao Wang
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences
Feng Cheng
Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences https://orcid.org/0000-0003-2982-9675

Article

Keywords: C. annuum, Capsicum spp., pepper variome

Posted Date: October 28th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-612022/v1

Abstract

Pepper (*Capsicum* spp.) is one of the earliest domesticated crops, providing a unique pungent sensation when eaten. Through the construction of the first pepper variome, we describe the main groups that emerged during domestication and breeding of *C. annuum*, their relations and temporal succession, and the molecular events underlying the main transitions. The initial differentiation in fruit shape and pungency, increase in fruit weight, and transition from erect to pendent fruits, and the recent appearance of blocky, large, sweet fruits (bell peppers), were accompanied by strong selection/fixation of key alleles and introgressions in two large genomic regions. Furthermore, we describe the identification of *Up*, a key domestication gene controlling erect vs pendent fruit orientation, encoding a BIG GRAIN protein involved in auxin transport, and *Flip1* associated with capsaicinoid content, encoding a protein involved in phospholipid flipping. The function of *Up* was confirmed by virus-induced gene silencing. These findings constitute a cornerstone for understanding the domestication and differentiation of a key horticultural crop.

Introduction

With $15.67 billion of production value (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 Mediterranean area, Middle and far East and the Americas. Its pungency is conferred by capsaicinoids, primarily capsaicin and dihydrocapsaicin[1], and is sensed by a vanilloid receptor also involved in heat and pain perception *(2)*. The fruits of wild peppers are extremely pungent, small, nearly round, brightly colored, and erect, discouraging mammalian herbivores, which are sensitive to pungency, and favoring seed dispersal by birds, which have impaired capacity to sense pungency and good color vision[2, 3].

About 35 species have been described in the *Capsicum* genus, including the five domesticated species: *C. annuum* L., *C. frutescens* L., *C. chinense* Jacq., *C. baccatum* L., and *C. pubescens* Ruiz & Pavon[4]. Among the domesticated species, *C. annuum* is the most widely cultivated one. Archeological microfossil evidence[5] indicates that cultivated pepper species have undergone distinct domestication events as early as 6,000 years ago in primary diversity centers in South and Meso-America[4, 5, 6, 7, 8]. Pepper was introduced from the West Indies into Europe in the late 15th and early 16th centuries, and then it was rapidly distributed to Africa and Asia, including China, where the earliest written record of pepper dates back to 1591 (Ming Period)[7, 8, 9]. During domestication and breeding, non-deciduous peppers with diverse fruit shapes, sizes, weights, pendent fruit orientation, and a range of pungency levels emerged[10]. The change in fruit position from erect to pendent was selected during early domestication and provides an adaptation to increased fruit size, better protection from sun exposure and from predation by birds. It is thus a key agronomic trait in different fruit-bearing crops[10]. A more recent selection was the emergence of very large, blocky, non-pungent fruits (sweet bell peppers), whose earliest record dates to the 1700's[8].

Studies exploiting the natural variability of pepper allowed the identification of several QTLs and candidate genes controlling capsaicinoid levels such as *Pun1*, *pAMT*, *CaKR1*, and *Pun3*[11, 12, 13, 14] or fruit

shape/size (*longifolia 1-like*)[15, 16, 17]. In contrast, the molecular basis of other key fruit traits, such as erect vs. pendent orientation or narrow vs. blocky types, is hitherto undescribed in pepper or in any other plant species.

The large variations in fruit size, shape, weight, orientation, and pungency found in the pepper germplasm offer an opportunity to explore the genomic events underlying 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[18, 19, 20], the understanding of the molecular evolution of this crop is lagging behind 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 allowed the reconstruction of the history of pepper domestication and breeding, and of the major genomic events and key genes that shaped the present-day diversity of this important horticultural species.

Results And Discussion

The main trajectories of C. annuum domestication

Three hundred forty-seven accessions from 12 species of *Capsicum*, collected from genebanks in Asia, the Americas, Africa, and Europe, of which 311 *C. annuum*, were resequenced to an average depth of ~9×, generating 10.1 trillion paired-end reads (**Table S1)**. 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) (**Table S2**). Variants were uniformly distributed along the 12 chromosomes, with the exception of a genomic region in chromosome 9 containing substantially more variants (**Fig. 1a**), and they were about twice as abundant in intergenic regions than in gene bodies (**Fig. S1**). The median heterozygosity of the accessions was 1.11% (**Fig. S2**), and 56,182 SNPs and 3,080 InDels caused changes in the protein sequences of coding genes (**Table S3**).

We used 33,346 synonymous SNPs located in genes to investigate the phylogenetic relations of the accessions. Different *Capsicum* species formed distinct branches (**Fig. 1b**), while the 311 *annuum* accessions formed nine groups (**Fig. 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 mainly composed of old landraces; III) cultivars with diverse geographical 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; IX) accessions from southwest China, collected from high-altitude areas in Yunnan, Guizhou, Sichuan, and Tibet (**Fig. S3**).

Groups I to IX represent the main domestication and breeding trajectories of pepper worldwide. Both the evolutionary relationships (**Fig. 1c-d**), and the genetic diversity (π) within each group and the genetic differentiation (F_{ST}) between groups (**Table 1**) 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), while 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 either the existence of minor genetic bottlenecks, or of 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, while more derived alleles by selection were present in group III (**Fig. S4**). Groups II and III gave rise, directly or indirectly, to all other groups (**Fig. 1c-d** 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).

Among the Chinese groups, group IX exhibited the highest genetic diversity (π=0.2831) (**Table 1** and a predominant genetic component (represented by dark green in **Fig. 1d**), present in significant levels in the ancestral groups I and II, which were possibly re-introduced 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 (**Fig. 1d**). All groups, with the exception of IV and VI, exhibit 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 (**Fig. 1d**). All groups present relatively high levels of genetic admixture (**Fig. 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).

The domestication and differentiation of narrow fruit peppers

The two wild accessions (*C. annuum* var. *glabriusculum*) have short, very small, waterdrop shaped, erect fruits, with high (839-1146 mg/Kg DW) capsaicinoid content. The early domesticates of group I present, compared to the wild accessions, a significant increase in fruit size, a large variation in fruit shape (olivary, short, conical), the appearance of pendent fruits (8 out of 10), and very large variation in capsaicinoid content (0-1972 mg/Kg DW), indicating a strong diversifying selection exerted on these traits during early domestication (**Table S1**).

During early domestication, average fruit length increased 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 (**Fig. 2a**). The Chinese peppers in groups VII, VIII, and IX showed comparable fruit lengths to those of group III. The increase in length, resulting 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 applied to this day to conserve chili peppers. In contrast, capsaicinoid levels, after the initial diversifying selection in early domesticates showed a multi-phasic trend, with a slight increase in group II and a clear reduction in group III (**Fig. 2b**). The pungency increased again later in groups VII, VIII, and IX, consistent with a 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 (**Table S1**).

Selective pressure generates genomic selection signals, measured as a reduction of nucleotide diversity [ROD][21]. 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 (**Fig. 2c** and **Table S4**). Three previously reported QTLs for fruit shape and length[22, 23] and four capsaicinoid biosynthesis genes (*PDH_E2-P3*, *PDH_E2-D1*, *CM1-D2*, and *a-CT-D1*)[18] are localized in these genomic regions. There are 348 gene units under selection in the transition of group I to group II, including the A-class gene flower homeotic gene *AP2-A* (*Capana04g002188*) (**Tables S5-S6**).

Genomic regions of five chromosomes were found to be under selection in the second transition (group II to group III) (**Fig. 2d** and **Table S4**). 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*, *pAMT-P5*) are localized in these regions, including a second A-class gene *AP2-A* (*Capana02g000700*) (**Tables S5-S6**).

Thus, it appears that early evolutionary transitions in narrow fruit peppers (group I to II, and group II to III) involved selection at 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. In contrast, transition from group II to the Chinese groups VII-IX involved a selection on a narrower group of genes (**Fig. 2e-f** and **Tables S5-S6**), consistent with the hypothesis that a genetic bottleneck was active during this transition, probably due to the transport *via* sea or land (the silk road) to mainland China of a subset of the group II genepool.

The recent emergence of blocky fruit, sweet peppers

Blocky fruit peppers (groups IV and VI) exhibit distinctive phenotypes, such as a large increase in fruit diameter and weight, decreased variation in fruit shape, reduction to almost zero of capsaicinoid levels, and pendent fruit orientation, which is necessary to support the large fruit (**Fig. 2a** and **Table S1**). As aforementioned, they also exhibit a very low genetic diversity (**Table 1**), consistent with their recent emergence[8] and a higher fraction of fixed alleles, either ancestral or derived, compared to the other groups (**Fig. S4**). 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 (**Fig. S5**).

By comparing groups IV and VI with group III, several genomic selection signals were identified using the ROD parameter (**Fig. 3a** and **Table S4**), overlapping with previously described QTLs for fruit shape, length or weight (*fs-8*, *fs10.1B*, *fs11.4*, *fl-8*, *fd-11*, *fw4.1*)[22, 23, 24, 25], and with two capsaicinoid biosynthesis genes (*ACS2-D1* and *pAMT-P5*)[18]. Given the recent emergence of blocky fruit peppers, parameters XP-EHH (cross

population extended haplotype homozygosity)[26], and Tajima's D[27] were used to find additional genomic selection signals (**Fig. 3b**), which were overlapping with QTLs *fd-3.1* for fruit diameter, *SAP* for flower and ovule development, and *qcap6.1* for pungency. *Capana07g001005*, an *Agamous* family gene regulating ovule development, *Capana10g000984* and *Capana10g001014*, encoding cyclin-dependent protein kinase regulators of cell cycle, and *Capana05g000060*, a member of the IQD family that includes *SUN*, regulating fruit shape in tomato[28], were localized in these genomic regions and found to be under strong selection (**Tables S5-S6**).

Two genomic regions, named F9 and F11, on chromosomes 9 and 11 showed very low XP-EHH values (**Fig. 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 (**Fig. S6**), 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 (**Fig. S7**). In F9, all blocky types except three, plus five conical fruit accessions from group V were highly homologous to each other, while most (92.46%) of the other non-blocky types diverged (**Fig. 3c**). In F11, all blocky types except four showed high homologies to each other, as well as four conical fruit peppers from group V, the two wild and five ancestral peppers from group I, while most (91.06%) of the other non-blocky types diverged (**Fig. 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 to the reference fragment than that of F11 as the former has a higher frequency of coding SNPs in comparison to the rest of the pepper genome (**Fig. 1a** and **3e**). We further compared genotypes of loci in F9 with the released sequences of *C. annuum* var. *glabriusculum*[18], and built a phylogenetic tree to trace their evolutionary relationships. These results support the conclusion that F9 was introgressed from this wild *C. annuum* (**Fig. S8**).

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

Fruit shape is an important agronomical trait and is controlled by a conserved network of interacting gene products in distantly related plants[29, 30]. 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 types. We found overlapping, strong association signals for fruit shape index, length, and diameter on chromosome 3. The most significant SNP overlapped with previously mapped QTLs for fruit shape and length (*fs-3.1*, *fl-3.2*), and caused a nonsynonymous Ile340Thr mutation in the *Capana03g002426* (*TRM25*) gene (**Fig. 4**), encoding a TONNEAU 1 Recruiting Motif protein. TRM proteins are part of a protein complex interacting with microtubules arrays and controlling cell division patterns, and are well-known regulators of fruit shape in tomato and cucumber[30]. *TRM25* was expressed in the early stage of pepper fruit development in both the pericarp and placenta tissues (**Fig. 4**). An additional gene, *Capana09g001401*, localized 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 ones (**Fig. S9)**.

Capana09g001401 encodes a glycine-rich cell wall structural protein (*GRP*) that is associated with cell elongation/expansion and differentiation in various tissues in rice[31]. The gene was found to be under selection in blocky fruit types (**Table S5**) and is thus a strong candidate for the control of blocky fruit peppers.

Several genes controlling pungency in pepper have been identified, encoding either structural genes in the capsaicin biosynthesis pathway or, in one case, a transcriptional regulator[11, 12, 13, 14]. GWAS analysis in narrow-fruited peppers showed a strong association signal on chromosome 6, at the *Capana06g001204* gene location. Two nonsynonymous mutations (Ile812Val, Thr495Ile), in strong LD to each other (r^2=0.99) were found in this gene and were significantly associated (P=8.71×10^{-11} and 6.16×10^{-11}) with the increased pungency phenotype (**Fig. 5a-b**). *Capana06g001204* encodes a phospholipid-flipping ATPase (flippase) and is highly expressed in the middle and late development stages of pepper fruit in the pericarp and placenta (**Fig. 5c**). We propose the name *Flip1* for this novel gene controlling capsaicinoid accumulation. Flippases translocate lipids (mainly phospholipids) across biological membranes through the hydrolysis of ATP, and are involved in a series of physiological responses such as membrane stabilization, vesicle-mediated metabolite transport, adaptation to temperature changes, defense, and lipid signaling[32]. The role of the FLIP1 protein in the control of pungency is intriguing: we hypothesize that it could be either directly involved in capsaicinoid transport across membranes, or in membrane protection against the destabilizing effects of high capsaicin concentrations[33].

Compared to narrow fruit peppers, blocky fruit peppers contain almost no capsaicin or dihydrocapsaicin. GWAS analysis found a strong association signal on chromosome 2 (**Fig. S10a-b**), close to the previously reported *Pun1* gene (*Capana02g002340*) mediating the last step in capsaicin biosynthesis[11]. A loss-of-function deletion in the recessive allele *pun1* was found using reads mapping information (**Fig. S10c**); this structural variation has the most significant association (P<2.23×10^{-308}) with the pungency trait.

Fruit orientation is an important agricultural trait in both vegetable crops and fruit trees, but its molecular basis is unknown. As aforementioned, fruit orientation transitioned from erect (up) in wild peppers to pendent (down) in domesticated large-fruited ones. The *up* locus controls fruit orientation in pepper (**Fig. 6a**), but the gene underlying this variation is unknown. We conducted a genome-wide association study (GWAS) for this trait and found a strong association signal on chromosome 12, where the *up* locus resides[34] (**Fig. 6b**). The most significant signal was in the promoter region of gene *Capana12g000954*, expressed in the flower pedicel and the placenta of pepper fruit (**Fig. S11**). *Capana12g000954* encodes a BIG GRAIN 1-like (BGL) protein, whose rice ortholog is expressed in vascular tissues and mediates auxin transport[35]. This gene was one of two genes considered previously as candidates for controlling pepper fruit orientation[17]. A 579-bp deletion was detected in the promoter region of the gene in the pendent accessions, with an extremely significant association with the fruit orientation trait (P=6.00×10^{-175}) and was confirmed in a test population composed of 241 samples (**Fig. 6c** and **Fig. S12**). RNA-Seq and quantitative Real Time (qRT) PCR analyses found that this deletion is associated with a high expression

level of the gene in pedicels of pendent fruits, but accessions with erect fruits exhibited low level of expression of the gene (**Fig. 6d**). *BG1-like* genes have been implicated mostly in controlling organ size and yield in rice, Arabidopsis and maize[35, 36]. Additional growth-related traits such as plant height, tiller angle, and gravitropism, as well as stress tolerance were affected by down or up regulation of these genes. The function of *BG1-like* genes has not been determined yet in fruit crops. The novel putative role of *BG1-like* in controlling fruit orientation in pepper is likely mediated by differential distribution of auxin and level of gravitropism response in the pedicle.

We crossed a wild pepper accession (erect) with a blocky pepper accession (pendent) and obtained a 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 signal found that it overlapped with the GWAS signal, where locates the gene *Capana12g000954* (**Fig. 6e**). We further verified the function of the *BG1-like* gene through virus-induced gene silencing (VIGS) (**Methods**). Plants infected with the TRV2::*up* vector showed erect fruits, compared to the pendent fruits of the wild-type accession and of the accession infected with an empty TRV2 vector (**Figs**. **6f** and **S13**). Expression of the *BG1-like* gene was suppressed in pedicels of erect fruits infected with the TRV::*up* vector, but not in pedicels of pendent fruits not infected, or infected with the empty TRV2 vector (**Fig. 6g**).

The key temporal sequence in pepper fruit domestication and diversification

Analysis of the pepper variome allows a temporal reconstruction of the key events that 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 high proportion in the ancestral group I, increased in groups III to IX and reached complete fixation in blocky groups IV and VI (**Fig. 7a**). Interestingly, the *flip1* mutation controlling 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* value=2.32×10^{-11}) which is not due to physical linkage, since 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 showed a very high association in all groups (*P*=1.12×10^{-21}).

Strong associations between unlinked loci can be explained by a series of different scenarios: i) reduced gene flow of the populations containing the associated regions with respect to the general genepool; this hypothesis is unlikely in the present case, since it would influence the linkage disequilibrium of additional unlinked loci, which does not seem to be the case; ii) simultaneous selection for two different traits, encoded by the associated loci; this seems to be the case for *up* and *flip1* during early domestication; and iii) cooperative action of the associated unlinked loci in determining a single phenotype; this seems to be the case for the F9 and F11 introgressions, which are almost always found together in blocky fruit types.

In contrast, the knock-out *pun1* allele controlling fruit pungency was extremely rare in narrow pepper groups I-IX, and its frequency increased progressively in groups IV and VI (blocky) (**Fig. 7a**). The most likely explanation is that early selection for blocky fruits co-opted accidentally the *pun1* sweet pepper allele in a subset of group IV accessions, and that the associated "sweet" phenotype was subsequently selected for to reach a complete fixation in group VI, which is the most recent blocky fruit group and presents 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 (**Fig. 7b**): i) all alleles found in one or more later groups were pre-existing in the ancestral group; ii) during early domestication (groups I→III), the *up* allele frequency increased to almost complete fixation, mediating the conversion from erect to pendent fruits; iii) the F9 and F11 introgressions were co-opted, leading to the appearance of blocky fruit peppers (groups IV and VI), which also became sweet due to the increase and fixation of *pun1*. In contrast, the genetic circuits controlling fruit elongation and pungency in narrow fruit peppers appear to be more complex: iv) fruit elongation between groups in groups II and IX was primarily mediated by the *trm25* allele, while in other groups the primary contribution appears to be due to the contribution of additional genes (**Fig. 2f**); v) 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 the complete fixation of *flip1* and also probably accompanied by selection at other loci controlling capsaicinoid content (**Fig. 2f**).

In conclusion, the first variome map of pepper described here, uncovered the main genomic events underlying the initial transition from small, almost round, erect, pungent fruits, to larger, more elongated fruits, with a larger variation in capsaicinoid content, followed by the 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 constitute a cornerstone for the further breeding and improvement of this important horticultural crop.

References

1. Srinivasan K. Biological Activities of Red Pepper (*Capsicum annuum*) and Its Pungent Principle Capsaicin: A Review. *Critical Reviews in Food Science and Nutrition* **56**, 1488-1500 (2016).

2. Jordt S-E, Julius D. Molecular Basis for Species-Specific Sensitivity to "Hot" Chili Peppers. *Cell* **108**, 421-430 (2002).

3. Tewksbury JJ, Nabhan GP. Directed deterrence by capsaicin in chillies. *Nature* **412**, 403-404 (2001).

4. Carrizo García C*, et al.* Phylogenetic relationships, diversification and expansion of chili peppers (*Capsicum*, Solanaceae). *Annals of Botany* **118**, 35-51 (2016).

5. Perry L*, et al.* Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. *Science* **315**, 986–988 (2007).

6. Aguilar-Melendez A, Morrell PL, Roose ML, Kim SC. Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum annuum*; Solanaceae) from Mexico. *American Journal of Botany* **96**, 1190-1202 (2009).

7. Andrews J. *Peppers: The Domesticated Capsicums*. University of Texas Press (1995).

8. Bosland PW, Votava EJ. *Peppers: Vegetable and Spice Capsicums*. CABI (2012).

9. Xuexiao Z, Yanqing M, Xiongze D, Xuefeng L, Sha aY. Spread and Industry Development of Pepper in China. *Acta Horticulturae Sinica* **47**, 1715-1716 (2020).

10. Paran I, van der Knaap E. Genetic and molecular regulation of fruit and plant domestication traits in tomato and pepper. *Journal of Experimental Botany* **58**, 3841-3852 (2007).

11. Stewart C, Kang BC, Liu K, Mazourek M, Jahn MM. The *Pun1* gene for pungency in pepper encodes a putative acyltransferase. *Plant Journal* **42**, 675-688 (2005).

12. Arce-Rodríguez ML, Ochoa-Alejo N. Biochemistry and molecular biology of capsaicinoid biosynthesis: recent advances and perspectives. *Plant Cell Reports* **38**, 1017-1030 (2019).

13. Han K*, et al.* QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in *Capsicum*. *Plant Biotechnology Journal* **16**, 1546-1558 (2018).

14. Zhu Z*, et al.* Natural variations in the MYB transcription factor *MYB31* determine the evolution of extremely pungent peppers. *New Phytologist* **223**, 922-938 (2019).

15. Chaim AB, Paran I, Grube RC, Jahn M, van Wijk R, Peleman J. QTL mapping of fruit-related traits in pepper (*Capsicum annuum*). *Theoretical and Applied Genetics* **102**, 1016-1028 (2001).

16. Colonna V*, et al.* Genomic diversity and novel genome-wide association with fruit morphology in *Capsicum*, from 746k polymorphic sites. *Scientific Reports* **9**, 10067 (2019).

17. Lee H-Y, Ro N-Y, Patil A, Lee J-H, Kwon J-K, Kang B-C. Uncovering Candidate Genes Controlling Major Fruit-Related Traits in Pepper *via* Genotype-by-Sequencing Based QTL Mapping and Genome-Wide Association Study. *Frontiers in Plant Science* **11**, 1100 (2020).

18. Qin C*, et al.* Whole-genome sequencing of cultivated and wild peppers provides insights into *Capsicum* domestication and specialization. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 5135-5140 (2014).

19. Kim S*, et al.* Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nature Genetics* **46**, 270-278 (2014).

20. Ou L*, et al.* Pan-genome of cultivated pepper (*Capsicum*) and its use in gene presence–absence variation analyses. *New Phytologist* **220**, 360-363 (2018).

21. Xu X*, et al.* Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes. *Nature Biotechnology* **30**, 105-111 (2012).

22. Han K*, et al.* An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (*Capsicum annuum*). *DNA Research* **23**, 81-91 (2016).

23. Yarnes SC*, et al.* Identification of QTLs for capsaicinoids, fruit quality, and plant architecture-related traits in an interspecific *Capsicum* RIL population. *Genome* **56**, 61-74 (2012).

24. Borovsky Y, Paran I. Characterization of fs10.1, a major QTL controlling fruit elongation in *Capsicum*. *Theoretical and Applied Genetics* **123**, 657-665 (2011).

25. Zygier S, Chaim AB, Efrati A, Kaluzky G, Borovsky Y, Paran I. 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. *Theoretical and Applied Genetics* **111**, 437-445 (2005).

26. Sabeti PC*, et al.* Genome-wide detection and characterization of positive selection in human populations. *Nature* **449**, 913-918 (2007).

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

28. Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science* **319**, 1527-1530 (2008).

29. van der Knaap E, Østergaard L. Shaping a fruit: Developmental pathways that impact growth patterns. *Seminars in Cell & Developmental Biology* **79**, 27-36 (2018).

30. Wu S*, et al.* A common genetic mechanism underlies morphological diversity in fruits and other plant organs. *Nature Communications* **9**, 4734 (2018).

31. Xu D, Lei M, Wu R. Expression of the rice *Osgrp1* promoter-*Gus* reporter gene is specifically associated with cell elongation/expansion and differentiation. *Plant Molecular Biology* **28**, 455-471 (1995).

32. Nintemann SJ, Palmgren M, Lopez-Marques RL. Catch You on the Flip Side: A Critical Review of Flippase Mutant Phenotypes. *Trends in Plant Science* **24**, 468-478 (2019).

33. Aranda FJ, Villalaín J, Gómez-Fernández JC. Capsaicin affects the structure and phase organization of phospholipid membranes. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1234**, 225-234 (1995).

34. Lefebvre*, et al.* Construction of an intraspecific integrated linkage map of pepper using molecular markers and doubled-haploid progenies. *Genome* **38**, 112-121 (1995).

35. Liu L*, et al.* Activation of *Big Grain1* significantly improves grain size by regulating auxin transport in rice. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 11102-11107 (2015).

36. Simmons CR*, et al.* Maize BIG GRAIN1 homolog overexpression increases maize grain yield. *Plant Biotechnology Journal* **18**, 2304-2315 (2020).

Declarations

Acknowledgments: We thank Prof. Xiaowu Wang for his helpful suggestions.

Funding:

China National Key Technology Research and Development Program (2016YFD0100200, 2016YFD0101700, 2018YFD1000800, and 2020YFD1001100)

National Natural Science Foundation of China (NSFC grants 31722048, 31972411, and 3170110879)

Science and Technology Innovation Program of the Chinese Academy of Agricultural Sciences (CAAS-ASTIP-IVFCAAS)

Earmarked fund for China Agriculture Research System (CARS-25)

CAAS-GAAS Coordinated Innovation Project of the Chinese Academy of Agricultural Sciences (2019)

Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, P.R. China

EU Horizon 2020 G2P-SOL project (grant n. 677379) to GG, IP and VL.

Author contributions:

Conceptualization: LW, FC

Formal analysis: FC, KZ, YC

Investigation: HY, DX, XL (X. Liu), YY, YJ (Y. Jing), YM, YC, HZ, ZZ, SC

Resources: YC, XG, HW, BZ, XL (X. Li)

Validation: HY, WZ, YJ (Y. Jin), DA

Writing – original draft: FC, LW, YC, KZ, HY

Writing – review & editing: GG, FC, LW, YC, KZ, RW, PWB, IP, VL

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: The raw resequencing data of this study has been deposited in the Genome Sequence Archive in 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.

Tables

Table 1. Genetic differentiation (estimated by FST) between each two of the nine pepper groups, and genetic diversity (estimated by π) within each pepper 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

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

Figures

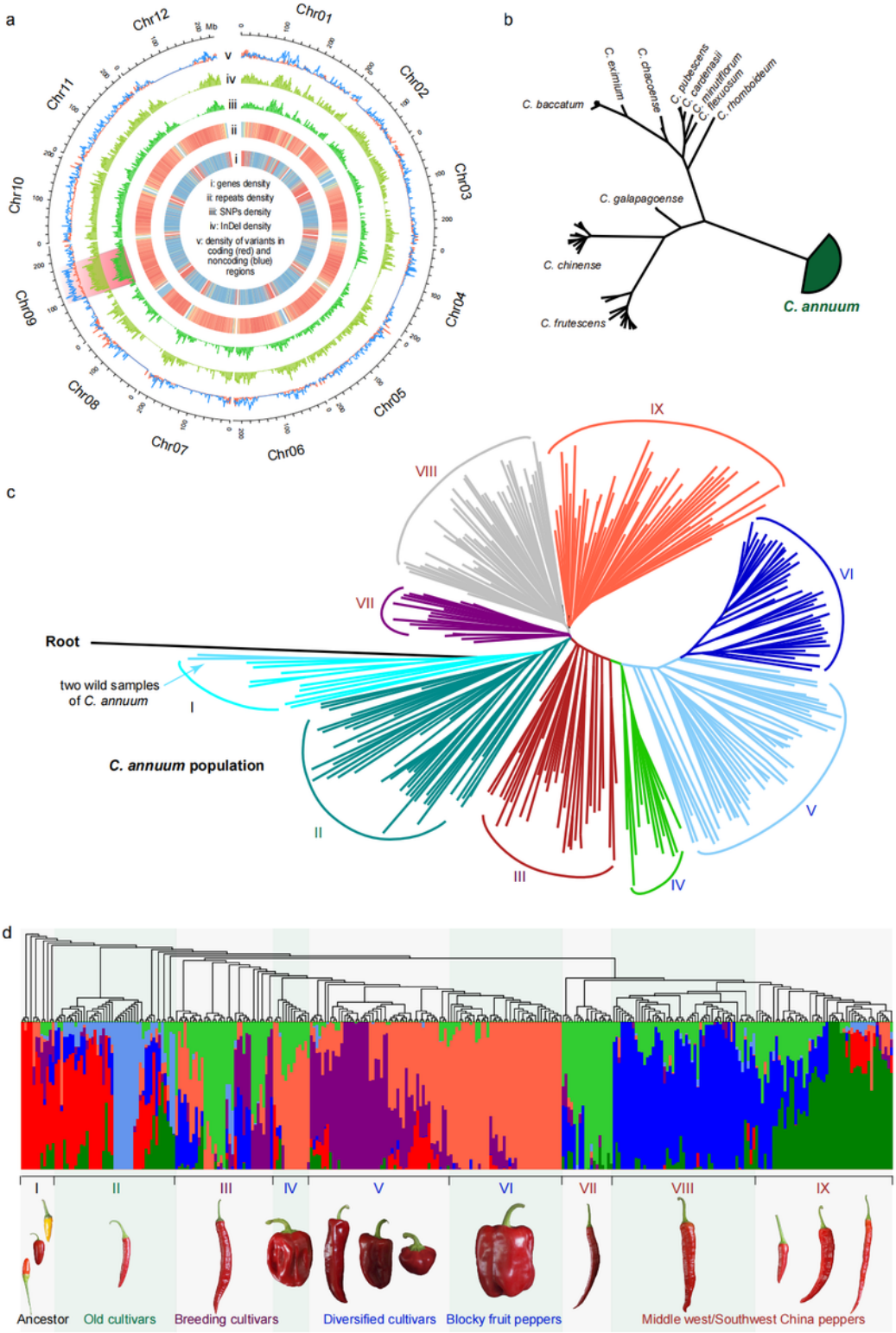

Figure 1

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.

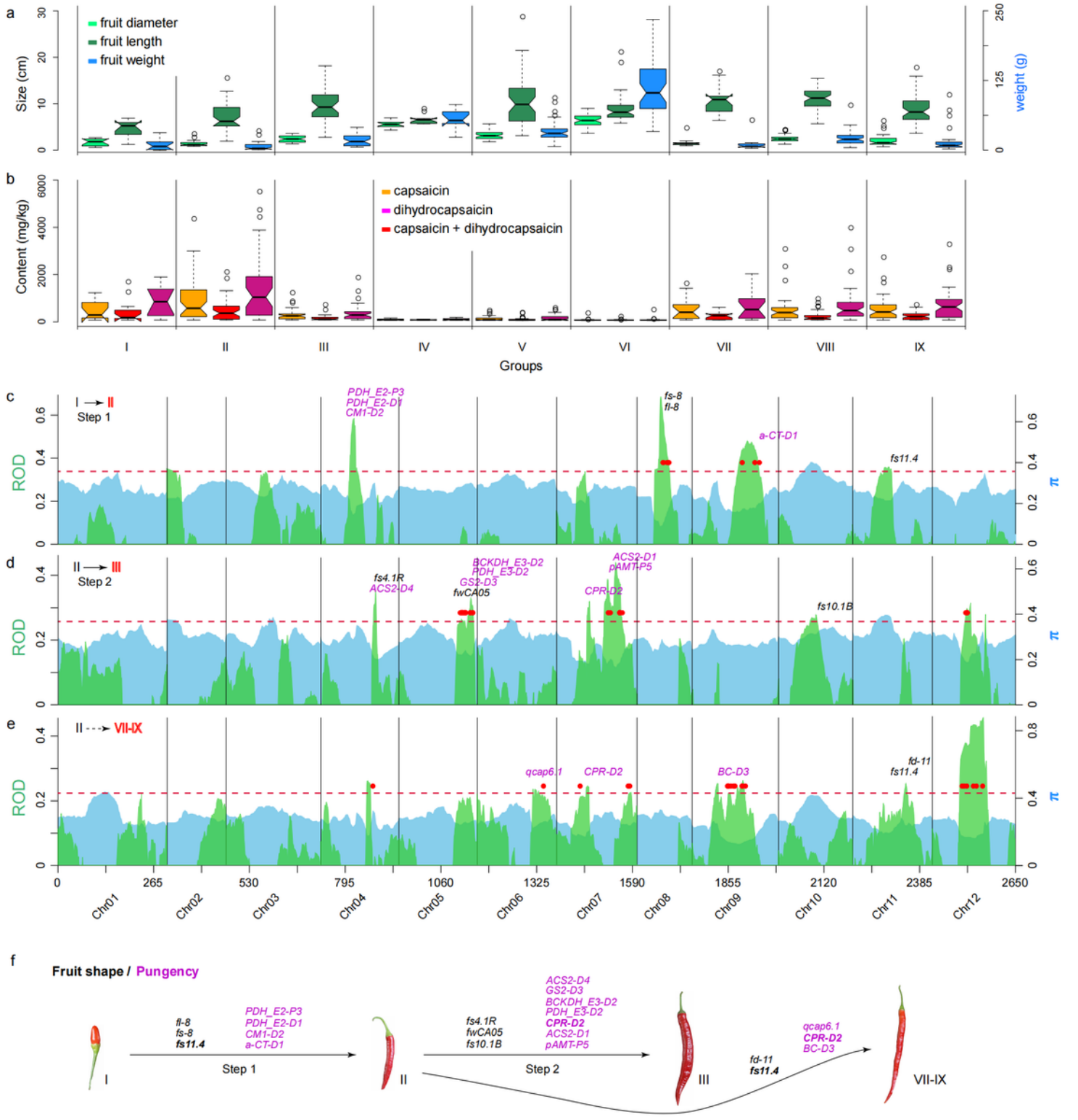

Figure 2

Selection for fruit shape and pungency in narrow fruit peppers. Box plot of fruit diameter, length and weight (a) and of capsaicinoid content (b), in the nine C. annuum groups. Genomic selection signals detected by π (colored in blue) and ROD (colored in green) in 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.

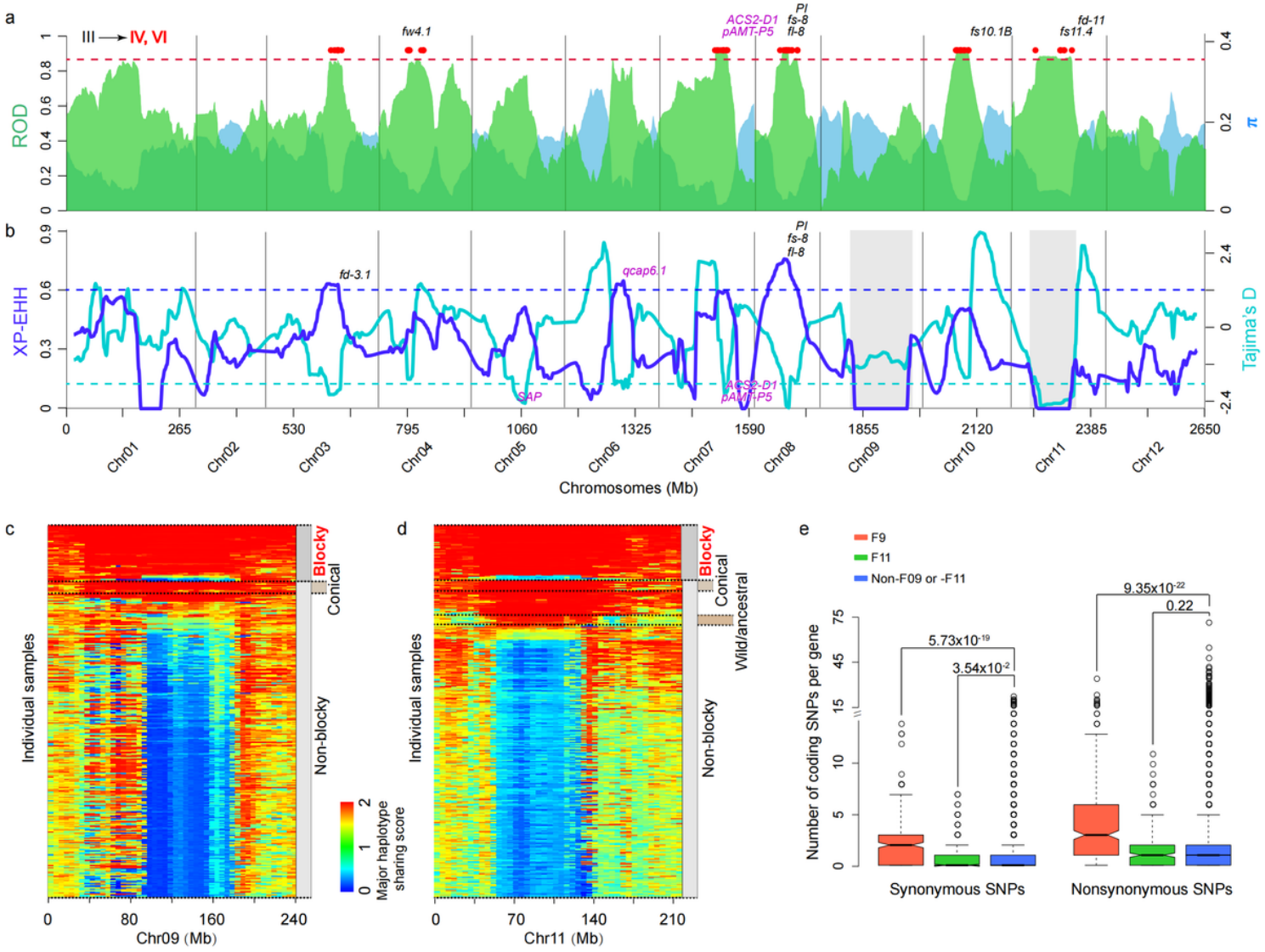

Figure 3

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 to group III. Dashed lines indicate genomic regions under selection, identified as top 5% outliers of ROD and XP-EHH or bottom 5% outliers of Tajima's D. Genetic loci on traits of fruit shape (colored in black) and pungency (purple) overlapping with genomic regions under selection are labeled at the corresponding positions. Major haplotype sharing scores on chromosomes 9 (c) or 11 (d) of blocky and non-blocky fruit peppers, using blocky haplotypes as a reference. (e) Differences in the numbers of coding SNPs in genes localized at the F9 and F11 introgressed regions compared to all other genomic regions.

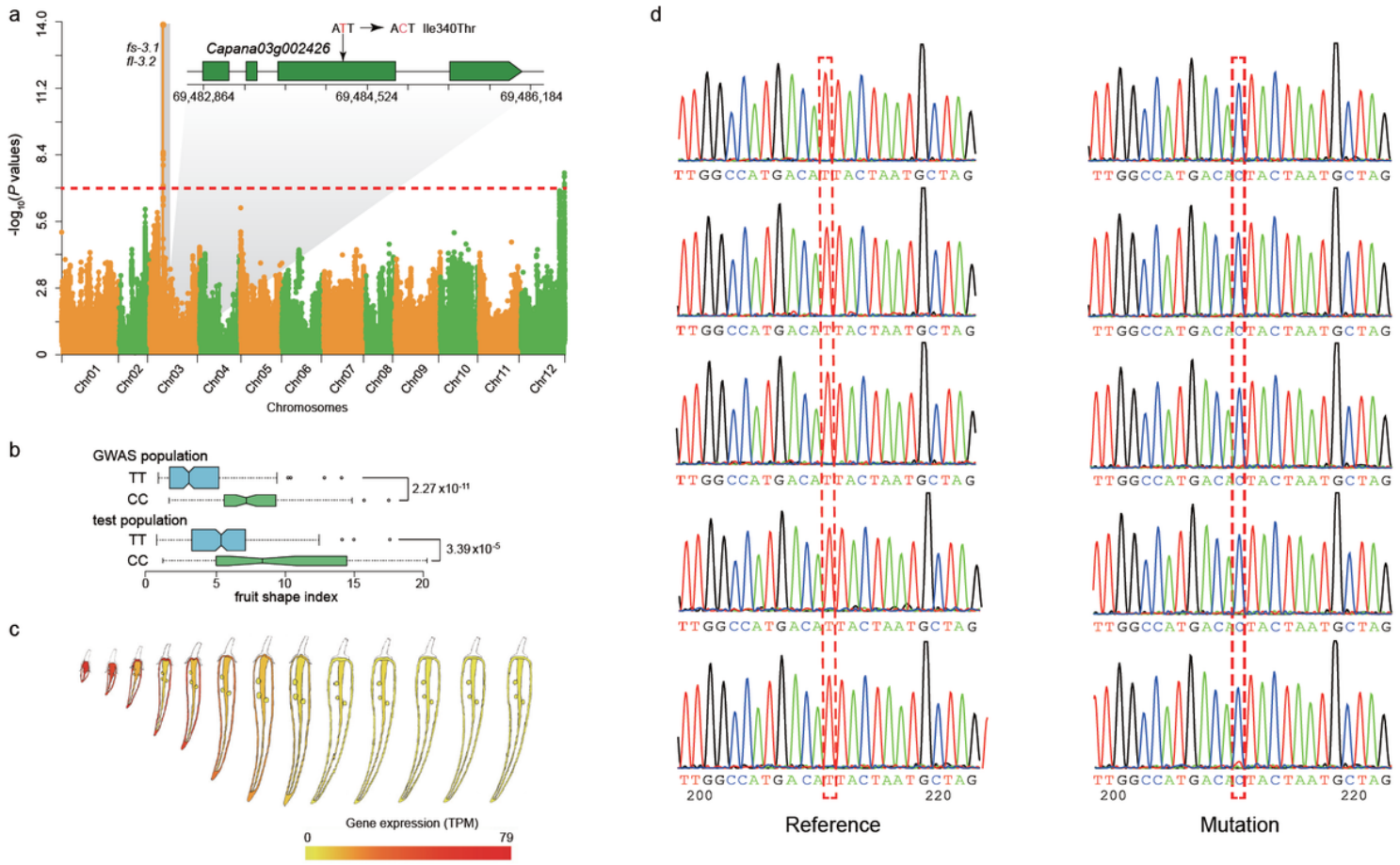

Figure 4

A missense mutation in the TRM25 gene is associated with elongated fruit shape. (a) GWAS association signal of fruit shape index with an Ile340Thr mutation in Capana03g002426 (TRM25); 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) TRM25 expression during development of the pepper fruit. (d) Sanger sequencing of the nonsynonymous mutation in TRM25 in the test population. The red dashed-line rectangle denotes the position of the mutation.

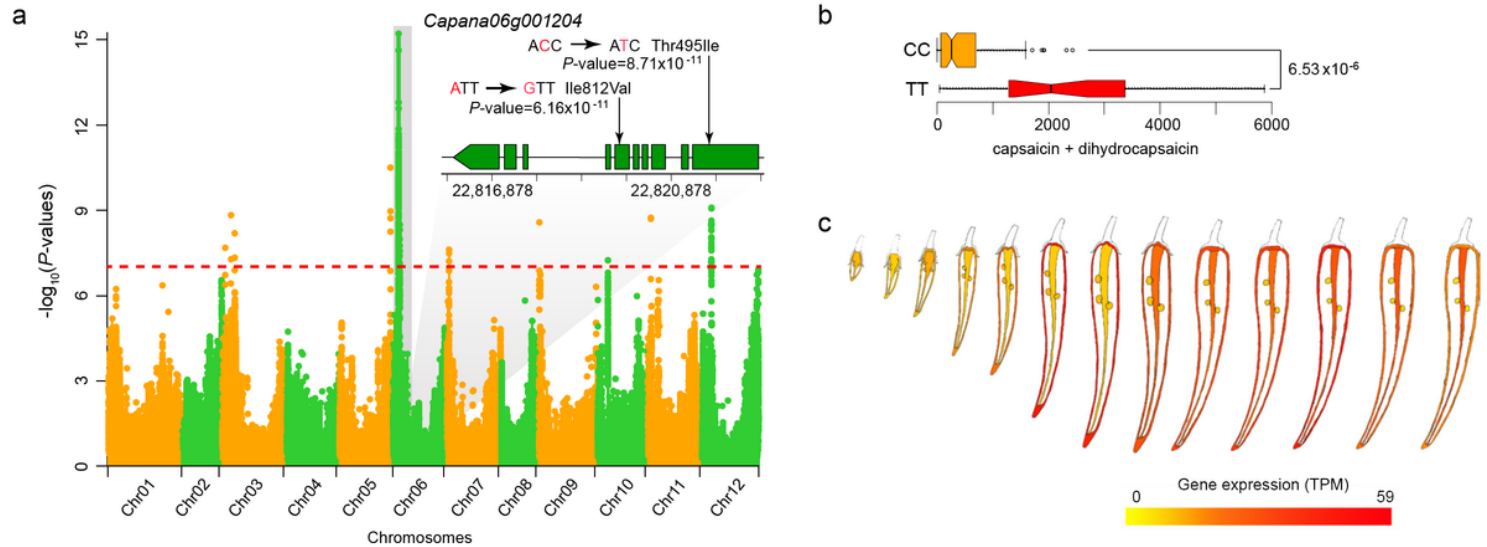

Figure 5

Two missense mutations in the Flip1 gene are associated with decreased capsaicinoid content. (a) GWAS association signal of capsaicinoid content with two (Thr495Ile and Ile812Val) mutations in Capana06g001204 (Flip1); (b) differences in capsaicinoid content in genotypes carrying the nonsynonymous mutations, in GWAS and test populations. (c) Flip1 expression during development of the pepper fruit.

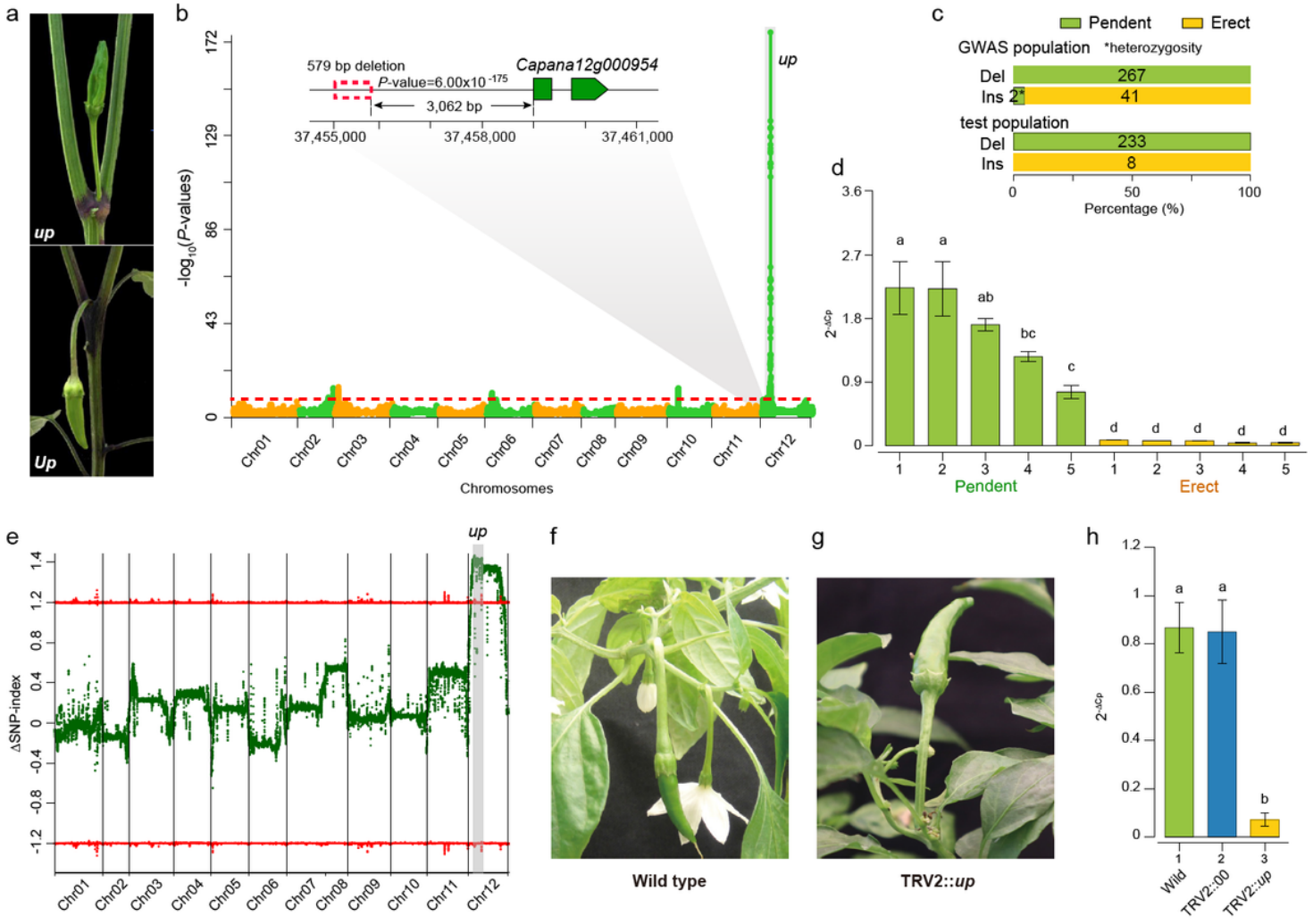

Figure 6

Identification of the Up gene controlling pepper fruit orientation. (a) C. annuum plants segregating for the up mutation, controlling 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 vs erect). Vertical bars: standard error. (e) BSA-Seq signal based on the F2 population of pendent and erect accessions. (f) the wild pendent fruit of the pepper accession “Changyang chili”; (g) erect fruit of “Changyang chili” after VIGS treatment with TRV2::up; (h) qRT-PCR analysis of up expression in fruit pedicels of controls (wild) and VIGS samples showing pendent and erect fruits, respectively. Vertical bars:

standard error. TRV2::00 and TRV2::up denote plants of “Changyang chili” treated with VIGS of empty vector and up gene (Capana12g000954).

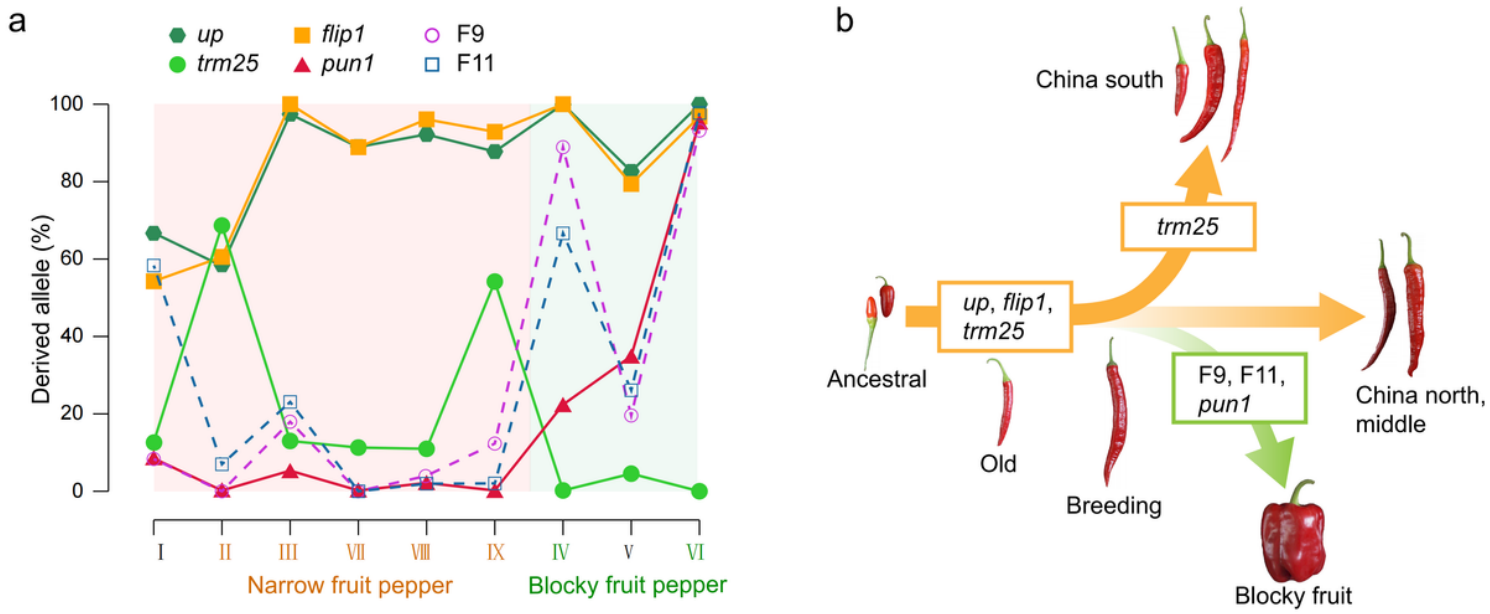

Figure 7

Key events that shaped pepper fruit domestication and diversification. (a) Frequency distributions of the mutations in the up and trm25 alleles, controlling respectively fruit orientation and shape, of the flip1 and pun1 alleles, controlling fruit pungency, and of the F9 and F11 introgressions, associated with blocky fruit shape, in the nine C. annuum groups. (b) Chronodiagram of the key genetic events controlling fruit characteristics during pepper domestication and differentiation.

Supplementary Files

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

- SupplementaryTables.xlsx
- MethodsSupplementaryfigs.docx