Biosynthesis, regulation, and domestication of bitterness in cucumber

Yi Shang *et al.*

Science **346**, 1084 (2014);
DOI: 10.1126/science.1259215

$\delta = [\Gamma_j(1-\beta\cos\theta)]^{-1}$, respectively, where β denotes the dimensionless shock velocity and θ the angle between the line of sight and the direction of the jet, ignoring the cosmological $(1+z)$ factor. The apparent bolometric luminosity differs from its isotropic co-moving–frame value by the factor δ^4.
17. M. L. Lister *et al*., *Astron. J.* **138**, 1874–1892 (2009).
18. C. M. Urry, P. Padovani, M. Stickel, *Astrophys. J.* **382**, 501 (1991).
19. M. Lyutikov, M. Lister, *Astrophys. J.* **7**, 197–203 (2010).
20. G. Ghisellini, F. Tavecchio, *Mon. Not. R. Astron. Soc.* **386**, L28–L32 (2008).
21. D. Giannios, D. A. Uzdensky, M. C. Begelman, *Mon. Not. R. Astron. Soc.* **395**, L29–L33 (2009).
22. F. C. Michel, *Phys. Rev. Lett.* **23**, 247–249 (1969).
23. M. Lyutikov, *Mon. Not. R. Astron. Soc.* **396**, 1545–1552 (2009).
24. J. G. Kirk, I. Mochol, *Astrophys. J.* **729**, 104 (2011).
25. J. Aleksić *et al*., *Astropart. Phys.* **35**, 435–448 (2012).
26. A. Neronov, D. Semikov, I. Vovk, *Astron. Astrophys.* **519**, L6 (2010).
27. J. Aleksić *et al*., *Astrophys. J.* **723**, L207–L212 (2010).
28. J. Aleksić *et al*., *Astron. Astrophys.* **563**, A91 (2014).
29. K. Gültekin *et al*., *Astrophys. J.* **698**, 198–221 (2009).
30. D. B. McElroy, *Astrophys. J. S.* **100**, 105 (1995).
31. A. Merloni, S. Heinz, T. di Matteo, *Mon. Not. R. Astron. Soc.* **345**, 1057–1076 (2003).
32. M. Kadler *et al*., *Astron. Astrophys.* **538**, L1 (2012).
33. D. Giannios, D. A. Uzdensky, M. C. Begelman, *Mon. Not. R. Astron. Soc.* **402**, 1649–1656 (2010).
34. W. Bednarek, R. J. Protheroe, *Mon. Not. R. Astron. Soc.* **287**, L9–L13 (1997).
35. M. V. Barkov, F. A. Aharonian, V. Bosch-Ramon, *Astrophys. J.* **724**, 1517–1523 (2010).
36. M. V. Barkov, V. Bosch-Ramon, F. A. Aharonian, *Astrophys. J.* **755**, 170 (2012).
37. F. M. Rieger, K. Mannheim, *Astron. Astrophys.* **353**, 473 (2000).
38. A. Neronov, F. A. Aharonian, *Astrophys. J.* **671**, 85–96 (2007).
39. A. Y. Neronov, D. V. Semikoz, I. I. Tkachev, *New J. Phys.* **11**, 065015 (2009).
40. A. Levinson, F. Rieger, *Astrophys. J.* **730**, 123 (2011).
41. V. S. Beskin, Y. N. Istomin, V. I. Parev, *SOVAST* **36**, 642 (1992).

ACKNOWLEDGMENTS

We thank the Instituto de Astrofisica de Canarias for the excellent working conditions at the Observatorio del Roque de los Muchachos in La Palma. The support of the German BMBF and MPG, the Italian INFN, the Swiss National Fund SNF, and the Spanish MICINN is gratefully acknowledged. This work was also supported by the CPAN CSD2007-00042 and MultiDark CSD2009-00064 projects of the Spanish Consolider-Ingenio 2010 program, by grant 127740 of the Academy of Finland, by the DFG Cluster of Excellence "Origin and Structure of the Universe", by the Croatian Science Foundation (HrZZ) Projects 09/176, by the University of Rijeka Project 13.12.1.3.02, by the DFG Collaborative Research Centers SFB823/C4 and SFB876/C3, and by the Polish MNiSzW grant 745/N-HESS-MAGIC/2010/0. We thank also the support by DFG WI 1860/10-1. J. S. was supported by ERDF and the Spanish MINECO through FPA2012-39502 and JCI-2011-10019 grants. E. R. was partially supported by the Spanish MINECO projects AYA2009-13036-C02-02 and AYA2012-38491-C02-01 and by the Generalitat Valenciana project PROMETEO/2009/104, as well as by the COST MP0905 action 'Black Holes in a Violent Universe.' The European VLBI Network is a joint facility of European, Chinese, South African and other radio astronomy institutes funded by their national research councils. The research leading to these results has received funding from the European Commission Seventh Framework Programme (FP/2007-2013) under grant agreement No. 283393 (RadioNet3). The MAGIC data are archived in the data center at the Port dÍnformació Cientfica (PIC) in Barcelona. The EVN data are available at the Data Archive at the Joint Institute for VLBI in Europe (JIVE).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6213/1080/suppl/DC1
Materials and Methods
Figs. S1 to S5
Tables S1 and S2
References (*42*–*69*)

16 May 2014; accepted 23 October 2014
10.1126/science.1256183

PLANT SCIENCE

Biosynthesis, regulation, and domestication of bitterness in cucumber

Yi Shang,[1,2]* Yongshuo Ma,[1,3]* Yuan Zhou,[1,4]* Huimin Zhang,[1,3]* Lixin Duan,[5] Huiming Chen,[6] Jianguo Zeng,[4] Qian Zhou,[1] Shenhao Wang,[1] Wenjia Gu,[1,7] Min Liu,[1,3] Jinwei Ren,[8] Xingfang Gu,[1] Shengping Zhang,[1] Ye Wang,[1] Ken Yasukawa,[9] Harro J. Bouwmeester,[10] Xiaoquan Qi,[5] Zhonghua Zhang,[1] William J. Lucas,[11] Sanwen Huang[1,2]†

Cucurbitacins are triterpenoids that confer a bitter taste in cucurbits such as cucumber, melon, watermelon, squash, and pumpkin. These compounds discourage most pests on the plant and have also been shown to have antitumor properties. With genomics and biochemistry, we identified nine cucumber genes in the pathway for biosynthesis of cucurbitacin C and elucidated four catalytic steps. We discovered transcription factors *Bl* (*Bitter leaf*) and *Bt* (*Bitter fruit*) that regulate this pathway in leaves and fruits, respectively. Traces in genomic signatures indicated that selection imposed on *Bt* during domestication led to derivation of nonbitter cucurbits from their bitter ancestors.

Plant specialized metabolites play essential roles in mediating interactions between the plant and its environment and constitute a valuable resource in discovery of economically important molecules. In the plant family *Cucurbitaceae*, a group of highly oxygenated tetracyclic and bitter triterpenes, the cucurbitacins, mediated the coevolution between cucurbits and herbivores. They serve either as protectants against generalists or feeding attractants to specialists (*1*–*3*). Widely consumed as vegetables and fruits, cucurbits were domesticated from their wild ancestors that had extremely bitter fruits. Drought and temperature stress can increase the bitterness in certain domesticated cultivars, which can affect fruit quality and marketability. Molecular insights into the occurrence and domestication of bitterness in cucurbits remain largely unknown.

Despite their presence in fruits as a negative agricultural taste, cucurbitacins have for centuries been exploited for anti-inflammatory and hepatoprotective activities, in the form of traditional herbal medicines (*4*, *5*). Bitter fruits and leaves of wild cucurbit plants have been used as a purgative and emetic in India (*6*). The bitter fruit stem of melon (in Chinese, "gua di") is prescribed as a traditional hepatoprotective medicine whose effect and usage were well documented in *Ben Cao Gang Mu*, the Chinese Encyclopedia of Botany and Medicines composed by the Ming Dynasty physician Li Shi-Zhen in 1590 CE. Recent studies revealed that cucurbitacins can cause cell-cycle arrest, apoptosis, and growth suppression of cancer cells through the specific inhibition of the Janus kinase–signal transducers and activators of transcription (JAK-STAT) pathway (*7*, *8*). At present, their low concentrations in plants and nonspecific cytotoxicity limit their pharmaceutical applications.

To date, plant metabolic diversification studies (*9, 10*), as well as recently reported gene clusters in plants [reviewed in (*11*)], indicate that clustering of functionally-related genes for the biosynthesis of secondary metabolites may well be a common feature of plant genomes. In cucumber, two interacting Mendelian loci were reported to control the bitterness, conferred predominantly by cucurbitacin C (CuC) (*3, 12*). The *Bi* gene (*1*) confers bitterness to the entire plant and is genetically associated with an operon-like gene cluster (*13*), similar to the gene cluster involved in thalianol biosynthesis in *Arabidopsis* (*14*). Fruit bitterness requires both *Bi* and the dominant *Bt* (*Bitter fruit*) gene. Nonbitterness of cultivated cucumber fruit is conferred by *bt*, an allele selected during domestication as indicated by population genomics (*15*). Exploiting these genetic clues, here we report the discovery of 11 genes involved in the biosynthesis, regulation, and domestication of cucumber bitterness.

First committed step in CuC biosynthesis

To identify genetic variants associated with *Bi*, a genome-wide association study was performed

[1]Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops of the Ministry of Agriculture, Sino-Dutch Joint Laboratory of Horticultural Genomics, Beijing 100081, China. [2]Agricultural Genomic Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen 518124, China. [3]College of Life Sciences, Nanjing Agricultural University, Nanjing 210095, China. [4]Horticulture and Landscape College, Hunan Agricultural University, National Chinese Medicinal Herbs Technology Center, Changsha 410128, China. [5]Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China. [6]Hunan Vegetable Research Institute, Hunan Academy of Agricultural Sciences, Changsha 410125, China. [7]College of Life Sciences, Wuhan University, Wuhan 430072, China. [8]Institute of Microbiology, Chinese Academy of Sciences, Beijing 100190, China. [9]School of Pharmacy, Nihon University, Tokyo 101-8308, Japan. [10]Laboratory of Plant Physiology, Wageningen University, Wageningen 6700, Netherlands. [11]Department of Plant Biology, College of Biological Sciences, University of California, Davis, CA 95616, USA.
*These authors contributed equally to this work. **†Corresponding author. E-mail: huangsanwen@caas.cn**

based on the variation map (*15*) of 115 diverse cucumber lines (Fig. 1A and table S1). The most significant single-nucleotide polymorphism (SNP) was located within the region where *Bi* had been mapped and resulted in a nonsynonymous change from cysteine (C) to tyrosine (Y) at residue 393 (C393Y) of the cucumber gene *Csa6G088690* (Fig. 1B). In the 115 lines, this SNP explained the phenotype in all but one line, CG7744. In-depth analysis of the variation map identified a 1–base pair (bp) deletion at *Csa6G088690* in CG7744 that resulted in a frameshift mutation at the 760th amino acid residue (FS760) (Fig. 1B). Genetic analysis pinpointed that *Csa6G088690* defines the Mendelian *Bi* gene (fig. S1A).

Bi is a member of the oxidosqualene cyclase (OSC) gene family. Phylogenetic analysis showed that *Bi* is the ortholog of cucurbitadienol synthase gene *CPQ* in squash (*Cucurbita pepo*) (*16*) (fig. S1B). We next used yeast to express *Bi*, as well as its two mutant alleles, C393Y and FS760, to test its biochemical function. As revealed by gas chromatography–mass spectrometry (GC-MS) analysis, formation of cucurbitadienol occurred only in the yeast strain expressing the wild-type gene (Fig. 1C and fig. S1, C and D). Thus, in cucumber, *Bi* encodes a cucurbitadienol synthase that catalyzes the cyclization of 2,3-oxidosqualene into the tetracyclic cucurbitane skeleton, the first committed step of CuC biosynthesis (fig. S1E).

A leaf-specific regulator of *Bi*

To investigate the molecular mechanism in regulating CuC biosynthesis, we searched for naturally occurring mutants and screened an ethylmethane sulfonate–induced cucumber mutant library and subsequently identified two nonbitter mutants (XY-3 and E3-231). The foliage expression level of *Bi* in the natural mutant XY-3

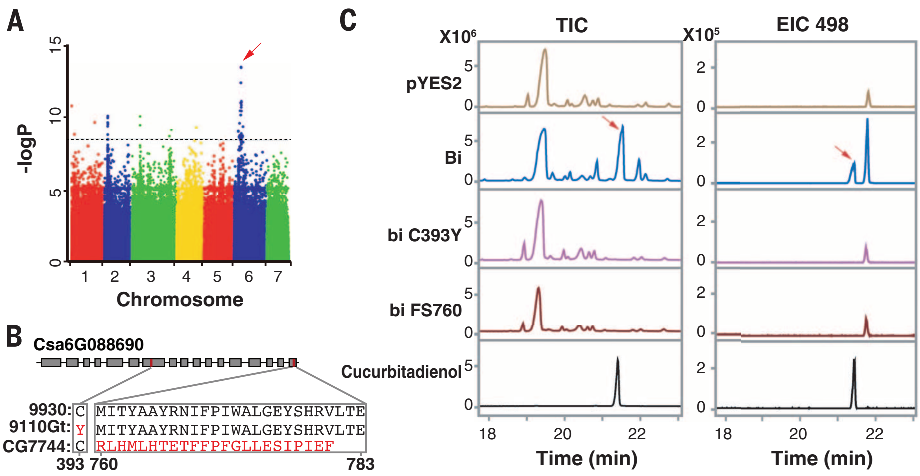

Fig. 1. The *Bi* gene. (**A**) Genome-wide association study for the bitter foliage trait. Red arrow, most significant association. Scale, –log10 of *P* value of SNPs. (**B**) Amino acid alignment between wild *Csa6G088690* and two mutant alleles. (**C**) GC-MS analysis of extracts prepared from yeast INVSc1 that harbored *Bi*, two mutant alleles (C393Y and FS760), empty vector, or an authentic standard. TIC, total ion chromatograms; EIC 498, extracted ion chromatograms at a mass/charge ratio (*m/z*) of 498 [M+TMS (trimethysilyl)].

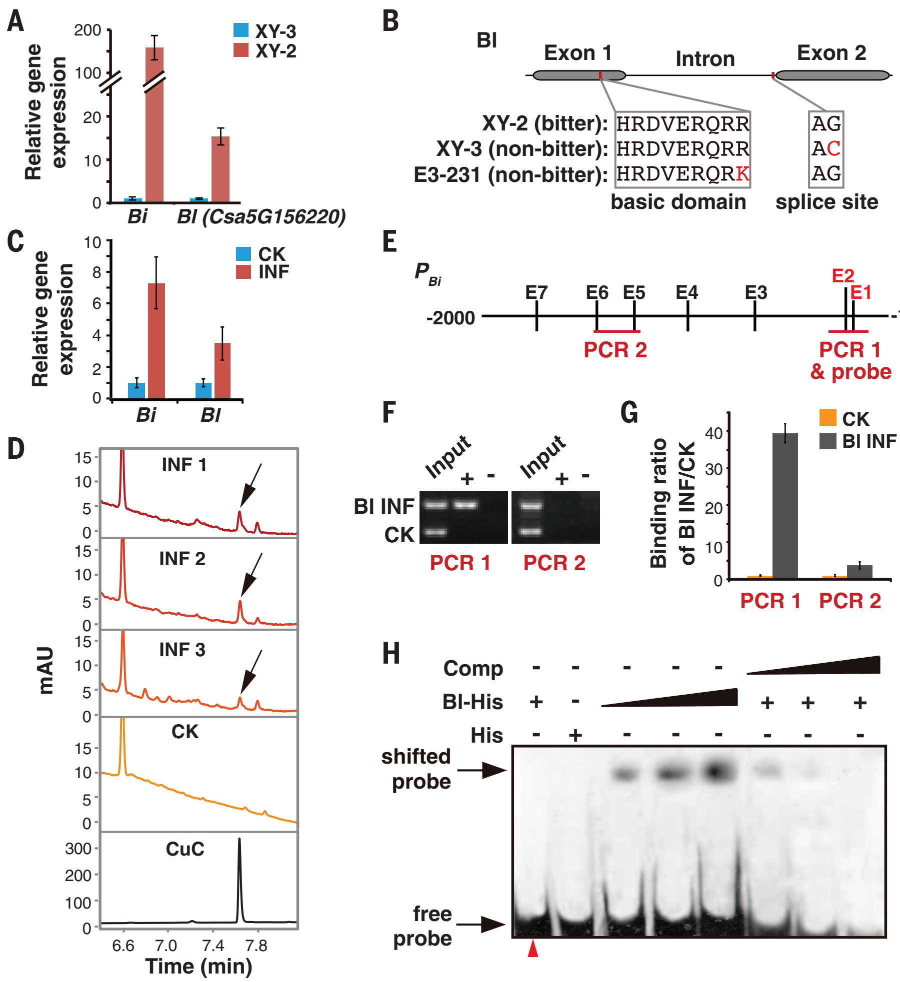

Fig. 2. The *Bl* gene. (**A**) Expression of *Bl* (*Csa5G156220*) and *Bi* in nonbitter mutant XY-3 and bitter XY-2 cucumber lines (means ± SEM, *n* = 3). (**B**) Sequence alignment between wild *Bl* and two mutated alleles. (**C** and **D**) Transient expression of *Bl* in cotyledons complemented the nonbitter phenotype of XY-3. (C) Expression of *Bl* and *Bi* determined 7 days after agroinfiltration (means ± SEM, *n* = 6). Value obtained from control (CK) was set to 1 and used to obtain relative values for the test sample. INF, sample infiltrated with *Bl*; CK, sample infiltrated with empty vector. (D) Presence or absence of CuC detected by high-performance liquid chromatography (HPLC) analysis of extracts prepared from *Bl* or control infiltrated cotyledons. mAU, milli–arbitrary units. (**E**) Schematic of the *Bi* promoter region (2000 bp upstream of the start codon). Black vertical lines indicate locations of E-box motifs, and red horizontal lines indicate regions amplified in ChIP assays or used in EMSA. Localization of mutated E-box used in EMSA is indicated in red. (**F**) ChIP analysis of Bl recruitment to the *Bi* promoter region by PCR. ChIP assays conducted with or without (+/–) Myc antibody. INF, sample infiltrated with *Bl*; CK, sample infiltrated with empty vector. (**G**) qPCR analysis of Bl recruitment to the indicated *Bi* promoter region (means ± SEM, *n* = 3). (**H**) EMSA showing that Bl-His specifically binds, in vitro, to the E-box region within the *Bi* promoter. Lane identified by a red triangle indicates that the E-box element within the probe has been mutated from CANNTG to GANNTG. Comp, competitor (unlabeled probe); His, His-tag; +/–, presence or absence of protein or competitor; closed triangle, increasing amount of protein or competitor.

was reduced to less than 1% of that in the bitter isogenic line XY-2 (Fig. 2A), which indicated that cucurbitacin biosynthesis is disrupted in XY-3. Genomes of XY-2 and XY-3 were resequenced and compared to identify possible mutations. A SNP in the cucumber gene *Csa5G156220* caught our attention, as it encodes a putative basic helix-loop-helix (bHLH) transcription factor (TF) expressed specifically in leaves (table S2). The mutation resides at the splicing site of the predicted intron that likely disrupts proper gene transcription (Fig. 2, A and B).

Resequencing of E3-231 revealed another SNP located within *Csa5G156220* that caused a change from arginine (R) to lysine (K) at the 85th amino acid residue (R85K), which is located inside the basic domain (Fig. 2B). This mutation may affect regulatory ability, as the basic domain is essential for DNA binding ability for bHLH TFs (*17*). Genetic analyses showed that the mutations in XY-3 and E3-231 are actually two recessive alleles of the same gene (fig. S2A). Increased expression of both *Bi* and *Csa5G156220* was also observed in cucumber plants either exposed to drought stress or treated with the phytohormone ABA (fig. S2, B and C), which indicated that abiotic stress may stimulate the bitterness biosynthesis in cucumber by up-regulation of *Csa5G156220*.

A cucumber cotyledon transient agro-infiltration expression system was developed to further confirm the in vivo function of *Csa5G156220* (*18*). Increasing expression of *Csa5G156220* in XY-3 cotyledons up-regulated expression of *Bi*, which in turn functionally complemented the nonbitter phenotype (Fig. 2, C and D, and fig. S3, A and B). Thus *Csa5G156220* regulates the bitterness biosynthesis in cucumber leaves, and hence, this gene was named *Bl* (*Bitter leaf*).

Next, we investigated how Bl regulates *Bi*. Analysis of the *Bi* promoter region revealed the occurrence of seven E-box (CANNTG) sequences (Fig. 2E), a cis-element to which bHLH TFs could potentially bind (*17*). Yeast one-hybrid (Y1H) assay and a tobacco transient reporter (luciferase) activation system showed that Bl indeed could bind to this promoter (fig. S2, D and E). Chromatin immunoprecipitation (ChIP) assays were performed by using formaldehyde-fixed cotyledons of XY-3 that were transiently expressing a Bl-Myc fusion protein. As revealed by the polymerase chain reaction (PCR) products and quantitative real-time PCR (qPCR), Bl was selectively recruited to the *Bi* promoter region containing E-box elements (Fig. 2, F and G). Electrophoretic mobility-shift assays (EMSAs) also confirmed selective binding of Bl to the E-box elements within the *Bi* promoter (Fig. 2H). Thus, Bl regulates cucurbitacin biosynthesis by activating transcription of *Bi* in cucumber leaves.

A cucumber domestication gene

Bt was previously mapped to a 442-kilobase (kb) region on chromosome 5 that harbors 67 predicted genes (*15*). *Bl* and its two homologs (*Csa5G157220* and *Csa5G157230*) are among these candidates and clustered in an 8.5-kb region (Fig. 3A). As *Bl* positively regulates *Bi* in cucumber leaves, we considered *Csa5G157230* to be a candidate for *Bt*, given that it is specifically expressed in the fruit of the wild line, PI 183967, consistent with the distribution of bitterness in these plants (Fig. 3A and table S2). In addition, positive correlations were observed between expression levels of *Csa5G157230* and *Bi*, and between fruit bitterness and gene expression in various cucumber lines, especially in those five extremely bitter wild lines (Fig. 3B). These studies established a correlation between *Csa5G157230* expression and accumulation of bitterness in the fruit.

Next, we performed a local association analysis within the 442-kb region to further identify genetic variants associated with the extremely bitter phenotype. This led to finding 11 signals at the regulatory region of *Csa5G157230*, including 10 SNPs and one structural variant, a 699-bp insertion 2195 bp upstream of the *Csa5G157230* start codon (SV-2195). Another variant was also identified at the regulatory region of *Csa5G157230*, a SNP at the 1601 bp upstream of the start codon (SNP-1601), which cosegregated with the *Bt* locus in a large F_2 population (n = 1822). In the 115 lines, 22 carrying a homozygous "A" at SNP-1601 all bear nonbitter fruits (table S3). These analyses indicated that selection at the regulatory region of *Csa5G157230* may down-regulate *Csa5G157230* expression in cultivated lines, which results in reduced fruit bitterness.

In some cucumber lines, fruits become bitter under stress conditions. For instance, the fruits of the cucumber line HAN become bitter when plants were grown at a low temperature (18°C

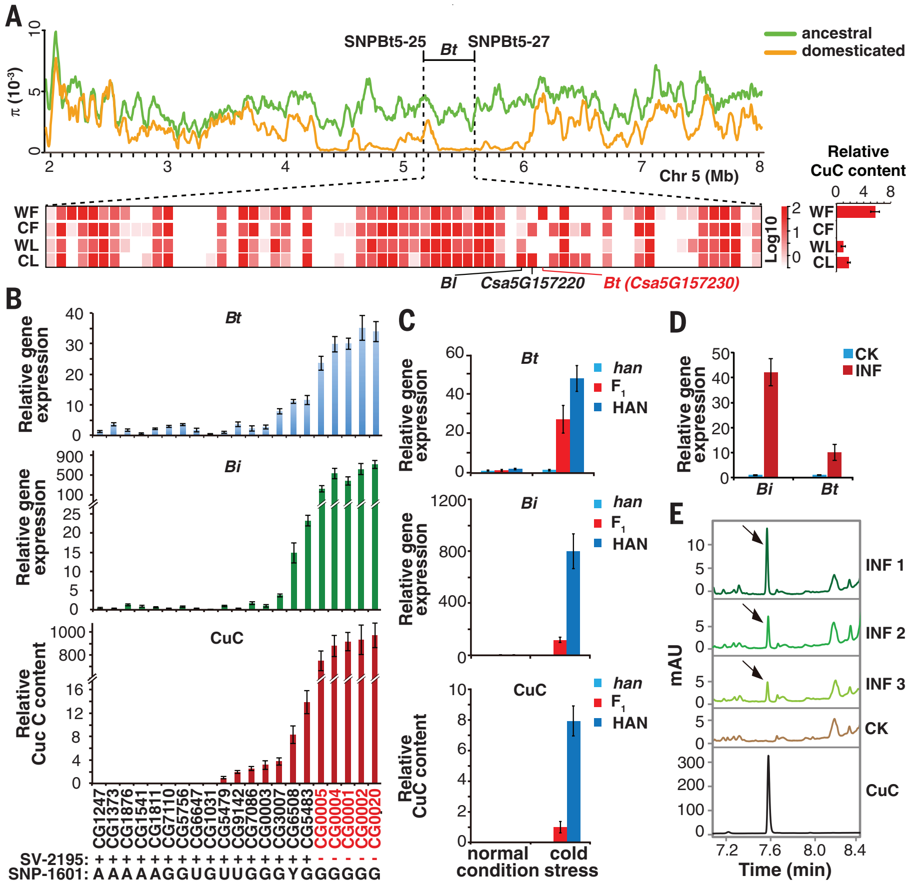

Fig. 3. The *Bt* gene. (**A**) The *Bt*-mapped region on chromosome 5 overlaps with a large domestication sweep region showing almost zero nucleotide diversity in the domesticated population (top). Differential expression profiles of genes predicted within the *Bt* region illustrated by a gradient in red (bottom). Numeric expression values of predicted genes are shown in table S2. Candidate *Bt* gene is indicated in red. CuC content of wild and cultivated cucumber was compared (means ± SEM, n = 3). WF, wild fruit; CF, cultivated fruit; WL, wild leaf; CL, cultivated leaf. (**B**) High consistency observed between expression of *Bt*, *Bi*, and the CuC content in 21 cucumber lines, including five extremely bitter lines (means ± SEM, n = 3, indicated in red). Presence or absence of SV-2195 indicated by +/−. Genotype of SNP-1601 (Y: A or G, U: unknown). (**C**) High consistency among cold-stress treatment: expression of *Bt*, *Bi*, and CuC content in fruit of HAN, *han*, and F_1 individual plants (means ± SEM, n = 3). (**D** and **E**) Transient expression of *Bt* in fruit complemented the nonbitter phenotype of cucumber line *XinTaiMiCi-2*. (D) Expression of *Bt* and *Bi* determined 15 days after agroinfiltration (means ± SEM, n = 3). Value obtained from control (CK) was set to 1 and used to obtain relative values for the test sample. INF, sample infiltrated with *Bt*; CK, sample infiltrated with empty vector. (E) Presence or absence of CuC detected by HPLC analysis of extracts prepared from *Bt* or control infiltrated fruits 15 days after agroinfiltration. mAU, milli–arbitrary units.

day, 12°C night), whereas, at a normal temperature (30°C day, 22°C night), the fruits are not bitter. We identified a natural HAN mutant (*han*), whose fruits were nonbitter even under such low temperature conditions. Resequencing both lines revealed a mutation corresponding to SNP-1601 (G in HAN and A in *han*). Genetic analysis showed that SNP-1601 cosegregates with the phenotype (fig. S4A). Our qPCR analysis indicated that SNP1601 is essential for regulating *Bi* expression in response to this environmental factor (Fig. 3C).

To confirm the in vivo function of *Csa5G157230*, a fruit transient gene expression system was developed (*18*). Expression of *Csa5G157230* activated transcription of *Bi* and promoted biosynthesis of CuC in the fruit (Fig. 3, D and E). In parallel experiments, we expressed *Csa5G157230* in XY-3 cotyledons, with the method described above. An increase in CuC content in the infiltrated XY-3 tissue was also observed (fig. S3C), which indicated that the TFs, *Bl*, and *Csa5G157230* have a similar biochemical function and that they control CuC biosynthesis in different organs. Next, we tested whether *Csa5G157230* could directly regulate the *Bi* gene. Here, we expressed the Myc-tagged protein in cotyledons of XY-3 to prepare sufficient material for ChIP assays. Similar to Bl, Csa5G157230 could bind to the E-box elements within the *Bi* promoter (fig. S4, B to F). Taken together, these studies provide strong support for the hypothesis that *Csa5G157230* is the *Bt* gene, which activates *Bi* and regulates CuC biosynthesis in the fruit.

Nine genes in CuC biosynthetic pathway

To catalyze the formation of CuC, cucurbitadienol has to be further modified with a series of oxidation reactions and acetylation, likely catalyzed by cytochrome P-450 enzymes (P450s) and an acyltransferase (ACT). On chromosome 6, *Bi* colocalizes with four P450 genes (*Csa6G088160*, *Csa6G088170*, *Csa6G088180*, and *Csa6G088710*) and one ACT gene (*Csa6G088700*) within a 35-kb genomic region. Except for *Csa6G088180*, all other genes shared nearly identical expression patterns, with high expression occurring in leaves of line 9930 as well as in fruits of wild line PI 183967 (Fig. 4, A and B). In addition, these coexpressed genes were down-regulated in leaves of XY-3 as compared with XY-2 and in fruits of *han* as compared to HAN, and they were up-regulated in cucumber leaves under ABA treatment or drought stress (Fig. 4, C to F, and table S4). Furthermore, our studies showed that Bl and Bt could specifically bind to the promoters of these coexpressed genes and could activate their transcription (Fig. 4G, and figs. S5 and S6). Mutation (R85K) within the basic domain of Bl appeared to affect its binding ability to the CuC biosynthetic genes (fig. S5, C and D), which in turn is likely to result in the nonbitter phenotype of cucumber (E3-231). Although the Y1H assay showed that Bt could also interact with the promoter of *Csa6G088180* (fig. S5A), Bt cannot activate *Csa6G088180*'s transcription (figs.S5B and S6C).

We failed in a search for the specific P450 within the cluster responsible for oxidizing cucurbitadienol, which suggests there should be other candidates located outside this 35-kb genomic region. We reasoned that other genes would be coexpressed with the *Bi* cluster and coregulated by *Bl* and *Bt*. Therefore, by applying the integrative bioinformatics and molecular biology approach described above, we identify four additional P450 genes (three on chromosome 3, *Csa3G698490*, *Csa3G903540*, and *Csa3G903550*, and one on chromosome 1, *Csa1G044890*) that are coexpressed with the *Bi* cluster and are activated by *Bl* and *Bt* in leaves and fruits, respectively (Fig. 4, A to F, and table S4).

The relation of CuC biosynthesis and these candidate tailoring enzymes was probed by using a transient RNA interference (RNAi) system acting on the bitter cotyledon of the cucumber line 9930 (*18*). RNAi-mediated target-specific down-regulation of transcripts for all these candidate genes resulted in a decrease in CuC content in the infiltrated cotyledons (Fig. 4H and fig. S7). Thus, Bl and Bt regulate bitterness formation in leaves and fruits,

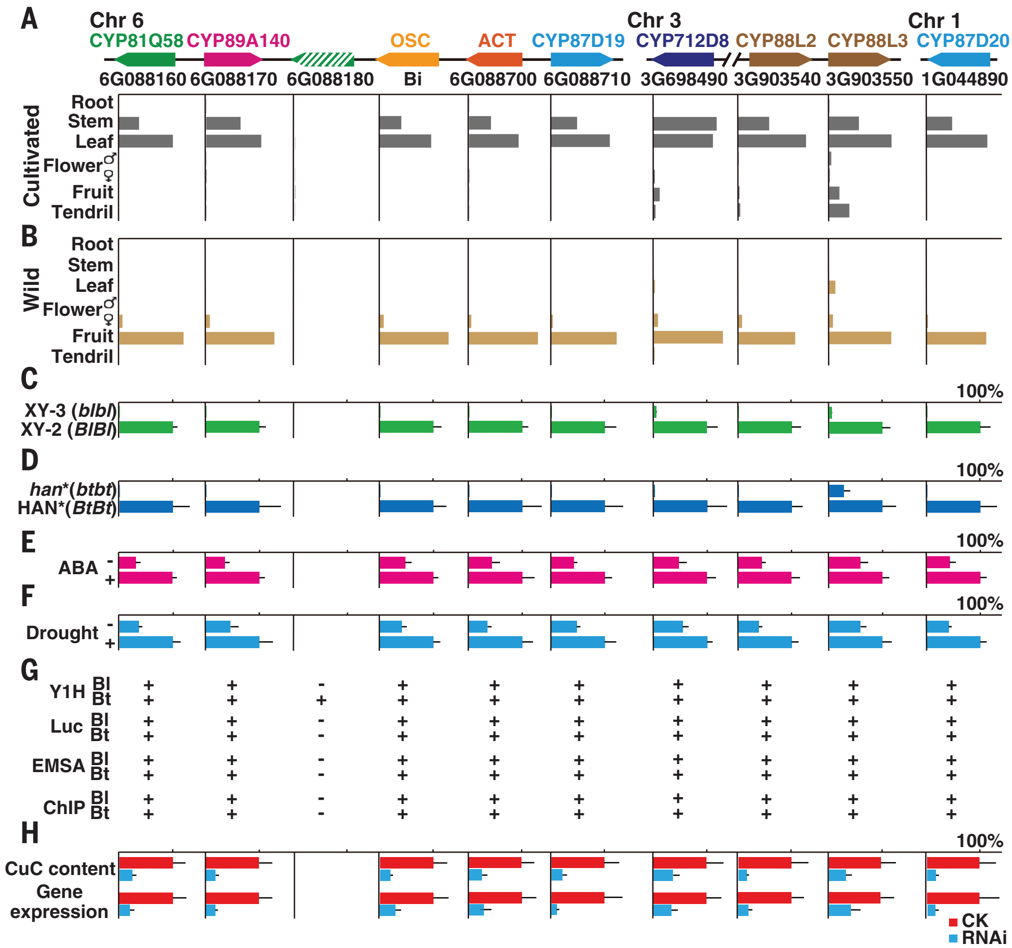

Fig. 4. Nine pathway genes that are coordinately regulated. (**A** and **B**) Identification of coexpressed candidate enzymes by analyzing transcriptomic data acquired from cultivar 9930 (A) and wild line PI 183967 (B). Candidate enzymes are indicated with different colors according to their annotations. Low-expressed gene *Csa6G088180* is indicated with hatched green and was used as a negative control in the following analyses. (**C** to **F**) Coregulation of candidate genes (means ± SEM, *n* = 3). Down-regulation of the nine genes in XY-3 as compared with XY-2 (C) and *han* as compared with HAN (D) (asterisk indicates samples prepared from plants grown under low temperature), and up-regulation of the nine genes in the presence of ABA treatment (E) or drought stress (F). (**G**) Summary of interaction between candidate gene promoter and Bl or Bt. Luc, luciferase trans-activation assay. (**H**) Function of enzymes elucidated by transient RNAi assays (means ± SEM, *n* = 6). RNAi sample in blue; CK in red. Value obtained from control (CK) was set to 1 and used to obtain relative values for the RNAi sample. CK, sample infiltrated with empty vector. More information is provided in figs. S5 to S7.

respectively, by direct transactivation of nine genes (one OSC, seven P450s, and one ACT) involved in the CuC biosynthetic pathway.

Three more steps in CuC biosynthesis

To characterize the biochemical function of these candidate P450s, we expressed each P450 in the engineered yeast (EY10) that accumulates 10 times as much cucurbitadienol as its original strain (*18*) (fig. S8). No expected product was detected from yeast extract at first (Fig. 5A). However, once an NADPH–cytochrome P450 oxidoreductase gene (*CPR*, *Csa1G423150*) was expressed with candidate P450 in the EY10, we detected a specific product catalyzed by *Csa3G903540* (a member of CYP88 family, located outside the *Bi* cluster) (Fig. 5A). The structure of this purified product (compound 1) was interrogated by nuclear magnetic resonance (NMR) spectroscopy (figs. S9 and S10), which indicated that it was a derivative of cucurbitadienol in which the 19-CH_3 was hydroxylated. The product of Csa3G903540 was named 19-hydroxy cucurbitadienol.

We continued to search for downstream P450s using this same approach. As revealed by liquid chromatography–mass spectrometry (LC-MS) assays, we identified an expected peak in the yeast expressing *Bi*, *CPR*, *Csa3G903540*, and *Csa6G088160* (a member of CYP81 family, located within the *Bi* cluster) (Fig. 5B). Tandem mass spectrometry (MS/MS) and NMR analysis revealed that a hydroxyl group was transferred to the C-25 position of 19-hydroxy cucurbitadienol and that the double bond between C-24,25 was shifted to the position of C-23,24 (figs. S11 and S12). The product (compound 2) of Csa6G088160 was named 19,25-dihydroxy cucurbitadienol.

From fresh bitter cucumber leaves, our NMR analysis identified a deacetyl CuC (figs. S13 and S14, compound 3). LC-MS analysis showed that the ACT enzyme (Csa6G088700) was able to acetylate this compound to yield CuC (Fig. 5C). These studies indicate that Csa6G088700 is the enzyme involved in the final step in the CuC biosynthetic pathway.

In summary, we discovered that two TFs regulate nine genes in the CuC biosynthetic pathway and propose a model as to how extremely bitter wild cucumber was domesticated into nonbitter cultivars (fig. S15). As revealed in this study, such regulators must contribute to the highly coordinated and efficient transcription of plant specialized metabolic pathways. The new knowledge on cucurbitacin biosynthesis will open a door for biological manufacturing and engineering of these triterpenoids as antitumor drugs, for example, in a manner similar to the biosynthesis of artemisinic acid, the antimalarial drug precursor (*19*, *20*).

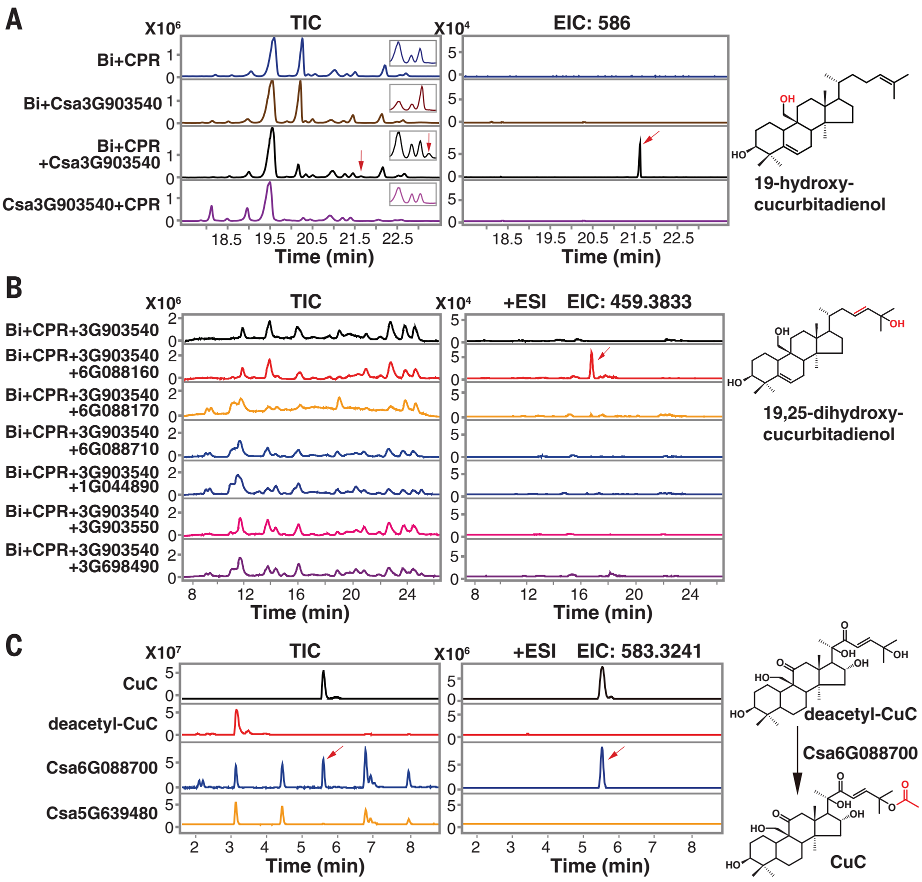

Fig. 5. Three more catalytic steps. (**A**) GC-MS analysis of putative product (red arrow) generated by Csa3G903540 in the engineered yeast (EY10). Partial enlarged details are shown as insets. TIC, total ion chromatograms; EIC 586, extracted ion chromatograms at *m/z* of 586. The product structure (right) was elucidated by NMR (*18*). (**B**) Ultra-performance liquid chromatography coupled with quadropole time-of-flight mass spectrometry (UPLC-qTOF-MS) analysis of yeast extracts with electrospray ionization (ESI) on positive mode. EIC 459.3833, extracted ion chromatograms of the accurate parent ion at *m/z* of 459.3833. The product (indicated by red arrow) structure was elucidated by MS/MS and NMR (*18*). (**C**) UPLC-qTOF-MS analysis of the acetyltransferase catalytic reaction product. Deacetyl-CuC is acetylated by Csa6G088700-His in vitro (indicated by red arrow). A leaf-specific ACT (Csa5G639480) served as a negative control. Schematic of this biosynthetic pathway from deacetyl-CuC to CuC is shown at right.

REFERENCES AND NOTES

1. C. P. Da Costa, C. M. Jones, *Science* **172**, 1145–1146 (1971).
2. R. L. Metcalf, R. A. Metcalf, A. M. Rhodes, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 3769–3772 (1980).
3. A. G. Balkema-Boomstra *et al.*, *J. Chem. Ecol.* **29**, 225–235 (2003).
4. X. Chen *et al.*, *Anticancer Drugs* **23**, 777–787 (2012).
5. J. C. Chen, M. H. Chiu, R. L. Nie, G. A. Cordell, S. X. Qiu, *Nat. Prod. Rep.* **22**, 386–399 (2005).
6. N. K. Dwivedi, O. P. Dhariwal, S. Gopala Krishnan, D. C. Bhandari, *Genet. Resour. Crop Evol.* **57**, 443–452 (2010).
7. M. A. Blaskovich *et al.*, *Cancer Res.* **63**, 1270–1279 (2003).
8. N. H. Thoennissen *et al.*, *Cancer Res.* **69**, 5876–5884 (2009).
9. L. Chae, T. Kim, R. Nilo-Poyanco, S. Y. Rhee, *Science* **344**, 510–513 (2014).
10. V. De Luca, V. Salim, S. M. Atsumi, F. Yu, *Science* **336**, 1658–1661 (2012).
11. H. W. Nützmann, A. Osbourn, *Curr. Opin. Biotechnol.* **26**, 91–99 (2014).
12. H. Horie *et al.*, *Jpn. Agric. Res. Q.* **41**, 65–68 (2007).
13. S. Huang *et al.*, *Nat. Genet.* **41**, 1275–1281 (2009).
14. B. Field, A. E. Osbourn, *Science* **320**, 543–547 (2008).
15. J. Qi *et al.*, *Nat. Genet.* **45**, 1510–1515 (2013).
16. M. Shibuya, S. Adachi, Y. Ebizuka, *Tetrahedron* **60**, 6995–7003 (2004).
17. G. Toledo-Ortiz, E. Huq, P. H. Quail, *Plant Cell* **15**, 1749–1770 (2003).
18. Materials and methods are available as supplementary material on *Science* Online.
19. D. K. Ro *et al.*, *Nature* **440**, 940–943 (2006).
20. C. J. Paddon *et al.*, *Nature* **496**, 528–532 (2013).

ACKNOWLEDGMENTS

We thank J. Bohlmann and D.-K. Ro for critical comments on the manuscript and J.-J. Qi, X.-Z. Lin, T. Lin, X.-F. Xue, and X.-Y. Liu for bioinformatic and experimental assistance. The P450s were named according to the alignment made by D. Nelson (http://drnelson.uthsc.edu/cytochromeP450.html). This work was funded by the National Program on Key Basic Research Projects in China (the 973 Program; 2012CB113900), National Science Fund for Distinguished Young Scholars (31225025), National Natural Science Foundation of China (31272161, 31322047, and 31101550), Agricultural Science and Technology Innovation Program, and National Key Technology R&D Program (the 863 Program; 2012BAI29B04). This work was also supported by the Shenzhen Municipal and Dapeng District Governments. The Institute of Flowers and Vegetables has three pending patent applications relating the genes reported in this study. Supplementary materials contain additional data. This whole-genome shotgun project has been deposited at DNA Data Bank of Japan/European Molecular Biology Laboratory/GenBank under the accession ACHR00000000. The version described in this paper is version ACHR02000000. Genes reported in the study are deposited in the National Center for Biotechnology Information (NCBI), NIH, with accession numbers (KM655851–KM655862, KM677686–KM677688). RNA-seq data may be obtained from NCBI with the accession number SRA046916.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/346/6213/1084/suppl/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 and S8
References (*21*–*24*)

25 July 2014; accepted 3 November 2014
10.1126/science.1259215