















# Natural variation in CRABS CLAW contributes to fruit length divergence in cucumber

Gen Che <sup>1,2,†</sup> Yupeng Pan <sup>3,†</sup> Xiaofeng Liu <sup>1,†</sup> Min Li <sup>1,†</sup> Jianyu Zhao <sup>1</sup> Shuangshuang Yan <sup>1</sup> Yuting He <sup>1</sup> Zhongyi Wang <sup>1</sup> Zhihua Cheng <sup>1</sup> Weiyuan Song <sup>1</sup> Zhaoyang Zhou <sup>1</sup> Tao Wu <sup>4</sup> Yiqun Weng <sup>3,5,\*</sup> and Xiaolan Zhang <sup>1,\*</sup>

- 1 State Key Laboratories of Agrobiotechnology, Beijing Key Laboratory of Growth and Developmental Regulation for Protected Vegetable Crops, MOE Joint Laboratory for International Cooperation in Crop Molecular Breeding, China Agricultural University, Beijing 100193, China
- 2 School of Life Science, Key Laboratory of Herbage & Endemic Crop Biology, Ministry of Education, Inner Mongolia University, Hohhot 010070, China
- 3 Horticulture Department, University of Wisconsin-Madison, 1575 Linden Drive, Madison, Wisconsin 53706, USA
- 4 College of Horticulture, Hunan Agricultural University, Changsha, China
- 5 USDA-ARS, Vegetable Crops Research Unit, 1575 Linden Drive, Madison, Wisconsin 53706, USA

\*Author for correspondence: zhxiaolan@cau.edu.cn (X.Z.), yiqun.weng@usda.gov (Y.W.)

<sup>†</sup>These authors contributed equally.

X.Z., and Y.W. designed the research; G.C., Y.P., X.L., and M.L. performed the experiments; G.C., Y.P., X.L., X.Z., and Y.W. wrote the paper; J.Z., S.Y., Z.W., Z.C., W.S., Y.H., T.W., and Z.Z. provided experimental assistance and helped with manuscript revisions.

The authors responsible for the distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plcell>) are Xiaolan Zhang (zhxiaolan@cau.edu.cn) and Yiqun Weng (yiqun.weng@usda.gov).

## Abstract

Fruit length is a key domestication trait that affects crop yield and appearance. Cucumber (*Cucumis sativus*) fruits vary from 5 to 60 cm in length. Despite the identification of several regulators and multiple quantitative trait loci (QTLs) underlying fruit length, the natural variation, and molecular mechanisms underlying differences in fruit length are poorly understood. Through map-based cloning, we identified a nonsynonymous polymorphism (G to A) in *CRABS CLAW* (*CsCRC*) as underlying the major-effect fruit size/shape QTL *FS5.2* in cucumber. The short-fruit allele *CsCRC<sup>A</sup>* is a rare allele that has only been found in round-fruited semi-wild Xishuangbanna cucumbers. A near-isogenic line (NIL) homozygous for *CsCRC<sup>A</sup>* exhibited a 34~39% reduction in fruit length. Introducing *CsCRC<sup>G</sup>* into this NIL rescued the short-fruit phenotype, and knockdown of *CsCRC<sup>G</sup>* resulted in shorter fruit and smaller cells. In natural cucumber populations, *CsCRC<sup>G</sup>* expression was positively correlated with fruit length. Further, *CsCRC<sup>G</sup>*, but not *CsCRC<sup>A</sup>*, targets the downstream auxin-responsive protein gene *CsARP1* to regulate its expression. Knockout of *CsARP1* produced shorter fruit with smaller cells. Hence, our work suggests that *CsCRC<sup>G</sup>* positively regulates fruit elongation through transcriptional activation of *CsARP1* and thus enhances cell expansion. Using different *CsCRC* alleles provides a strategy to manipulate fruit length in cucumber breeding.

## IN A NUTSHELL

**Background:** Cucumber is an important vegetable crop cultivated worldwide. The fruit length in cucumber varies from 5 to 60 cm in different germplasms, and thus serves as a key domestication trait affecting yield and appearance quality. Although several regulators and multiple QTLs controlling fruit length have been identified, the natural variations and molecular mechanisms underlying fruit length divergence are largely unknown.

**Question:** How can cucumber vary so widely in its fruit length? *FS5.2* was previously identified as a major-effect QTL specifying fruit length in cucumber, while the underlying controlling gene and its regulating mechanism are unclear.

**Findings:** A nonsynonymous SNP (G to A) in *CsCRC* was identified to control the fruit length variation caused by the *FS5.2* locus. Allelic analysis of 165 cucumber accessions showed that the *CsCRC<sup>A</sup>* allele only exists in 7 Xishuangbanna cucumbers with short or round fruits, while the *CsCRC<sup>G</sup>* allele was shared by all the remaining wild and cultivated cucumbers. Replacing the *CsCRC<sup>G</sup>* allele with *CsCRC<sup>A</sup>* by backcrossing resulted in a 34~39% reduction in fruit length. Genetic transformation-based functional characterization revealed that *CsCRC<sup>G</sup>* is a positive regulator of fruit length elongation and cell size expansion. *CsCRC<sup>G</sup>*, but not *CsCRC<sup>A</sup>*, targets the downstream auxin-responsive protein gene *CsARP1* (a positive mediator of fruit length), to promote its expression and stimulate fruit elongation in cucumber.

**Next step:** Manipulation of fruit length can be achieved by either modulating the expression levels of *CsCRC<sup>G</sup>* or utilizing different *CsCRC* alleles (*CsCRC<sup>G</sup>* or *CsCRC<sup>A</sup>*) in cucumber breeding practices. The regulatory mechanism of *CsARP1* and its relation with auxin await further characterization in plants.

## Introduction

Fruits are the seed-bearing structures in vascular plants and account for a substantial portion of human food resources. Fruit size, which is usually measured by fruit length (L), fruit diameter (D), and the L/D ratio (also called fruit shape index, FSI) (Weng et al., 2015), is a crucial domestication trait under intense selection for increasing human consumption. Cucumber (*Cucumis sativus* L.) is an important vegetable crop cultivated worldwide since its domestication around 3,000 years ago (Staub et al., 2008; Weng, 2021). The cucumber fruit belongs to the pepo fruit type that can be consumed either fresh or processed (pickles); and fruit length is a predominant trait that directly affects crop yield and product appearance, a key aspect of quality (Zhao et al., 2019). Wild cucumber (*C. sativus* var. *hardwickii*) bears fruits of only 3–5 cm in length, while cultivated cucumbers exhibit dramatic variation in fruit length, which was probably caused by domestication and improvement under diverse local adaptation pressures, production purposes, and consumer preferences. For example, European pickling cucumbers have a length of 6–15 cm; US slicing cucumbers are 20–22 cm in length; north China fresh-market cucumbers are usually 30–40 cm in length (Weng, 2021).

A longstanding question is how cucumber fruits vary so widely in length. More than 20 consensus quantitative trait loci (QTLs) have been identified for fruit size/shape in cucumber (Weng et al., 2015; Che and Zhang, 2019; Pan et al., 2020; Sheng et al., 2020). Very few of these QTLs have been cloned. Among them, the homologs of the tomato (*Solanum lycopersicum*) fruit shape genes *SUN* (*CsSUN* in cucumber) and *TRMS* (*TONNEAU1 Recruiting Motif*) (*CsTRMS*) have been shown to be the candidate genes for the QTLs *FS1.2*, and *FS2.1*, respectively (Pan et al., 2017a; Wu et al., 2018). In addition, utilizing natural populations and ethyl methanesulfonate (EMS) mutants, fruit length regulators

(*SHORT-FRUIT 1* [*SF1*], *SF2*, *SF3*, and *FRUITFUL1<sup>A</sup>* [*CsFUL1<sup>A</sup>*]) have also been identified in cucumber (Xin et al., 2019; Zhao et al., 2019; Zhang et al., 2020; Wang et al., 2021a). *SF1*, which encodes a cucurbit-specific RING-type E3 ligase, modulates cucumber fruit length by fine-tuning endogenous ethylene dosage (Xin et al., 2019). *SF2* encodes a HISTONE DEACETYLASE COMPLEX 1 (HDC1) protein that promotes cell proliferation through promoting histone deacetylation and several biological processes (Zhang et al., 2020). A nonsynonymous mutation in *SF3*, which is a homolog of the Katanin  $\beta$  subunit gene *KTN1*, resulted in short fruit due to reduced cell number (Wang et al., 2021a; Song et al., 2022). *CsFUL1<sup>A</sup>* is a functional allele of the MADS-box transcription factor gene *FRUITFULL* that is predominant in Asian long-fruited cucumbers. The encoded protein *CsFUL1<sup>A</sup>* negatively regulates fruit elongation by suppressing the expression of its downstream targets *SUPERMAN* (*CsSUP*) and the auxin transporter genes *PIN-FORMED1* (*CsPIN1*) and *CsPIN7*, thus inhibiting auxin accumulation and cell division/expansion during fruit development (Zhao et al., 2019). Given the important developmental nature and economic value of cucumber fruit, further investigations are required to better understand the regulatory pathways underlying the dramatic fruit length variation observed in cucumber.

CRC is a member of the YABBY protein family that contains conserved C2C2 zinc-finger and DNA-binding YABBY (High mobility group [HMG]-box-like) domains, and functions in carpel development, nectary initiation, floral meristem determination, and leaf midrib formation depending on the specific species (Eshed et al., 1999; Siegfried et al., 1999; Bowman, 2000; Eckardt, 2010; Finet et al., 2016). In *Arabidopsis* (*Arabidopsis thaliana*), CRC acts as a major determinant of carpel and nectary development.

Loss-of-function alleles of *CRC* result in unfused carpels, lack of nectary, wide and short siliques (Bowman and Smyth, 1999; Lee et al., 2005a). Overexpression of *CRC* in *Arabidopsis* confers some carpel characters to sepals, while the gynoecium transforms into small solid cylindrical structures composed of style and stigma (Eshed et al., 1999). Consistent with its associated phenotypes, *CRC* is specifically expressed in gynoecia and nectary, and such expression depends on the B-, C-, and E-class floral patterning genes (Lee et al., 2005b). The style defects seen in the *crc* mutant can be partially rescued by treatment with an auxin transport inhibitor or by local auxin production (Staldal et al., 2008). *CRC* functions in auxin homeostasis through the transcriptional repression of *TORNADO2* (*TRN2*), which encodes a tetraspanin protein controlling auxin maxima in *Arabidopsis* (Yamaguchi et al., 2017). Similarly, in petunia (*Petunia hybrida*), two closely related *CRC* homologs, PhCRC1 and PhCRC2, function in nectary and carpel development, and their activation relies on C-class genes (Lee et al., 2005a). Virus-induced gene silencing (VIGS) of *EcCRC* from California poppy (*Eschscholzia californica*) showed that *EcCRC* is required for floral meristem termination, gynoecium differentiation, and ovule initiation (Orashkova et al., 2009). However, in monocot species such as rice (*Oryza sativa*) and maize (*Zea mays*), *CRC* was found to regulate leaf development and flower determinacy. The *CRC* homologous gene in rice *DROOPING LEAF* (*DL*) is expressed uniformly in carpel primordia and the central region of leaf primordia (Nagasawa et al., 2003; Yamaguchi et al., 2004). In *dl* mutants, the blade and sheath fail to form midrib and the plants produce drooping leaves and abnormal gynoecium with extra floral organ (Nagasawa et al., 2003; Yamaguchi et al., 2004). *DL* was shown to genetically interact with *OsMADS3* in the regulation of floral meristem determinacy and floral organ identity (Li et al., 2011). *Arabidopsis CRC* has two orthologs in maize, *drooping leaf 1* (*drl1*) and *drl2*, which are expressed in leaf primordia and lateral floral primordia, and regulate plant architecture development and floral organ determinacy, respectively (Ishikawa et al., 2009; Strable et al., 2017; Strable and Vollbrecht, 2019; Tian et al., 2019). *DRL1* was shown to physically interact with *LIGULELESS1* (*LG1*), and such interaction can repress the *LG1*-mediated activation of a B3-domain transcription factor (*ZmRAVL1*) transcription, leading to decreased endogenous brassinosteroids level and reduced leaf angle in maize (Tian et al., 2019). Although *CRC* has been shown to play important roles in the development of flowers, nectary, and leaves, the function of *CRC* during fruit development is largely unknown. Cucumber bears typical pepo fruits with unique developmental patterns; many fundamental genes undergo functional divergence between cucumber and other plant species, which provides a good opportunity to understand *CRC* function in fruit growth (Pan et al., 2020; Wang et al., 2020; Liu et al., 2021).

In our previous study (Pan et al., 2017b), using an  $F_{2:3}$  mapping population derived from a cross between the semi-wild Xishuangbanna (XIS) cucumber accession “WI7167” and the

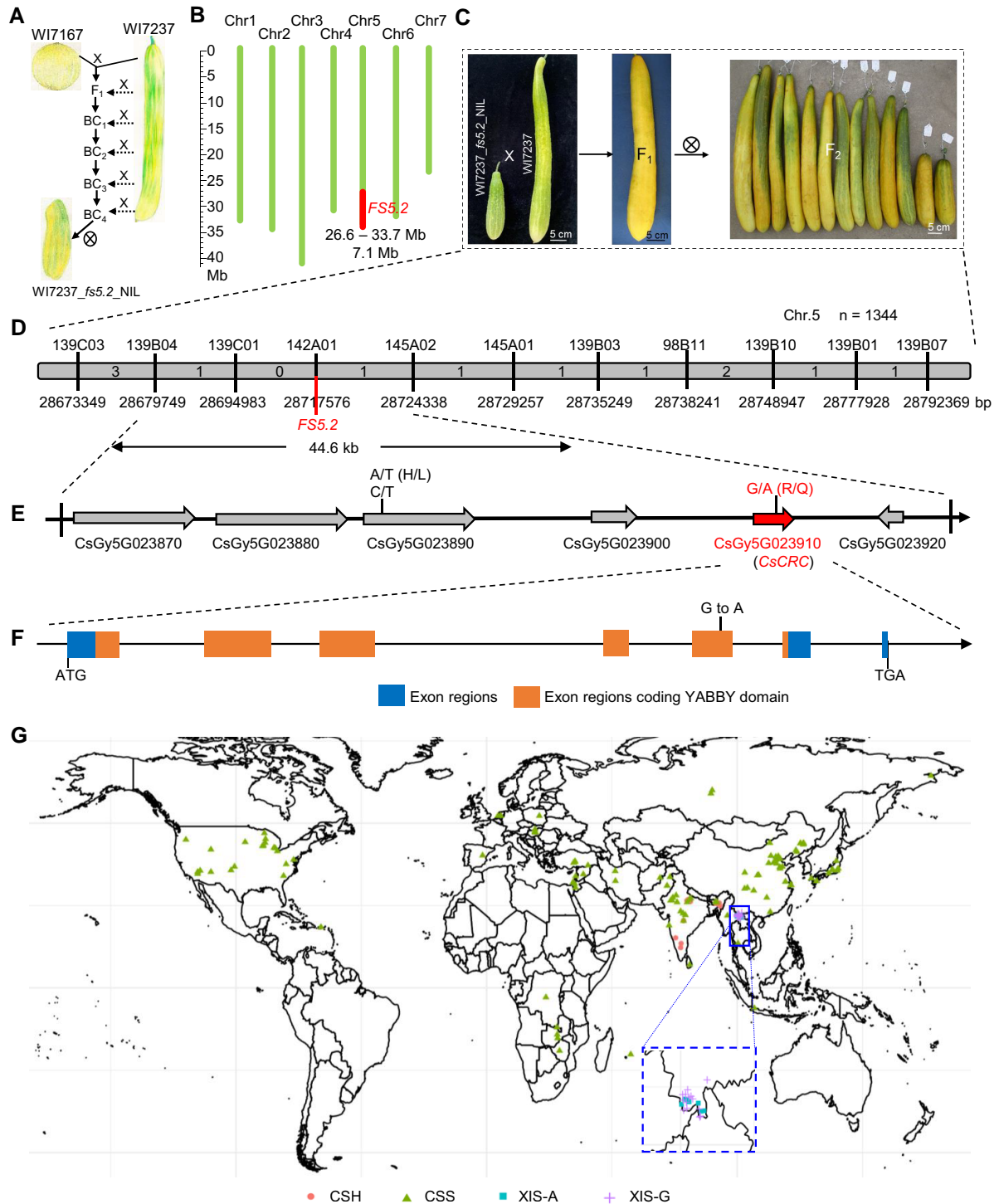
inbred line “WI7200” derived from plant introduction (PI) line PI 249561, we identified the major-effect fruit shape/size QTL *FS5.2* ( $R^2 = 36.7\%–51.0\%$ ), in which the WI7167 allele was associated with shorter fruit ( $R^2 = 17.8\%–24.7\%$ ). In this study, we cloned this QTL and identified a nonsynonymous G-to-A single-nucleotide polymorphism (SNP) in the homolog of *Arabidopsis CRABS CLAW* (designated *CsCRC*) that was associated with fruit length variation conferred by the *FS5.2* locus. The *CsCRC<sup>A</sup>* allele only exists in a few semi-wild Xishuangbanna (XIS) cucumbers with short or round fruits. Expression analysis showed that *CsCRC* transcripts specifically accumulated in the gynoecium and nectary. Overexpression of *CsCRC<sup>G</sup>* rescued the short-fruit phenotype caused by *CsCRC<sup>A</sup>*. Knockdown of *CsCRC* in cucumber led to reduced fruit length. Further, we determined that the auxin-responsive protein gene *CsARP1* is a downstream target of *CsCRC* that promotes fruit elongation mainly through promoting cell expansion. We propose a working model to explain the functions of different *CsCRC* alleles in regulating fruit length.

## Results

### The candidate gene of *FS5.2* is a homolog of *Arabidopsis CRABS CLAW*

To investigate the roles of *FS5.2* in fruit growth, we developed a pair of near-isogenic lines (NIL) WI7237\_*fs5.2*\_NIL via marker-assisted backcrossing by introducing the WI7167 *FS5.2* allele into WI7237, which is a landrace with very long (>50 cm) fruit (Figure 1A). Genotyping-by-sequencing (GBS) of the NIL pair suggested that there is a 7.1-Mb target region from WI7167 that was introgressed into WI7237\_*fs5.2*\_NIL, while all other regions were the same as the recurrent parent WI7237 (Figure 1B). Compared to WI7237, the WI7237\_*fs5.2*\_NIL (*fs5.2*\_NIL hereinafter) displayed a significant reduction in fruit length. This change was clearly visible from the ovary stage but became most obvious at the mature fruit stage. The averaged mature fruit length of *fs5.2*\_NIL was 15~20 cm shorter than that of WI7237 under either greenhouse or open-field growing conditions, corresponding to an approximately 34~39% reduction in fruit length (Figure 1C).

For fine-mapping of the major-effect QTL *FS5.2*, we genotyped a large NIL-derived  $F_2$  population with 1,344 individuals with flanking markers 139B07 and 139C03, which were physically 119 kb apart (Figure 1D). Among the 1,344  $F_2$  plants, 1,023 plants had long fruits and the remaining 321 set short fruits, which was consistent with the expected 3:1 segregation ratio ( $P = 0.5298$  in  $\chi^2$  test), and suggested that a single recessive gene controls fruit length variation in this population. Genotyping with these two flanking markers identified 12 recombinants. We developed new polymorphic insertion/deletion (InDel) or SNP markers in the 119-kb region and used them to genotype the 12 recombinant individuals (Supplemental Table S1). Several rounds of recombinant identification and marker discovery narrowed



**Figure 1** Fine-mapping of the major-effect fruit shape QTL *F55.2*. **A**, Schematic diagram of the generation of WI7237\_fs5.2\_NIL (NIL). **B**, Whole-genome sequencing results of WI7237\_fs5.2\_NIL, using the draft genome of Gy14 V2.0 as reference. Green box, WI7237 background; red box, WI7167. **C**, Fruits of WI7237\_fs5.2\_NIL, WI7237, their F<sub>1</sub>, and F<sub>2</sub> plants. **D–F**, Linkage mapping with 1,344 F<sub>2</sub> plants delimiting *F55.2* to a 44.6-kb region with six predicted genes in the Gy14v2.0 reference genome, in which the *CsCRC* (red) is the most possible candidate of *F55.2*. **F**, *CsCRC* has seven exons. **G**, Worldwide geographical distribution of the 165 cucumber lines used for investigating the *CsCRC*<sup>C</sup> and *CsCRC*<sup>A</sup> alleles among a natural population of wild cucumbers (CSH), cultivated cucumbers (CSS), and semi-wild cucumbers (XIS). XIS-A and XIS-G indicate XIS cucumbers carrying the *CsCRC*<sup>A</sup> allele and *CsCRC*<sup>C</sup> allele, respectively.

FS5.2 to a 44.6-kb region between markers 139B04 and 145A02, with the two markers 139C01 and 142A01 co-segregating with FS5.2 (Figure 1D). This 44.6-kb region contains six annotated genes based on the reference genome (Gy14v2.0, <http://cucurbitgenomics.org/organism/16>) (Figure 1E and Supplemental Table S2). To identify the possible candidate locus behind FS5.2, we looked for sequence variation in this region between WI7167 and WI7237 by alignment of the resequencing reads against the Gy14v2.0 genome. We identified 35 SNPs or InDels, of which 32 were located in intergenic regions, two were in the first exon of CsGy5G023890 encoding a ribosome maturation factor, and one in the 5th exon of CsGy5G023910 (gene #5 in the interval). We examined allelic variation at the two SNPs in CsGy5G023890 in 11 semi-wild Xishuangbanna (XIS) inbred lines with varying fruit lengths and observed no association between the type of allele and fruit length. We also examined the expression patterns of the genes in the 44.6-kb region (all except gene #5) and did not find any differences in their expression levels between the NILs (Supplemental Figure S1). These data supported CsGy5G023910 as the most possible candidate for fs5.2. The nonsynonymous SNP (G-to-A) in CsGy5G023910, a homolog of Arabidopsis CRABS CLAW (CRC), results in an amino acid substitution from Arginine (R) to Glutamine (Q) in the conserved DNA-binding YABBY domain. Genotyping of this SNP revealed that the A allele is specific for the XIS inbred lines bearing short and round fruits, suggesting that CsCRC is the likely candidate for fs5.2. CsCRC is predicted to have seven exons and encodes a YABBY protein of 175 amino acids (Figure 1F). All YABBY family proteins contain a conserved zinc finger and a YABBY domain (Siegfried et al., 1999) (Supplemental Figure S2A). Similar to its Arabidopsis and rice counterparts, CRC is present as a single-copy gene in cucumber (Supplemental Figure S2B).

### The CsCRC<sup>A</sup> allele is specific for the XIS semi-wild cucumber lineage with short or round fruits

FS5.2/CsCRC of short fruited WI7167 and fs5.2\_NIL carried the A allele (short allele, CsCRC<sup>A</sup>), while long-fruited WI7237 harbored the G allele (long allele, CsCRC<sup>G</sup>). To explore the allelic distribution of CsCRC among a natural cucumber population, we examined the genotypes for this SNP in 165 cucumber accessions consisting of 13 wild (CSH), 131 cultivated (CSS), and 21 semi-wild (XIS) cucumbers (Supplemental Data Set S1). Seven of the 165 lines carried the A allele and all other had the G allele. Interestingly, all cucumbers (XIS01 to XIS07) with the A allele belonged to the semi-wild XIS cucumbers, including WI7167 (XIS01) (Figure 1G); five of the seven accessions set round fruit. However, some round-fruited XIS cucumbers carried the CsCRC<sup>G</sup> allele (e.g. XIS10, XIS11, and XIS19). These observations suggest that the round-fruit CsCRC<sup>A</sup> allele is unique to the semi-wild XIS lineage, which appears to have become

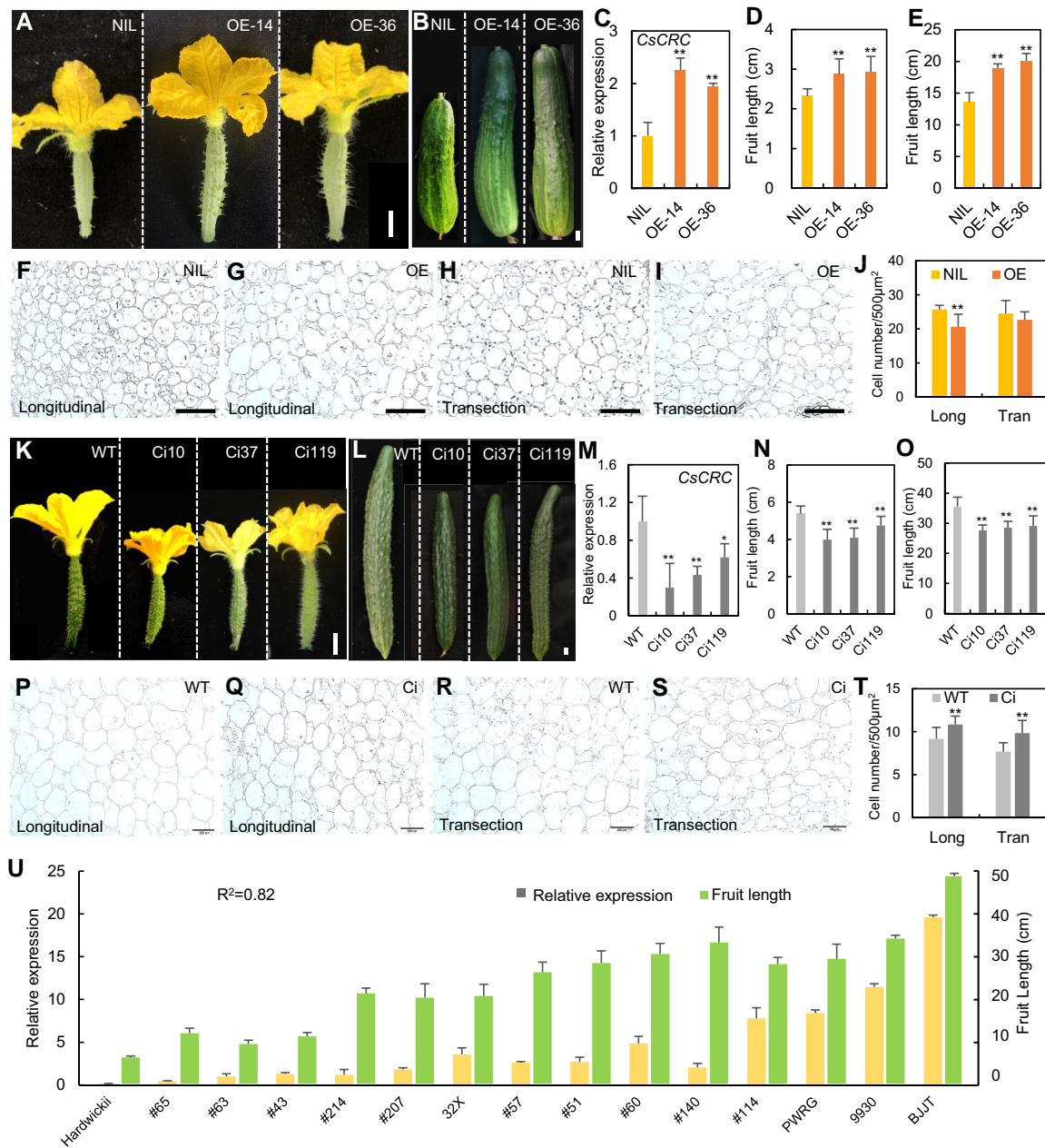
a rare allele due to free gene flow within the XIS population or between XIS and other non-XIS populations.

### CsCRC<sup>G</sup> overexpression rescues the fruit length phenotype in NIL

To verify the role of CsCRC in fruit length regulation, we performed a complementation test by introducing a construct harboring the CsCRC<sup>G</sup> allele driven by the cauliflower mosaic virus (CaMV) 35S promoter into the short-fruit fs5.2\_NIL (NIL). Compared to the NIL plants, CsCRC<sup>G</sup>-OE lines produced elongated fruits at anthesis day (AD) (Figure 2A), the commercial stage (10 DAA, day after anthesis) (Figure 2B), and the mature stage (30 DAA) (Supplemental Figure S3, A and B). In the transgenic lines, CsCRC<sup>G</sup> expression showed approximately a two-fold increase compared to the NIL; accordingly, the fruit length at 10 DAA of OE-14 ( $18.7 \pm 1.0$  cm) and OE-36 ( $20.8 \pm 1.5$  cm) was significantly longer than that of the NIL control ( $13.2 \pm 1.9$  cm), approximate 57.6% increase in fruit length (Figure 2, C and E). We observed the longer fruit length at different developmental stages (Figure 2, A–E), suggesting that CsCRC<sup>G</sup> positively regulates fruit elongation in cucumber. Under a microscope, longitudinal sections of the mesocarp of OE-14 fruit at 10 DAA showed larger but fewer cells ( $20.7 \pm 3.6$  per  $500 \mu\text{m}^2$  cell area) relative to the NIL control ( $25.7 \pm 1.2$ ) (Figure 2, F and G). The cross-sections revealed no difference between NIL and OE lines (Figure 2, H–J). These data suggest that CsCRC<sup>G</sup> promotes fruit elongation mainly through longitudinal cell expansion in cucumber.

### CsCRC<sup>G</sup> is a positive regulator for fruit elongation in cucumber

To verify the function of CsCRC in fruit length regulation, we transformed an RNA interference (RNAi) construct containing a specific fragment (351 bp) of CsCRC (CsCRC-RNAi) into the long-fruit cucumber inbred line R1461 (referred to as wild type [WT] thereafter, harboring the CsCRC<sup>G</sup> allele). We obtained 15 independent transgenic lines, of which we selected three representative lines (Ci10, Ci37, and Ci119) for further characterization. Compared to the WT, the fruit length in the transgenic lines was greatly reduced, with an apparent positive correlation between fruit length and CsCRC expression level ( $R^2 = 0.95$ ) (Figure 2, K–M). The average fruit length at 10 DAA was  $35.6 \pm 3.2$  cm in WT plants, but only  $27.7 \pm 1.7$  cm in the most severe transgenic line Ci10, representing a 28.5% decrease in fruit length (Figure 2N). Fruits at AD and 30 DAA were also significantly shorter in CsCRC<sup>G</sup>-RNAi plants than in WT plants (Figure 2O; Supplemental Figure S3C). Longitudinal and transverse sections of mesocarp at 10 DAA showed reduced cell size in the transgenic lines (Figure 2, P–T). Therefore, CsCRC<sup>G</sup> acted as an important activator for fruit elongation by promoting cell expansion in cucumber. Interestingly, in addition to fruit length suppression, we also observed abnormal development of flowers and nectaries in CsCRC<sup>G</sup>-RNAi plants. Compared to



**Figure 2** *CsCRC*<sup>G</sup> positively regulates fruit length elongation in cucumber. A–J, Overexpression of *CsCRC*<sup>G</sup> (OE) in WI7237\_55.2\_NIL (NIL) results in longer fruit. A, Cucumber female flowers at anthesis day. B, Cucumber fruits at the commercial stage. C, Relative *CsCRC* expression levels in cucumber ovaries. D and E, Mean fruit length at anthesis day (AD) and 10 days after anthesis (10 DAA). F–I, Longitudinal sections (F and G) and transverse sections (H and I) in the mesocarp of NIL and OE fruits at 10 DAA. J, Cell numbers in the mesocarp of NIL and OE fruits at 10 DAA. K–T, Knockdown of *CsCRC*<sup>G</sup> by RNAi (Ci) in cucumber leading to shorter fruit. K, Cucumber female flowers at anthesis day. (L) Cucumber fruits at the commercial stage. (M, Relative *CsCRC* expression levels in cucumber ovaries. (N and O) Mean fruit length at anthesis day (AD) and 10 DAA. P–S, Longitudinal sections (P and Q) and transverse sections (R and S) in the mesocarp of WT and Ci fruits at 10 DAA. (T) Cell numbers in mesocarp of WT and *CsCRC*<sup>G</sup>-RNAi fruits at 10 DAA. (U, *CsCRC*<sup>G</sup> expression and fruit length in natural cucumber lines. Scale bars, 1 cm in A, B, K, L and 200  $\mu$ m in F–I, P–S. Long, longitudinal section; Tran, transverse section. Data are means  $\pm$  SE. \* $P < 0.05$ ; \*\* $P < 0.01$ . Three biological replicates and three technique replicates were performed for each RT-qPCR analysis.

WT plants, stigmas were smaller and malformed, nectar production was lower in female flowers of Ci plants, while degenerated carpels started to produce style-like tissues in male flowers of transgenic plants (Supplemental Figure S3, D–F).

To further confirm the effect of the *CsCRC*<sup>G</sup> and *CsCRC*<sup>A</sup> alleles on fruit elongation, we transformed the cucumber inbred line R1461 with the constructs 35S:*CsCRC*<sup>G</sup> and 35S:*CsCRC*<sup>A</sup> into (Supplemental Figure S4). We obtained four *CsCRC*<sup>G</sup> and nine

CsCRC<sup>A</sup> overexpressing transgenic lines, and we selected two representative lines for each construct. Overexpression of CsCRC<sup>G</sup> led to longer fruit (Supplemental Figure S4, A–D). The average fruit length at 10 DAA was  $33.6 \pm 1.0$  cm in CsCRC<sup>G</sup> overexpression (G1) plants, compared to  $30.6 \pm 0.7$  cm in the WT (Supplemental Figure S4D). However, elevated expression of CsCRC<sup>A</sup> did not cause any visible change in fruit length in the transgenic lines (Supplemental Figure S4, E–J). Therefore, CsCRC<sup>G</sup>, but not CsCRC<sup>A</sup>, acts as a positive regulator for fruit elongation in cucumber.

We further investigated CsCRC<sup>G</sup> expression in 15 cucumber lines with different fruit lengths, including the wild cucumber *hardwickii* and several cultivated inbred lines (Figure 2U). Consistent with data presented above, CsCRC<sup>G</sup> expression was positively correlated with fruit length ( $R^2 = 0.82$ ) among these lines. For example, the expression of CsCRC<sup>G</sup> in the Chinese Long cucumber 9,930 was 89-fold higher than that in the wild progenitor *hardwickii* with short-fruit (Figure 2U). These data further support our conclusion that CsCRC<sup>G</sup> positively promotes fruit elongation in cucumber.

### CsCRC is highly expressed in the carpel and nectary of cucumber

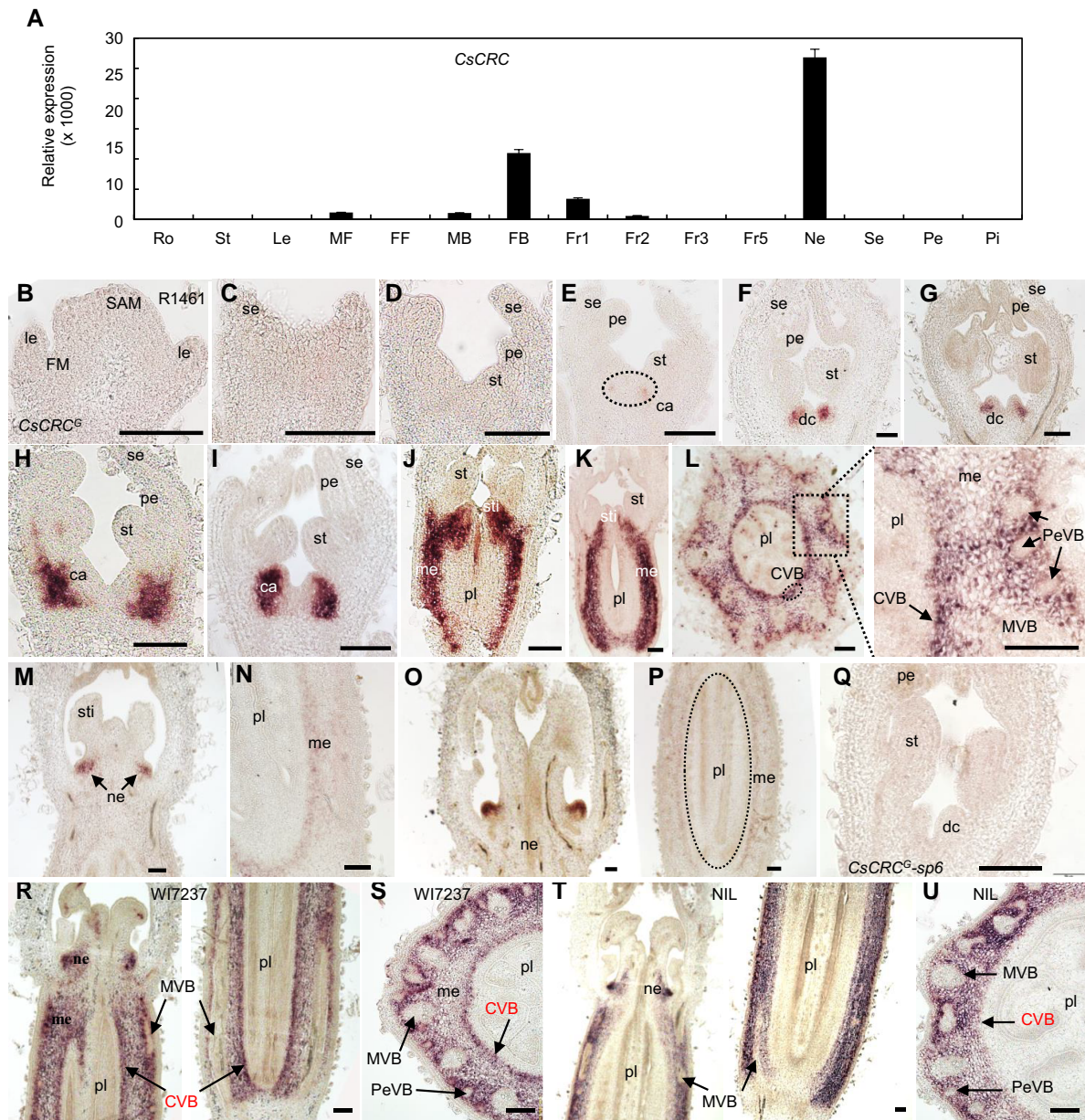
To investigate the expression pattern of CsCRC, we performed reverse transcription quantitative PCR (RT-qPCR) and in situ hybridization in different tissues of the inbred line R1461. RT-qPCR analysis showed that CsCRC transcript levels are specifically high in reproductive organs including flowers and fruits, while we detected no signal in vegetative organs such as stems and leaves (Figure 3A). We measured the highest expression of CsCRC in nectary, followed by female flower buds and young ovaries. As cucumber fruit development progressed, we witnessed a gradual decrease in CsCRC expression (Figure 3A). mRNA in situ hybridization in R461 revealed the detailed expression pattern of CsCRC in floral organs (Figure 3, B–Q). We first detected CsCRC transcripts in cells that would later develop into gynoecium primordia at stage 5 (Figure 3, B–E). During male flower development, CsCRC transcripts accumulated abaxially in degenerated carpels (Figure 3, F and G). At the female flower differentiation stages, we detected strong CsCRC signal in the abaxial side of gynoecium primordia (Figure 3, H and I). With the elongation of carpel primordia, we observed that CsCRC transcripts are abundant in the mesocarp of developing ovary at stages 7 and 8 (Figure 3, J–L). In cross-sections of young fruits, CsCRC was expressed in pulp cells among mesocarp and the carpellary vascular bundle (CVB), while we detected no hybridization signal in the main vascular bundle (MVB) or peripheral vascular bundle (PeVB) (Figure 3L). Consistent with the RT-qPCR analysis, CsCRC was expressed throughout nectary primordia development (Figure 3, M and O). Starting from stage 8-3, CsCRC signal in the mesocarp became weaker (Figure 3, N and P), but remained strong in the nectary (Figure 3O). We detected no signal when using the CsCRC sense probe for hybridization (Figure 3Q).

We also performed mRNA in situ hybridization in young ovaries from the NIL pair (Figure 3, R–U). In WI7237, the CsCRC<sup>G</sup> expression pattern was the same as that in R1461 cucumber, with transcripts accumulating in the mesocarp, CVB, and nectary (Figure 3, R and S). Interestingly, CsCRC<sup>A</sup> expression in the NIL was similar to that of WI7237, except for its absence in CVB (Figure 3, T and U).

### CsARP1 is the downstream target of CsCRC

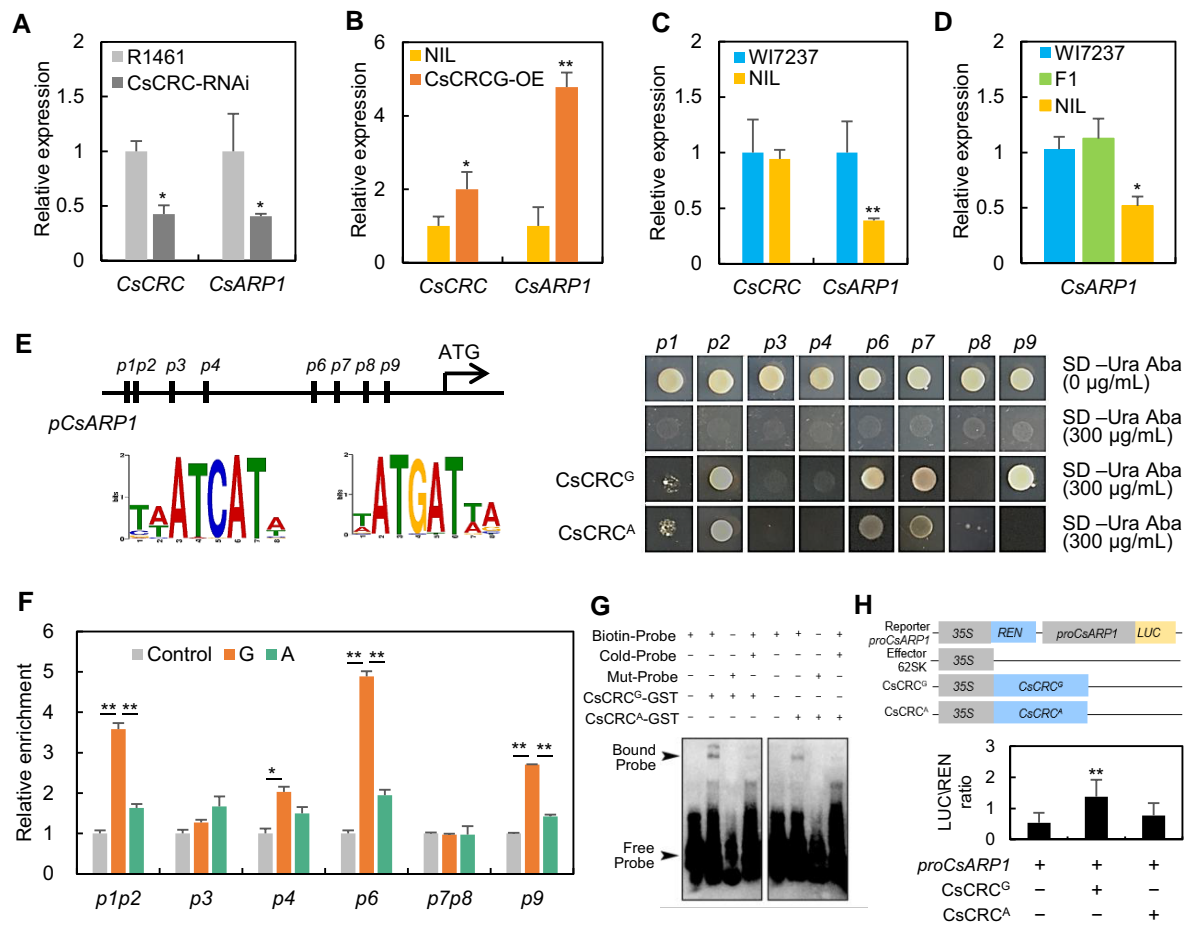
To identify putative downstream targets and gene pathways mediated by CsCRC, we performed transcriptome deep sequencing (RNA-seq) analysis using young ovaries from R1461 and CsCRC-RNAi10 line (*Ci10*) as samples. We identified 80 genes as being significantly downregulated and four genes as being upregulated in the CsCRC-RNAi line compared with WT (Supplemental Data Set S2). Of these, we noticed that the auxin-responsive protein gene *CsARP1* was significantly downregulated in the CsCRC-RNAi line (Supplemental Data Set S2). Relative expression of CsCRC and *CsARP1* in different cucumber lines showed that *CsARP1* transcript levels are low in the CsCRC-RNAi line and high in the CsCRC<sup>G</sup>-OE lines (Figure 4, A and B). *CsARP1* expression was also significantly lower in the NIL (harboring the A allele) compared to WI7237 (harboring the G allele) (Figure 4, C and D). These data suggest that *CsARP1* is a good downstream target of CsCRC during fruit length regulation in cucumber.

To confirm *CsARP1* as a direct target of CsCRC, we analyzed the promoter region of *CsARP1*, which harbored and nine putative CRC-binding sites (p1–p9) in the ~2,000-bp region upstream of the transcription start site (Figure 4E). The nucleotide sequences of putative CRC-binding sites are [A/T] ATCAT[A/T] or [T/A]ATGAT[T/A] (O'Malley et al., 2016). A yeast one-hybrid (Y1H) assay showed that CsCRC<sup>G</sup> can bind to five fragments (p1, p2, p6, p7, and p9) of *CsARP1*, while CsCRC<sup>A</sup> failed to bind to the p9 fragment, and also exhibited a weaker binding affinity towards the p6 and p7 fragments (Figure 4E). We also performed chromatin immunoprecipitation-quantitative PCR (ChIP-qPCR) using an anti-Myc antibody to pull down Myc-tagged CsCRC<sup>G</sup> and CsCRC<sup>A</sup> proteins from cucumber calli. We determined that the fragments p1, p2, p4, p6, and p9 containing the putative CRC-binding motif are enriched in the chromatin associated with CsCRC<sup>G</sup>, whereas only the p6 fragment was significantly enriched following CsCRC<sup>A</sup> immunoprecipitation (Figure 4F). Electrophoretic mobility shift assays (EMSA) also showed stronger binding to the *CsARP1* p6 element by recombinant CsCRC<sup>G</sup> fused to glutathione S-transferase (GST) than by CsCRC<sup>A</sup>-GST (Figure 4G). To investigate whether CsCRC<sup>G</sup> and CsCRC<sup>A</sup> regulate *CsARP1* expression in vivo, we performed transient expression assays in *Nicotiana benthamiana* leaves, whereby a 2,506-bp *CsARP1* promoter fragment was cloned upstream of the firefly luciferase (*LUC*) reporter gene. We also placed the CsCRC<sup>G</sup> and CsCRC<sup>A</sup> coding sequence under the control of the 35S promoter as the effectors (Figure 4H). We determined that



**Figure 3** Expression analyses of CsCRC in cucumber. **A**, Relative CsCRC expression levels in different organs of cucumber as detected by RT-qPCR. Ro, Root; St, Stem; Le, leaf; MF, Male Flower; FF, Female Flower without ovary; MB, Male Bud; FB, Female Bud; Fr1, 1-cm Fruit; Fr2, 2-cm Fruit; Fr3, 3-cm Fruit; Fr5, 5-cm Fruit; Fr1~Fr5 indicate ovaries without corollas; Ne, Nectary; Se, Sepal; Pe, Petal; Pi, Pistil. **B–Q**, *In situ* hybridization of CsCRC in inbred cucumber line R1461 carrying the G allele. No signal was detected in SAM (**B**), the 2nd (**C**) or 4th (**D**) floral bud differentiation stage. **E**, CsCRC signals were detected in cells where the gynoecium primordia will initiate at stage 5. **F** and **G**, CsCRC transcripts accumulated abaxially in the degenerated carpels in the male flower bud at stage 7 (**F**) and 8 (**G**). **H–P**, CsCRC expression pattern among female organs at different developmental stages. **H–I**, CsCRC was strongly expressed in carpel primordia at stage 6. **J** and **K**, CsCRC transcripts accumulated in the mesocarp region among young ovaries at stage 7 (**J**) and 8-1 (**K**). **L**, Cross-section of young ovaries at stage 8-2 showing that CsCRC is specifically expressed in mesocarp and carpellary vascular bundles (CVB), and absent in main and peripheral vascular bundles (MVB and PeVB). **M–P**, CsCRC signals in nectary primordia and ovaries at stage 8-3 (**M** and **N**) and stage 9 (**O** and **P**). (**Q**) CsCRC sense probe was used as a negative control in Male flower buds at stage 7. **R–U**, CsCRC expression pattern in female flower buds at stage 8 in WI7237 harboring the G allele (**R** and **S**) and NIL harboring the A allele (**T–U**). **R–S**, Longitudinal (**R**) and cross (**S**) section of WI7237 ovary showing that CsCRC<sup>G</sup> signals accumulate in CVB and the mesocarp region without MVB and PeVB. **T** and **U**, Longitudinal (**T**) and cross (**U**) section of NIL ovary showing that CsCRC<sup>A</sup> signals accumulate in the mesocarp region without MVB and PeVB, but absent in CVB. SAM, shoot apical meristem; le, leaf; FM, flower meristem; se, sepal; pe, petal; st, stamen; ca, carpel; dc, degenerated carpel; me, mesocarp; pl, placenta; CVB, carpellary vascular bundle; MVB, main vascular bundle; PeVB, peripheral vascular bundle; ne, nectary. Scale bars, 100  $\mu$ m.





**Figure 4** CsCRC<sup>G</sup> directly activates the transcription of the putative auxin-response protein CsARP1 in cucumber. A–D, Relative CsCRC and CsARP1 expression levels in young ovaries of different cucumber lines. E, Y1H assay showing that CsCRC<sup>G</sup> and CsCRC<sup>A</sup> bind to the CsARP1 promoter. The binding sites were [A/T]ATCAT[A/T] or [T/A]ATGAT[T/A] (O'Malley et al., 2016). F, ChIP-qPCR assay showing *in vivo* binding of CsCRC<sup>G</sup> and CsCRC<sup>A</sup> to the CsARP1 promoter. G, EMSA showing that the binding of CsCRC<sup>G</sup> to CsARP1-P6 is stronger than that of CsCRC<sup>A</sup>. Labeled probes containing the p6 element were synthesized and labeled with biotin. The cold (unlabeled) probe and mutated p6 element (labeled) were used as competitor and mutant probes, respectively. H, Schematic diagram showing the reporter and effector constructs for the LUC/REN assays. LUC activities were measured after transient expression of *proCsARP1:LUC* with 35S:CsCRC<sup>G</sup> or 35S:CsCRC<sup>A</sup> in *N. benthamiana* leaves. Data are means  $\pm$  SE. \* $P < 0.05$ ; \*\* $P < 0.01$ .

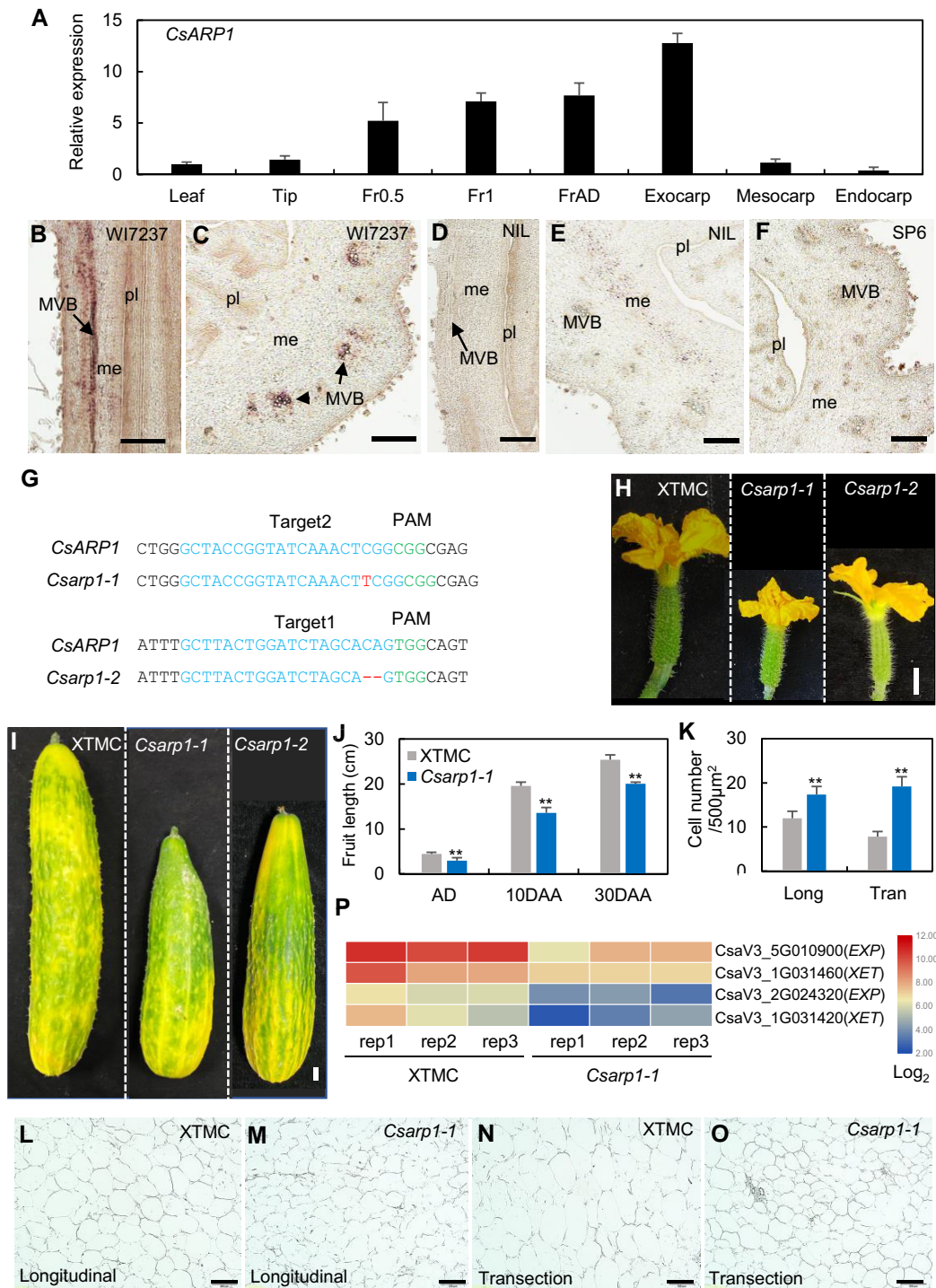
relative LUC activity is significantly higher upon co-infiltration of 35S:CsCRC<sup>G</sup> with *ProCsARP1:LUC*, whereas co-infiltrating 35S:CsCRC<sup>A</sup> with *ProCsARP1:LUC* resulted in no increase in LUC activity over that seen in the negative control (Figure 4H). These data support the notion that CsCRC<sup>G</sup>, but not CsCRC<sup>A</sup>, directly promotes the transcription of CsARP1. This hypothesis is consistent with the down-regulation of CsARP1 in the NIL (harboring the A allele) compared to WI7237 (carrying the G allele), in which the expression of CsCRC was unchanged (Figure 4C).

### CRISPR/Cas9-engineered mutations in CsARP1 resulted in decreased fruit length

To explore whether CsARP1 and CsCRC display overlapping expression patterns, we performed RT-qPCR analysis, which established that CsARP1 expression is high in fruit exocarp

(Figure 5A). mRNA in situ hybridization revealed that CsARP1 mRNA signals accumulate in young ovaries (Figure 5, B–F). In WI7237, CsARP1 was expressed in the pulp cells and main vascular bundles in fruits (Figure 5, B and C). CsARP1 expression levels were significantly lower in the NIL, with weak signal detected in the mesocarp cells of fruits (Figure 5, D and E).

To gain insight into the function of CsARP1 in cucumber fruit development, we generated a construct targeting CsARP1 for clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated nuclease 9 (Cas9)-mediated genome editing and introduced this construct into the cucumber inbred line Xin Tai Mi Ci (XTMC). Sequencing identified two *Csarp1* lines, with *Csarp1-1* having a 1-bp insertion in the target 2 region, while *Csarp1-2* harbored a 2-bp deletion in the target 1 region (Figure 5G). Compared to the XTMC control, fruit length



**Figure 5** Knockout of *CsARP1* by CRISPR/Cas9 result in shorter fruit in cucumber. **A**, Relative *CsARP1* expression levels in cucumber lateral organs. **B–F**, *In situ* hybridization of *CsARP1* in young ovaries of WI7237 and NIL. Longitudinal section (**B**) and transverse section (**C**) of young ovaries showing that *CsARP1* is specifically expressed in the mesocarp and main vascular bundles in WI7237. **D** and **E**, In NIL, *CsARP1* is weakly expressed in the mesocarp. **F**, *CsARP1* sense probe was used as negative control. **G**, *CsARP1* knockout lines by CRISPR/Cas9 (*Csarp1*) identified by sequencing. InDels are indicated by red dashed lines and letters. Blue letters, sgRNA targets; green letters, protospacer-adjacent motif (PAM). **H**, Cucumber female flowers at anthesis day. **I**, Cucumber fruits at commercial stage. **J**, Mean fruit length at anthesis day (AD) and 10 DAA. **K**, Cell number in the mesocarp of XTMC and *Csarp1* fruits at 10 DAA. Longitudinal sections (**L** and **M**) and transverse sections (**N** and **O**) in the mesocarp of XTMC and *Csarp1* fruits at 10 DAA. **P**, Heatmap showing that the expression of cell expansion-related *EXP* and *XET* genes is significantly lower in *Csarp1* plants compared to the XTMC control. MVB, main vascular bundle; me, mesocarp; pl, placenta. Scale bars, 100 µm in **B–F**, 1 cm in **H** and **I**, 200 µm in **L–O**. Data are means ± SE. \**P* < 0.05; \*\**P* < 0.01. Three biological replicates and three technique replicates were performed for each RT-qPCR analysis.

was significantly reduced in the *Csarp1* lines (Figure 5, H–J). The average fruit length at 10 DAA was  $19.6 \pm 1.4$  cm in XTMC and  $13.5 \pm 2.2$  cm in *Csarp1*, or a 31.1% reduction in fruit length (Figure 5J). Longitudinal sections and transverse sections of mesocarp from 10-DAA fruit showed that cell size is greatly reduced in the *Csarp1-1* mutant (Figure 5, K–O), suggesting that CsARP1 promotes fruit elongation through cell expansion.

To identify putative downstream genes of CsARP1, we performed an RNA-seq analysis using the *Csarp1-1* mutant, with three biological replicates. We identified 16 downregulated cell wall-related genes in the *Csarp1-1* mutant relative to the wild-type control (Supplemental Data Set S3). Gene Ontology (GO) term enrichment analysis showed that cell wall-related genes are significantly enriched among the downregulated genes in both the *Csarp1* mutant and *CsCRC*-RNAi plants (Supplemental Figure S5). For example, xyloglucan endotransglucosylase (XET) and expansins (EXPs) have been reported to mediate plant growth by reorganization of cell wall and facilitating non-enzymatic cell wall loosening, respectively (Van Sandt et al., 2007; Mayorga-Gómez and Nambesan, 2020). Indeed, from RNA-seq data, the expression of two XET and two EXP genes was significantly downregulated in the *Csarp1-1* mutant compared to the XTMC control (Figure 5P; Supplemental Data Set S3). Similarly, both XET genes were expressed to lower levels in the *CsCRC*-RNAi line relative to WT (Supplemental Data Set S2). These data support the notion that *CsCRC* and CsARP1 promote fruit elongation through cell wall-mediated cell expansion in cucumber.

## Discussion

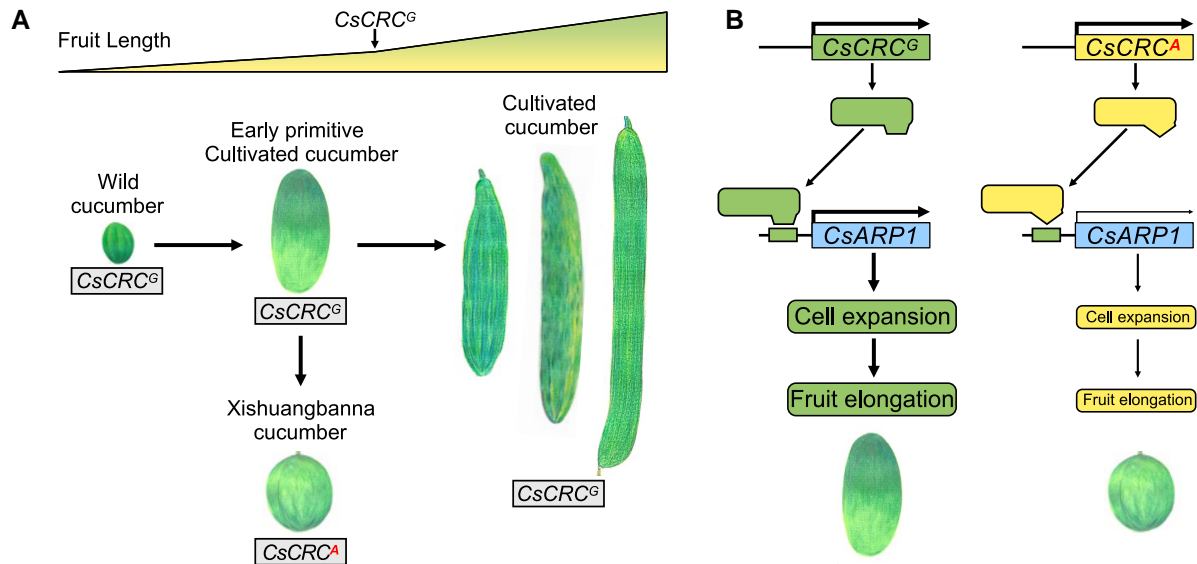
### Allelic variation at *CsCRC* regulates fruit elongation in cucumber

Natural variation in regulatory genes plays particularly important roles in crop trait divergence. In the two subspecies of rice, *Oryza sativa* subs. *indica* and sub. *japonica*, natural variation in the nitrate transporter gene *NRT1.1B-indica* (with two SNPs in the coding region) contributes to higher nitrogen-use efficiency and improved grain (Hu et al., 2015). In maize, two InDels and three SNPs upstream of *ZmDREB2.7* (*DEHYDRATION RESPONSE ELEMENT-BINDING PROTEIN 2.7*) affect drought tolerance (Liu et al., 2013). A 2-bp sequence polymorphism in *UPA2* (*Upright Plant Architecture 2*) is associated with the different binding affinity of *DRL1* and downstream pathway genes, with the teosinte *UPA2* allele showing narrow leaf angles and promoting yield under high-density planting in maize (Tian et al., 2019). Therefore, natural variations and their favorable alleles can be utilized as important genetic resources for trait improvement in crop breeding.

Fruit length is a key agronomic trait affecting fruit shape and fruit size in cucumbers. The extent and nature of the

genetic variation involved in cucumber domestication are largely unknown. Despite the identification of more than 20 consensus fruit length QTLs in cucumber (Weng et al., 2015; Che and Zhang, 2019; Gao et al., 2020; Pan et al., 2020; Sheng et al., 2020), very few have been fine-mapped or functionally characterized. Here, we showed that *CsCRC* is the best candidate gene for the major-effect fruit length QTL *FS5.2* (Figures 1 and 2). A nonsynonymous G-to-A SNP at *CsCRC* differentiated the two parental lines, with the *CsCRC<sup>A</sup>* allele from WI7167 contributing to short fruit, and the *CsCRC<sup>G</sup>* allele from WI7167 being associated with long fruit. The *fs5.2\_NIL* displayed a 34~39% reduction in fruit length relative to the recurrent parent line WI7237, suggesting that the *CsCRC<sup>A</sup>* allele had a negative effect on fruit elongation in cucumber. Complementation tests showed that overexpression of *CsCRC<sup>G</sup>* in the NIL resulted in longer fruit with larger cells, whereas *CsCRC<sup>G</sup>* knockdown led to shorter fruit and inhibited cell expansion (Figure 2). Overexpression of *CsCRC<sup>G</sup>*, but not *CsCRC<sup>A</sup>*, also produced longer fruits in cucumber R1461 (Supplemental Figure S4). In natural cucumber lines, *CsCRC<sup>G</sup>* expression was positively correlated with fruit length (Figure 2U). These results indicate that the *CsCRC<sup>G</sup>* allele displays positive effects on fruit elongation in cucumber.

During domestication from the wild progenitor to cultivated cucumber, longer fruits were preferred and selected for to achieve higher yields (Bo et al., 2015). The *CsCRC<sup>A</sup>* allele is a rare allele specific to the semi-wild *XIS* lineage, while most cultivated or wild cucumbers examined here carried the long *CsCRC<sup>G</sup>* allele (Figure 1, Supplemental Data Set S1). The *CsCRC<sup>A</sup>* allele may result in a point mutation in the *XIS* cucumber and lead to reduced fruit length (Figure 6A). Selection of this A allele was likely from a more primitive cultivated cucumber carrying the G allele, because *XIS* cucumber is known to have been domesticated from cultivated cucumber rather than from wild cucumber (Bo et al., 2015; Pan et al., 2017a). Not all round-fruited *XIS* lines carried the short allele (*CsCRC<sup>A</sup>*) however. For those cucumber accessions with short-fruit harboring the *CsCRC<sup>G</sup>* allele, we can offer two explanations: one is that the expression level of *CsCRC<sup>G</sup>* is too low to promote fruit elongation, just as in the wild progenitor *hardwickii* (Figure 2U). Considering that 135 fruit size (FS), 30 fruit shape (FSI), and 20 fruit weight (FW) QTLs have been identified in cucumber (Pan et al. 2020), a more plausible explanation is that other genetic factors also contribute to regulating fruit size/shape in these lines. In fact, even in WI7167, multiple QTLs were detected that regulate fruit size variation (Pan et al. 2017b). Nevertheless, due to the relatively large effect of *FS5.2/CsCRC* on fruit elongation, variation at *CsCRC* can be utilized to manipulate cucumber fruit length as desired, by introgressing the *CsCRC<sup>A</sup>* allele into cultivated long cucumbers to decrease fruit length, while modulating the expression level of the *CsCRC<sup>G</sup>* allele can be used to increase fruit length in modern cucumbers.



**Figure 6** A model depicting roles of *CsCRC* alleles associated with fruit length variation in cucumber. A, Schematic diagram of the origin of the *CsCRC<sup>A</sup>* allele in cucumber. The semi-wild Xishuangbanna cucumber was domesticated from a primitive cultivated cucumber that was accompanied with a mutation and selection of the A allele, resulting in shorter fruit. In cultivated cucumber, *CsCRC<sup>G</sup>* expression is positively correlated with fruit length. B, A putative working model for *CsCRC* fine-tuning fruit elongation in cucumber. *CsCRC<sup>G</sup>* positively regulates cell expansion and fruit elongation through promoting *CsARP1* activity by direct DNA-binding at its promoter region. *CsCRC<sup>A</sup>* protein is unable to activate the transcription of *CsARP1*, resulting in reduced cell expansion and decreased fruit elongation.

### *CsARP1* is a downstream target gene of *CsCRC<sup>G</sup>*

The phytohormone auxin plays a fundamental role in plant growth and development through transcriptional regulation of auxin-responsive genes, such as those encoding auxin receptors, auxin-response factors (ARFs), INDOLE-3-ACETIC-ACID (Aux/IAA) proteins, and other downstream target genes (Krupinski and Jönsson, 2010; Stewart and Nemhauser, 2010; Enders and Strader, 2015; Lavy et al., 2016). Silencing of *SIIAA17* in tomato resulted in larger fruit with increased cell expansion (Su et al., 2014). In rice, the *OsIAA3-OsARF25* interaction positively regulates cell expansion and promotes grain length (Zhang et al., 2018). In Arabidopsis, CRC generates an auxin maxima in the medial region of the developing gynoecium via direct transcriptional repression of *TRN2* in carpel development (Staldal et al., 2008; Yamaguchi et al., 2017). ARF3 was shown to have pleiotropic effects on flower development, including the apical–basal patterning of gynoecium (Nemhauser et al., 2000), and was speculated to be the downstream target of the CRC-*TRN2*-mediated auxin pathway (Yamaguchi et al., 2017).

In this study, RNA-seq analyses in WT and *CsCRC<sup>G</sup>*-RNAi transgenic plants identified the differentially expressed gene *CsARP1*, which is the homolog of an auxin-responsive protein gene from Arabidopsis (At5g47530) that is activated during auxin-induced lateral root formation (Supplemental Data Set S2) (Neuteboom et al., 1999). We determined that *CsARP1* expression was lower in *CsCRC<sup>G</sup>*-RNAi and NIL, and higher in *CsCRC<sup>G</sup>*-OE cucumber plants (Figure 4 and Supplemental Data Set S2), suggesting that *CsCRC<sup>G</sup>*

promotes the expression of *CsARP1*. Biochemical analyses showed that indeed, *CsCRC<sup>G</sup>* directly bound to the *CsARP1* promoter to activate its transcription (Figure 4). The binding sites of *CsCRC* in cucumber included the ATCAT(ATGAT) sequence with one or two more A/T on either side, displaying the same core conserved nucleotides and certain flexibility as in Arabidopsis. *CsARP1* displayed overlapping expression regions with *CsCRC* in pulp cells of fruit mesocarp (Figures 3 and 5), indicating that, *CsCRC* directly activated *CsARP1* function in fruit mesocarp. Moreover, *CsARP1* also showed a strong signal in main vascular bundles (MVB) where *CsCRC* transcripts were absent, suggesting that *CsARP1* may participate in an additional regulatory pathway in MVB in a *CsCRC*-independent manner. Knockout of *CsARP1* by CRISPR/Cas9-mediated genome editing in cucumber led to shorter fruit and smaller cells, concomitantly with the downregulation of 16 cell wall-related genes (see Supplemental Data Set S3), suggesting a positive role for *CsARP1* in cell expansion and fruit elongation (Figure 5). Consistently, cell wall-related genes were significantly enriched among the downregulated DEGs in both *Csarp1* mutants and *CsCRC*-RNAi plants (Supplemental Figure S5). Therefore, we propose that *CsCRC<sup>G</sup>* positively regulates fruit elongation by directly promoting *CsARP1* transcription, thus enhancing cell expansion in pulp cells of fruit mesocarp in cucumber (Figure 6). Interestingly, cucumber *TRN2* and *ARF3* were not found among the DEGs in both RNA-seq results of *Csarp1* and *CsCRC*-RNAi plants, neither did we detect any difference between NIL and W17237 by RT-qPCR analysis, suggesting that CRC may act through different downstream

targets in regulating distinct biological processes and/or in different plant species.

In addition, CsARP1 contains a conserved cytochrome b561 domain and an additional DOMON domain, according to its annotation in the cucumber genome database. Previous studies have indicated that CYB561-like CYBDOM domain-containing proteins may function in transmembrane electron transport, which can adjust the redox balance of the apoplast and modulates cell wall biosynthesis and modification for cell expansion (Asard et al., 2001; Verelst and Asard, 2003; Rahman et al., 2013; Picco et al., 2015). GO term enrichment analysis showed that genes encoding iron-binding proteins and heme-binding proteins, as well as cell wall-related genes, were significantly downregulated in *Csarp1* plants (Supplemental Figure S5). Thus, CsARP1 may regulate both transmembrane electron transport and cell wall modification to contribute to fruit elongation in cucumber. Sequence alignment and phylogenetic tree analyses of CsARP1 with well-characterized b561 proteins from Arabidopsis and maize showed that CsARP1 is divergent across the conserved domains and distantly related to traditional b561s (Supplemental Figure S6). Therefore, CsARP1 appears to be a novel protein that awaits further characterization of its specific functions and underlying mechanisms in plants.

### Expression patterns and functional divergence of CRC homologs

The CRC homologs in basal angiosperms, eudicots (especially in ancestor plants and the Gramineae), monocots, and dicots displayed somewhat divergent expression patterns and functional differentiation. CRC has a single origin in angiosperms, and in the most basal angiosperm species *Amborella trichopoda*, *AmbCRC* is expressed in carpel wall and stamen filaments (Fourquin et al., 2005; Finet et al., 2016). In the basal eudicot red columbine (*Aquilegia formosa*), *AfCRC* is expressed in cells surrounding the central vascular bundle in carpel (Lee et al., 2005a). In core eudicots, such as *P. hybrida* and *N. benthamiana*, silencing of CRC homologs resulted in reduced floral meristem determinacy in both species, and loss of nectary and fusion of carpels in *N. benthamiana* (Lee et al., 2005a). These expression patterns suggest an ancestral role for CRC in carpel identity and its derived function in lateral organ development (Lee et al., 2005a; Fourquin et al., 2007). In dicots, the phenotype of Arabidopsis *crc* mutants suggests that CRC functions in carpel identity establishment, nectary formation, and floral meristem determinacy (Eshed et al., 1999). In monocots, similar to *DL* in rice, *LiCRC* in lily (*Lilium longiflorum*) is highly expressed in carpels but only weakly in leaves, and its ectopic expression rescued the drooping leaf phenotype of the rice *dl* mutant (Wang et al., 2009). In Asparagales (*Asparagus asparagoides*), *AaDL* is expressed abaxially in the ovary wall and centrally in the phloem of leaves (Nakayama et al., 2010). The expression of CRCs in leaves in monocots is consistent with its

important role in midrib formation and thus affecting plant architecture.

Here, we established that *CsCRC* expression was restricted in the mesocarp, exocarp, CVB, and nectary, but was absent in other lateral organs in cucumber (Figure 3). *CsCRC* was strongly expressed in the abaxial carpel primordia in an early stage and accumulated in the mesocarp in later developing gynoecium, suggesting that cucumber mesocarp cells may differentiate from the abaxial carpel cells. Interestingly, *CsCRC<sup>A</sup>* expression was absent in carpellary vascular bundles in the NIL, indicating that allelic variation may contribute to functional diversification by changing expression pattern in cucumber. *CsCRC<sup>C</sup>-RNAi* plants showed phenotypes of abnormal stigma occurrence and reduced nectar production (Supplemental Figure S3), suggesting that a role for CRC in carpel and nectary development is conserved in cucumber. Unlike other eudicots and monocots, *CsCRC* does not appear to be involved in floral meristem determinacy or midrib formation. Moreover, the great effect of *CsCRC* alleles on fruit length regulation in cucumber has not been reported in other agricultural crops yet.

## Materials and methods

### Plant materials and map-based cloning of *FSS.2*

For map-based cloning of the *FSS.2* candidate gene, one pair of NILs, WI7237 (fruit length 50–60 cm) and WI7327\_*fs5.2\_NIL* (fruit length 30–35 cm) were used to develop an F<sub>2</sub> population segregating only at the *FSS.2* locus. All plant materials were grown in the Walnut Street Greenhouse (WSGH) of the University of Wisconsin–Madison under natural photoperiodic conditions. Fruit size measurement was performed at the mature fruit stage, which was approximately 35 days after pollination.

Fine-mapping of *FSS.2* followed our established procedures for QTL mapping and cloning (Wang et al., 2019, 2021b). In brief, recombinants were identified from a large NIL-derived F<sub>2</sub> population using the two closest markers flanking the *FSS.2* locus previously identified. The genotypes of these recombinants at the *FSS.2* locus were determined based on segregation of fruit length and width in self-pollinated F<sub>3</sub> families in replicated field trials. At least 30 plants per F<sub>3</sub> family were examined. The genotype and phenotype data allowed the target gene to be narrowed down to a smaller region, in which new polymorphic markers were explored from resequencing reads of WI7167 and WI7237 (Pan et al., 2015). This process was repeated until *FSS.2* was delineated to a 44.6-kb region with six annotated genes (Gy14v2.0). DNA sequences between WI7167 and WI7237 within this 44.6 kb were aligned to confirm the candidate gene for QTL *FSS.2*. Genomic DNA and cDNA sequences of the *FSS.2* candidate gene were cloned from *fs5.2\_NIL* and WI7237 and sequenced by Sanger sequencing. The GenBank accession numbers for the genomic DNA sequences for the WI7237 and WI7167 alleles are MW413374

and MW413375, respectively. DNA extraction, PCR amplification of molecular markers and gel electrophoresis were conducted as described previously (Pan et al., 2015). Information on all markers used in map-based cloning is provided in Supplemental Table S1.

Cucumber inbred line R1461, *fs5.2\_NIL*, and XTMC were used for genetic transformation. Cucumber seeds were germinated at 28°C in the dark overnight and grown in a growth chamber (350  $\mu\text{mol photons/m}^2/\text{s}$ , GXZ-500D-LED, JIANGNAN Instrument; 25°C/18°C) under a 16-h light/8-h dark cycle. After the first true leaf was fully opened, the seedlings were transferred to the greenhouse at China Agricultural University in Beijing under standard growth conditions.

### Allelic diversity of *FS5.2/CsCRC* locus in a natural cucumber population

The short fruit of WI7167 and *fs5.2\_NIL* was due to a G-to-A SNP within *CsCRC*. Allelic diversity at this SNP was examined among 165 cucumber accessions consisting of 13 wild (*C. s. var. hardwickii*, CSH), 131 cultivated (*C. s. var. sativus*, CSS), and 21 semi-wild (*C. s. var. xishuangbannensis*, XIS) cucumbers. Whole-genome resequencing reads of 113 lines were obtained from the National Center for Biotechnology Information (NCBI) database (Qi et al., 2013); 52 lines were resequenced (>10 $\times$  coverage) with Illumina platforms. The genotypes of the 165 accessions at the causal SNP were inferred from resequencing data through alignment using GS Reference Mapper (V2.8) software (<https://wikis.utexas.edu/display/bioiteam/GS+Reference+mapper>), using the genomic DNA sequence of *CsCRC* cloned from WI7237 as the reference.

### Gene cloning and phylogenetic analysis

Total RNA was extracted from young fruits of R1461 and WI7237\_ *fs5.2\_NIL* using RNA extraction protocol (Liu et al., 2018). First-strand cDNAs were synthesized using TianScript II RT Kit (TIANGEN Biotech, Beijing, China). The full-length coding sequence of *CsCRC* was obtained from the Cucurbit Genomic Database (<http://cucurbitgenomics.org/>), and then cloned using gene-specific primers (Supplemental Data Set S4). Protein alignment was performed using ClustalW in the software package MEGA5.2, and conserved domains were predicted using the BoxShade website online ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html)). The phylogenetic tree was constructed using the neighbor-joining method with bootstrapping (1,000 replicates) in MEGA 10 (Supplemental File S1–S4). The evolutionary distance was calculated using Poisson's correction in units of amino acid substitution at each site. The primers used for gene cloning were listed in Supplemental Data Set S4.

### Quantitative Rt-PCR

Samples were collected from cucumber root, stem, leaf, female bud, male bud, female flower, male flower, nectary,

sepal, petal, pistil at anthesis day; and ovaries of different stages from R1461, XTMC, WI7237\_ *fs5.2\_NIL*, WI7237, transgenic plants (*CsCRC<sup>G</sup>-OE*, *CsCRC<sup>G</sup>-RNAi*, and *Csarp1*) and used for total RNA extraction (Waryoung, China, <http://www.huayueyang.com/>). First-strand cDNA was synthesized and used as the template for qPCR with a SYBR Premix Taq Mix (Takara) on an Applied Biosystems 7,500 real-time PCR system. The cucumber *UBIQUITIN EXTENSION PROTEIN* (*UBI*, Csa000874) was used as reference transcript. For each qPCR analysis, three biological replicates from different plants and three technique replicates were performed for each gene. The ovaries at anthesis were used for expression analyses of *CsCRC* in 15 natural cucumber lines. Statistical analysis was conducted with two-tailed Student's *t* tests (\**P* < 0.05, \*\**P* < 0.01) (Supplemental Data Set S5). Gene-specific primers are listed in Supplemental Data Set S4.

### In situ hybridization

Cucumber shoot apex, female buds, male buds, and young ovaries at early development stages were sampled from cucumber inbred line R1461, WI7237, or *fs5.2\_NIL*. All samples were fixed, sectioned, and hybridized as previously described (Liu et al., 2018). Sense and antisense probes for *CsCRC* and *CsARP1* were designed using the full-length coding sequence (Supplemental Data Set S4).

### Cucumber transformation

A specific 350-bp fragment of the *CsCRC<sup>G</sup>* coding sequence was cloned in reverse orientation into the pFGC1008 vector to generate the *CsCRC<sup>G</sup>-RNAi* construct. The full-length coding sequence of *CsCRC<sup>G</sup>* and *CsCRC<sup>A</sup>* was cloned into the pBI121 vector to generate the overexpression constructs. To generate the CRISPR/Cas9 vector for *CsARP1*, two sequences of 19-bp targeting exon 1 and 2 of *CsARP1* were assembled into the binary CRISPR–Cas9 vector pKSE402G using *BsaI* enzyme and T4 Ligase (Hu et al., 2017). The *CsCRC<sup>G</sup>-RNAi*, *CsCRC<sup>G</sup>-OX*, *CsCRC<sup>A</sup>-OX*, and *CsARP-CR* constructs were introduced into *Agrobacterium tumefaciens* via chemical transformation (Wise et al., 2006) and then transformed into cucumber as described previously (Hu et al., 2017). The ovaries at 5 days before anthesis (N1 stage) were collected from different transgenic plants for DNA extraction and expression analysis. PCR, sequencing, and immunoblotting were used to characterize transgenic lines (Wang et al., 2021c). In each line, 10~20 fruits at 10 DAA were measured. The primers used for vector construction and genotyping are listed in Supplemental Data Set S4.

### Histology

Fruit tissues at 10 DAA were fixed, embedded, sectioned to 6- $\mu\text{m}$  thickness, and dewaxed as described (Zhao et al., 2019). Images were taken under an Olympus light microscope. Cell numbers were counted in different fields (500  $\mu\text{m}^2$ ) in the fruit mesocarp under the microscope.

Four fields were observed for each sample, and six biological samples for each representative line were included.

### Transcriptome sequencing (RNA-seq)

Ovaries at the N1 stage from R1461, CsCRC-RNAi transgenic plants, XTMC, *Csarp1-1* were collected for RNA-seq. Three biological replicates from different plants were prepared for each sample. RNA-seq libraries were prepared following the manufacturer's instructions as previously described (Jiang et al., 2015). Library sequencing was performed on an Illumina NovaSeq 6,000 platform. Bioinformatics analysis was performed as previously described (Zhao et al., 2016). Transcriptome data were analyzed with the R package topGO and the JAVA software MapMan (Alexa et al., 2006). Differentially expressed genes (DEGs) were identified as significant with a *P*-value below 0.05 and an absolute fold-change of at least 2. Raw sequencing data were deposited at the NCBI under accession number PRJNA863475 and PRJNA863565.

### Yeast one-hybrid assay

The full-length CsCRC<sup>G</sup> and CsCRC<sup>A</sup> coding sequence were individually cloned into the pGADT7 vector (Clontech, CA, USA). The putative CRC-binding sites identified at the Arabidopsis DAP-seq website (<http://neomorph.salk.edu/PlantCistromeDB>) were identified in the CsARP1 promoter and the promoter fragments were cloned into the pAbAi vector (Clontech, CA, USA). The resulting constructs were transformed into the yeast Y1H Gold strain according to the manufacturer's instructions and screened for optimal AbA (Aureobasidin A) concentration on synthetic defined (SD) medium lacking leucine (SD –Leu) (Clontech, CA, USA). Primers for oligonucleotide synthesis are listed in Supplemental Data Set S4.

### Electrophoretic mobility shift assay

The full-length coding sequences of CsCRC<sup>G</sup> and CsCRC<sup>A</sup> were individually cloned into the pGEX-6P-1 vector in-frame with the sequence encoding GST. The resulting constructs were introduced into *Escherichia coli* BL21 (DE3). Labeled probes containing the p4, p6, or p9 elements from the CsARP1 promoter were synthesized and labeled with biotin. The recombinant proteins were purified using glutathione. EMSA was performed using a Light Shift Chemiluminescent EMSA kit (Thermo Fisher Scientific) following the manufacturer's instructions. The primers are listed in Supplemental Data Set S4.

### LUC/REN assay

A 2500-bp promoter fragment upstream of CsARP1 was cloned into pGreenII 0800 upstream of the firefly luciferase (*LUC*) reporter gene as reporter construct, and the full-length coding sequences of CsCRC<sup>G</sup> and CsCRC<sup>A</sup> were cloned into pGreenII 62-SK as effector constructs (Liu et al., 2018). The constructs together with pSUPER and the p19 plasmid were individually introduced into Agrobacterium strain

GV3101 (Hellens et al., 2005). Bacterial cultures were resuspended in infiltration buffer (MES buffer pH5.7), and mixed to 10 ml in a 9:1 ratio (effector: reporter). Then, 1/3 volume of p19 solution was added, mixed, and infiltrated into young *N. benthamiana* leaves. The injected sample was grounded in liquid nitrogen, then added PLB buffer and centrifuged at 4°C. The supernatant was taken and mixed with LAR II buffer in the 1.5 ml centrifuge tube to detect the LUC value in the GLO-MAX 20/20 Luminometer (Promega, USA). Then we added 40 µl stop reagent buffer into the sample and read the *Renilla reiformis luciferase* (*REN*) value to measure the LUC/REN ratio. The CsARP1<sup>pro</sup>:LUC reporter construct was co-infiltrated with the empty effector vector as negative control. Primers for vector construction are listed in Supplemental Data Set S4.

### Chromatin immunoprecipitation-quantitative PCR

ChIP-qPCR was performed as described previously (Zhao et al., 2019). 35S:CsCRC<sup>G</sup>-Myc and 35S:CsCRC<sup>A</sup>-Myc constructs were infiltrated in germinated cucumber seeds and incubated at 23°C for calli development (Wang et al., 2021c). The calli samples were fixed by 1% formaldehyde, and added 0.125 M glycine to stop the cross-linking. The sonicated chromatin from the vector control, 35S:CsCRC<sup>G</sup> and 35S:CsCRC<sup>A</sup> transgenic cucumber calli were purified with a QIA quick PCR Purification Kit (QIAGEN, Germany) and used as input in qPCR. In the immunoprecipitation procedure, the 8 µg/sample of anti-myc antibody (ab32, Abcam) was used to pull down the chromatin-protein complex, which was then captured on protein A beads (Abcam, UK). The final purified DNA served as the template of experimental group in qPCR (Wang et al., 2021c). Three technical and three biological replications were performed for each fragment of the CsARP1 promoter.

### Accession numbers

Raw sequencing data were deposited at the NCBI under accession number PRJNA863475 and PRJNA863565. The accession numbers for genes in this study: CsCRC (Csa5G606780), CsARP1 (Csa7G041870), and CsUBI (Csa000874).

## Supplemental data

**Supplemental Figure S1.** Expression analysis of the other five genes located within the mapping interval between WI7237 and NIL.

**Supplemental Figure S2.** Sequence alignment and phylogenetic tree analysis of CsCRC.

**Supplemental Figure S3.** Phenotypic characterization of CsCRC overexpression (OE) and CsCRC-RNAi (Ci) transgenic plants.

**Supplemental Figure S4.** Phenotypic characterization of CsCRC<sup>G</sup> and CsCRC<sup>A</sup> overexpression transgenic plants.

**Supplemental Figure S5.** Transcriptome analyses of CsCRC<sup>G</sup>-RNAi and *Csarp1-1* downstream genes.

**Supplemental Figure S6.** Sequence alignment and phylogenetic analysis of CsARP1.

**Supplemental Table S1.** Information on the markers used for fine-mapping of *FS5.2*.

**Supplemental Table S2.** Annotated genes in the 44.6-kb candidate region (Gy14v2.0).

**Supplemental File S1.** Multiple protein sequence alignment from CRC homologs used to generate the phylogenetic tree shown in **Supplemental Figure S2B**.

**Supplemental File S2.** Newick file format of the phylogenetic tree shown in **Supplemental Figure S2B**.

**Supplemental File S3.** Multiple protein sequence alignment from ARP homologs used to generate the phylogenetic tree shown in **Supplemental Figure S6B**.

**Supplemental File S4.** Newick file format of the phylogenetic tree shown in **Supplemental Figure S6B**.

**Supplemental Data Set S1.** Cucumber materials used in the present study for the allelic diversity analysis of the causal SNP of *CsCRC*.

**Supplemental Data Set S2.** Differentially expressed genes identify by RNA-Seq in fruits of *CsCRC*-RNAi vs WT.

**Supplemental Data Set S3.** Differentially expressed genes identify by RNA-Seq in fruits of *CsARP1*-CR vs WT.

**Supplemental Data Set S4.** List of primers used in this study.

**Supplemental Data Set S5.** Results of Student *t*-tests of quantitative data.

## Acknowledgments

This study was supported by grants from the National Natural Science Foundation of China [32025033] and [31930097], Hunan High-level Talents Gathering Project-innovative Talents [2021RC5006], 111 Project [B17043], and the Construction of Beijing Science and Technology Innovation and Service Capacity in Top Subjects [CEFF-PXM2019\_014207\_000032] to XZ. Work in YW's lab was supported by the Agriculture and Food Research Initiative competitive grant no. 2017-67013-26195 and 2020-51181-32139 of the USDA National Institute of Food and Agriculture. We thank Kristin Haider for technical help. We are also grateful to Drs. Jinsheng Lai, Lubin Tan, Feng Qin, and Xuexian Li for critical reading and valuable comments on the manuscript.

*Conflict of interest statement.* The authors declare that they have no conflicts of interest.

## References

- Alexa A, Rahnenführer J, Lengauer T** (2006) Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**(13): 1600–1607
- Asard H, Kapila J, Verelst W, Bérczi A** (2001) Higher-plant plasma membrane cytochrome b561: a protein in search of a function. *Protoplasma* **217**(1-3): 77–93
- Bo K, Ma Z, Chen J, Weng Y** (2015) Molecular mapping reveals structural rearrangements and quantitative trait loci underlying traits with local adaptation in semi-wild Xishuangbanna cucumber (*Cucumis sativus* L. var. *xishuangbannanensis* Qi et Yuan). *Theor Appl Genet* **128**(1): 25–39
- Bowman JL** (2000) The YABBY gene family and abaxial cell fate. *Curr Opin Plant Biol* **3**(1): 17–22
- Bowman JL, Smyth DR** (1999) *CRABS CLAW*, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development* **126**(11): 2387–2396
- Che G, Zhang X** (2019) Molecular basis of cucumber fruit domestication. *Curr Opin Plant Biol* **47**: 38–46
- Eckardt NA** (2010) YABBY Genes and the development and origin of seed plant leaves. *Plant Cell* **22**(7): 2103
- Enders TA, Strader LC** (2015) Auxin activity: past, present, and future. *Am J Bot* **102**(21): 180–196
- Eshed Y, Baum SF, Bowman JL** (1999) Distinct mechanisms promote polarity establishment in carpels of Arabidopsis. *Cell* **99**(2): 199–209
- Finet C, Floyd SK, Conway SJ, Zhong B, Scutt CP, Bowman JL** (2016) Evolution of the YABBY gene family in seed plants. *Evol Dev* **18**(2): 116–126
- Fourquin C, Vinauger-Douard M, Chambrier P, Berne-Dedieu A, Scutt CP** (2007) Functional conservation between *CRABS CLAW* orthologues from widely diverged angiosperms. *Ann Bot* **100**(3): 651–657
- Fourquin C, Vinauger-Douard M, Fogliani B, Dumas C, Scutt CP** (2005) Evidence that *CRABS CLAW* and *TOUSLED* have conserved their roles in carpel development since the ancestor of the extant angiosperms. *Proc Natl Acad Sci U S A* **102**(12): 4649–4654
- Gao Z, Zhang H, Cao C, Han J, Li H, Ren Z** (2020) QTL Mapping for cucumber fruit size and shape with populations from long and round fruited inbred lines. *Hortic Plant J* **6**(3): 132–144
- Hellens RP, Allan AC, Friel EN, Bolitho K, Grafton K, Templeton MD, Karunairetnam S, Gleave AP, Laing WA** (2005) Transient expression vectors for functional genomics, quantification of promoter activity and RNA silencing in plants. *Plant methods* **1**(1): 13
- Hu B, Li D, Liu X, Qi J, Gao D, Zhao S, Huang S, Sun J, Yang L** (2017) Engineering non-transgenic gynococious cucumber using an improved transformation protocol and optimized CRISPR/Cas9 system. *Mol Plant* **10**(12): 1575–1578
- Hu B, Wang W, Ou S, Tang J, Li H, Che R, Zhang Z, Chai X, Wang H, Wang Y, et al.** (2015) Variation in *NRT1.1B* contributes to nitrate-use divergence between rice subspecies. *Nat Genet* **47**(7): 834–838
- Ishikawa M, Ohmori Y, Tanaka W, Hirabayashi C, Murai K, Ogihara Y, Yamaguchi T, Hirano HY** (2009) The spatial expression patterns of *DROOPING LEAF* orthologs suggest a conserved function in grasses. *Genes Genet Sys* **84**(2): 137–146
- Jiang L, Yan S, Yang W, Li Y, Xia M, Chen Z, Wang Q, Yan L, Song X, Liu R, et al.** (2015) Transcriptomic analysis reveals the roles of microtubule-related genes and transcription factors in fruit length regulation in cucumber (*Cucumis sativus* L.). *Sci Rep* **5**(1): 8031
- Krupinski P, Jönsson H** (2010) Modeling auxin-regulated development. *Cold Spring Harb Perspect Biol* **2**(2): a001560
- Lavy M, Prigge MJ, Tao S, Shain S, Kuo A, Kirchsteiger K, Estelle M** (2016) Constitutive auxin response in *Physcomitrella* reveals complex interactions between Aux/IAA and ARF proteins. *eLife* **5**: e13325
- Lee JY, Baum SF, Alvarez J, Patel A, Chitwood DH, Bowman JL** (2005b) Activation of *CRABS CLAW* in the nectaries and carpels of Arabidopsis. *Plant Cell* **17**(1): 25–36
- Lee JY, Baum SF, Oh SH, Jiang CZ, Chen JC, Bowman JL** (2005a) Recruitment of *CRABS CLAW* to promote nectary development within the eudicot clade. *Development* **132**(22): 5021–5032
- Li H, Liang W, Yin C, Zhu L, Zhang D** (2011) Genetic interaction of *OsMADS3*, *DROOPING LEAF*, and *OsMADS13* in specifying rice floral organ identities and meristem determinacy. *Plant Physiol* **156**(1): 263–274
- Liu X, Chen J, Zhang X** (2021) Genetic regulation of shoot architecture in cucumber. *Hortic Res* **8**(1): 143
- Liu X, Ning K, Che G, Yan S, Han L, Gu R, Li Z, Weng Y, Zhang X** (2018) *CsSPL* functions as an adaptor between HD-ZIP III and



- CsWUS transcription factors regulating anther and ovule development in *Cucumis sativus* (cucumber). *Plant J* **94**(3): 535–547
- Liu S, Wang X, Wang H, Xin H, Yang X, Yan J, Li J, Tran LS, Shinozaki K, Yamaguchi-Shinozaki K, et al. (2013) Genome-wide analysis of *ZmDREB* genes and their association with natural variation in drought tolerance at seedling stage of *Zea mays* L. *Plos Genet* **9**(9): e1003790
- Mayorga-Gómez A, Nambesuan SU (2020) Temporal expression patterns of fruit-specific  $\alpha$ -EXPANSINS during cell expansion in bell pepper (*Capsicum annuum* L.). *BMC Plant Biol* **20**(1): 241
- Nagasawa N, Miyoshi M, Sano Y, Satoh H, Hirano H, Sakai H, Nagato Y (2003) SUPERWOMAN1 And DROOPING LEAF genes control floral organ identity in rice. *Development* **130**(4): 705–718
- Nakayama H, Yamaguchi T, Tsukaya H (2010) Expression patterns of *AaDL*, a CRABS CLAW ortholog in *Asparagus asparagoides* (*Asparagaceae*), demonstrate a stepwise evolution of CRC/DL subfamily of YABBY genes. *Am J Bot* **97**(4): 591–600
- Nemhauser JL, Feldman LJ, Zambryski PC (2000) Auxin and *ETTIN* in Arabidopsis gynoecium morphogenesis. *Development* **127**(18): 3877–3888
- Neuteboom LW, Ng JM, Kuyper M, Clijdesdale OR, Hooykaas PJ, van der Zaal BJ (1999) Isolation and characterization of cDNA clones corresponding with mRNAs that accumulate during auxin-induced lateral root formation. *Plant Mol Biol* **39**(2): 273–287
- O'Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR (2016) Cistrome and episcistrome features shape the regulatory DNA landscape. *Cell* **165**(5): 1280–1292
- Orshakova S, Lange M, Lange S, Wege S, Becker A (2009) The CRABS CLAW ortholog from California poppy (*Eschscholzia californica*, *Papaveraceae*), *EcCRC*, is involved in floral meristem termination, gynoecium differentiation and ovule initiation. *Plant J* **58**(4): 682–693
- Pan Y, Bo K, Cheng Z, Weng Y (2015) The loss-of-function *GLABROUS 3* mutation in cucumber is due to LTR-retrotransposon insertion in a class IV HD-ZIP transcription factor gene *CsGL3* that is epistatic over *CsGL1*. *BMC Plant Biol* **15**(1): 302
- Pan Y, Liang X, Gao M, Liu H, Meng H, Weng Y, Cheng Z (2017a) Round fruit shape in W17239 cucumber is controlled by two interacting quantitative trait loci with one putatively encoding a tomato *SUN* homolog. *Theor Appl Genet* **130**(3): 573–586
- Pan Y, Qu S, Bo K, Gao M, Haider KR, Weng Y (2017b) QTL Mapping of domestication and diversifying selection related traits in round-fruited semi-wild Xishuangbanna cucumber (*Cucumis sativus* L. var. *xishuangbannanensis*). *Theor Appl Genet* **130**(7): 1531–1548
- Pan Y, Wang Y, McGregor C, Liu S, Luan F, Gao M, Weng Y (2020) Genetic architecture of fruit size and shape variation in cucurbits: a comparative perspective. *Theor Appl Genet* **133**(1): 1–21
- Picco C, Scholz-Starke J, Festa M, Costa A, Sparla F, Trost P, Carpaneto A (2015) Direct recording of trans-plasma membrane electron currents mediated by a member of the cytochrome b561 family of soybean. *Plant Physiol* **169**(2): 986–995
- Qi J, Liu X, Shen D, Miao H, Xie B, Li X, Zeng P, Wang S, Shang Y, Gu X, et al. (2013) A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nat Genet* **45**(12): 1510–1515
- Rahman MM, Nakanishi N, Sakamoto Y, Hori H, Hase T, Park SY, Tsubaki M (2013) Roles of conserved Arg(72) and Tyr(71) in the ascorbate-specific transmembrane electron transfer catalyzed by *Zea mays* cytochrome b561. *J Biosci Bioeng* **115**(5): 497–506
- Sheng Y, Pan Y, Li Y, Yang L, Weng Y (2020) Quantitative trait loci for fruit size and flowering time-related traits under domestication and diversifying selection in cucumber (*Cucumis sativus*). *Plant Breed* **139**(1): 176–191
- Siegfried KR, Eshed Y, Baum SF, Otsuga D, Drews GN, Bowman JL (1999). Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. *Development* **126**(18), 4117–4128
- Song M, Fu W, Wang Y, Cheng F, Zhang M, Chen J, Lou Q (2022) A mutation in *CsKTN1* for the katanin p60 protein results in miniature plant in cucumber, *Cucumis sativus* L. *Veg Res* **2**(1): 1–9
- Staldal V, Sohlberg JJ, Eklund DM, Ljung K, Sundberg E (2008) Auxin can act independently of *CRC*, *LUG*, *SEU*, *SPT* and *STY1* in style development but not apical-basal patterning of the Arabidopsis gynoecium. *New Phytol* **180**(4): 798–808
- Staub JE, Robbins MD, Wehner TC (2008). Cucumber. In Prohens J. and Nuez F., eds, *Vegetables I: Asteraceae, Brassicaceae, Chenopodiaceae, and Cucurbitaceae*. Springer, New York, NY, pp. 241–282
- Stewart JL, Nemhauser JL (2010) Do trees grow on money? Auxin as the currency of the cellular economy. *Cold Spring Harb Perspect Biol* **2**(2): a001420
- Strable J, Vollbrecht E (2019) Maize YABBY genes *drooping leaf1* and *drooping leaf2* regulate floret development and floral meristem determinacy. *Development* **146**(6): dev171181
- Strable J, Wallace JG, Unger-Wallace E, Briggs S, Bradbury PJ, Buckler ES, Vollbrecht E (2017) Maize YABBY genes *drooping leaf1* and *drooping leaf2* regulate plant architecture. *Plant Cell* **29**(7): 1622–1641
- Su L, Bassa C, Audran C, Mila I, Cheniclet C, Chevalier C, Bouzayen M, Roustan JP, Chervin C (2014) The auxin *SI-IAA17* transcriptional repressor controls fruit size via the regulation of endoreduplication-related cell expansion. *Plant Cell Physiol* **55**(11): 1969–1976
- Tian J, Wang C, Xia J, Wu L, Xu G, Wu W, Li D, Qin W, Han X, Chen Q, et al. (2019) Teosinte ligule allele narrows plant architecture and enhances high-density maize yields. *Science* **365**(6454): 658–664
- Van Sandt VS, Stieperaere H, Guisez Y, Verbelen JP, Vissenberg K (2007) XET Activity is found near sites of growth and cell elongation in bryophytes and some green algae: new insights into the evolution of primary cell wall elongation. *Ann Bot* **99**(1): 39–51
- Verelst W, Asard H (2003) A phylogenetic study of cytochrome b561 proteins. *Genome Biol* **4**(6): R38
- Wang Y, Bo K, Gu X, Pan J, Li Y, Chen J, Wen C, Ren Z, Ren H, Chen X, et al. (2020) Molecularly tagged genes and quantitative trait loci in cucumber with recommendations for QTL nomenclature. *Hortic Res* **7**(1): 3
- Wang Y, Jiang B, Dymerski R, Xu X, Weng Y (2021b) Quantitative trait loci for horticulturally important traits defining the Sikkim cucumber, *Cucumis sativus* var. *sikkimensis*. *Theor Appl Genet* **134**(1): 229–247
- Wang H, Sun J, Yang F, Weng Y, Chen P, Du S, Wei A, Li Y (2021a) *CsKTN1* for a katanin p60 subunit is associated with the regulation of fruit elongation in cucumber (*Cucumis sativus* L.). *Theor Appl Genet* **134**(8): 2429–2441
- Wang Y, Tan J, Wu Z, VandenLangenberg K, Wehner TC, Wen C, Zheng X, Owens K, Thornton A, Bang HH, et al. (2019) STAYGREEN. STAY HEALTHY: a loss-of-susceptibility mutation in the *STAYGREEN* gene provides durable, broad-spectrum disease resistances for over 50 years of US cucumber production. *New Phytol* **221**(1), 415–430
- Wang A, Tang J, Li D, Chen C, Zhao X, Zhu L (2009) Isolation and functional analysis of *LiYAB1*, a YABBY family gene, from lily (*Lilium longiflorum*). *J Plant Physiol* **166**(9): 988–995
- Wang Z, Wang L, Han L, Cheng Z, Liu X, Wang S, Liu L, Chen J, Song W, Zhao J, et al. (2021c) *HECATE2* Acts with *GLABROUS3* and *tu* to boost cytokinin biosynthesis and regulate cucumber fruit wart formation. *Plant Physiol* **187**(3): 1619–1635
- Weng Y (2021). *Cucumis sativus* chromosome evolution, domestication, and genetic diversity. In I Goldman, ed, *Plant Breeding Reviews*, Ed 2, Vol 44. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp 79–111
- Weng Y, Colle M, Wang Y, Yang L, Rubinstein M, Sherman A, Ophir R, Grumet R (2015) QTL mapping in multiple populations and development stages reveals dynamic quantitative trait loci for fruit size in cucumbers of different market classes. *Theor Appl Genet* **128**(9): 1747–1763
- Wise A, Liu Z, Binns A (2006) Three methods for the introduction of foreign DNA into Agrobacterium. *Methods Mol Biol* **2006**(343): 43–53
- Wu S, Zhang B, Keyhaninejad N, Rodríguez GR, Kim HJ, Chakrabarti M, Illa-Berenguer E, Taitano NK, Gonzalo MJ, Díaz A, et al. (2018)

- A common genetic mechanism underlies morphological diversity in fruits and other plant organs. *Nat Commun* **9**(1): 4734
- Xin T, Zhang Z, Li S, Zhang S, Li Q, Zhang ZH, Huang S, Yang X** (2019) Genetic regulation of ethylene dosage for cucumber fruit elongation. *Plant Cell* **31**(5): 1063–1076
- Yamaguchi N, Huang J, Xu Y, Tanoi K, Ito T** (2017) Fine-tuning of auxin homeostasis governs the transition from floral stem cell maintenance to gynoecium formation. *Nat Commun* **8**(1): 1125
- Yamaguchi T, Nagasawa N, Kawasaki S, Matsuoka M, Nagato Y, Hirano HY** (2004) The YABBY gene *DROOPING LEAF* regulates carpel specification and midrib development in *Oryza sativa*. *Plant Cell* **16**(2): 500–509
- Zhang Z, Li J, Tang Z, Sun X, Zhang H, Yu J, Yao G, Li G, Guo H, Li J, et al.** (2018) Gnp4/LAX2, a RAWUL protein, interferes with the OsIAA3-OsARF25 interaction to regulate grain length via the auxin signaling pathway in rice. *J Exp Bot* **69**(20): 4723–4737
- Zhang Z, Wang B, Wang S, Lin T, Yang L, Zhao Z, Zhang Z, Huang S, Yang X** (2020) Genome-wide target mapping shows *histone deacetylase complex1* regulates cell proliferation in cucumber fruit. *Plant Physiol* **182**(1): 167–184
- Zhao J, Jiang L, Che G, Pan Y, Li Y, Hou Y, Zhao W, Zhong Y, Ding L, Yan S, et al.** (2019) A functional allele of *CsFUL1* regulates fruit length through repressing *CsSUP* and inhibiting auxin transport in cucumber. *Plant Cell* **31**(6): 1289–1307
- Zhao J, Li Y, Ding L, Yan S, Liu M, Jiang L, Zhao W, Wang Q, Yan L, Liu R, et al.** (2016) Phloem transcriptome signatures underpin the physiological differentiation of the pedicel, stalk and fruit of cucumber (*Cucumis sativus* L.). *Plant Cell Physiol* **57**(1): 19–34