

# ***SELF PRUNING 3C* is a flowering repressor that modulates seed germination, root architecture and drought responses**

Julienne dos Reis Moreira<sup>1</sup>; Alejandra Quiñones<sup>1</sup>; Bruno Silvestre Lira<sup>2</sup>; Jessenia M. Robledo<sup>1</sup>; Shaun J. Curtin<sup>3,4,5,6</sup>; Mateus H. Vicente<sup>7</sup>; Dimas M. Ribeiro<sup>1</sup>; Malgorzata Ryngajllo<sup>8</sup>; José M. Jiménez-Gómez<sup>8,9</sup>; Lázaro Eustáquio Pereira Peres<sup>7</sup>; Maria Magdalena Rossi<sup>2</sup>; Agustín Zsögön<sup>1,‡</sup>

<sup>1</sup>Departamento de Biologia Vegetal, Universidade Federal de Viçosa, Viçosa, Brazil.

<sup>2</sup>Departamento de Botânica, Universidade de São Paulo, São Paulo, Brazil.

<sup>3</sup>United States Department of Agriculture, Plant Science Research Unit, St Paul, MN 55108, USA

<sup>4</sup>Department of Agronomy and Plant Genetics, University of Minnesota, St. Paul, MN 55108, USA

<sup>5</sup>Center for Plant Precision Genomics, University of Minnesota, St. Paul, MN 55108, USA

<sup>6</sup>Center for Genome Engineering, University of Minnesota, St. Paul, MN 55108, USA

<sup>7</sup>Departamento de Ciências Biológicas, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, Brazil.

<sup>8</sup>Max Planck Institute for Plant Breeding Research, Cologne, Germany

<sup>9</sup>Centro de Biotecnología y Genómica de Plantas (CBGP), Madrid, Spain

‡Corresponding author: [agustin.zsogon@ufv.br](mailto:agustin.zsogon@ufv.br)

JdRM: [julienemoreira14@gmail.com](mailto:julienemoreira14@gmail.com)

AQ: [aquinon3@ncsu.edu](mailto:aquinon3@ncsu.edu)

BSL: [bslira.bsl@gmail.com](mailto:bslira.bsl@gmail.com)

JMR: [kjmoncaleano@gmail.com](mailto:kjmoncaleano@gmail.com)

SJC: [shaun.curtin@usda.gov](mailto:shaun.curtin@usda.gov)

MHV: [mhvicente21@gmail.com](mailto:mhvicente21@gmail.com)

DMR: [dimas.ribeiro@ufv.br](mailto:dimas.ribeiro@ufv.br)

MR: [malgorzata.ryngajllo@p.lodz.pl](mailto:malgorzata.ryngajllo@p.lodz.pl)

JMJG: [jose.jimenez.gomez@csic.es](mailto:jose.jimenez.gomez@csic.es)

LEPP: [lazaro.peres@usp.br](mailto:lazaro.peres@usp.br)

MMR: [mmrossi@usp.br](mailto:mmrossi@usp.br)

AZ: [agustin.zsogon@ufv.br](mailto:agustin.zsogon@ufv.br)

## Highlight

The *SELF PRUNING 3C* (*SP3C*) gene is a repressor of flowering in tomato and exhibits additional functions, acting as a repressor of seed germination and modulating root architecture.

## Abstract

Allelic variation in the *CETS* (*CENTRORADIALIS*, *TERMINAL FLOWER 1*, *SELF PRUNING*) gene family has been shown to control agronomically important traits in many crops. *CETS* genes encode phosphatidylethanolamine binding proteins (PEBPs) that have a central role in flowering time control as florigenic and anti-florigenic signals. The great expansion of *CETS* genes in many species suggests that the functions of this family go beyond flowering. Here, we characterize the tomato *SELF PRUNING 3C* (*SP3C*) gene, and show that besides acting as a flowering repressor it also regulates seed germination and modulates root architecture. We show that loss of *SP3C* function in CRISPR/Cas9-generated mutant lines accelerates seed germination and increases root length with lower root side branching. Higher *SP3C* expression in transgenic lines promotes the opposite effects and also improves tolerance to water stress in seedlings. These discoveries provide insights into the role of *SP* paralogs in agronomically relevant traits and support future exploration of the involvement of *CETS* genes in abiotic stress responses.

## Keywords

Breeding, *CETS* genes, domestication, flowering repressor, germination, root architecture, *Solanum lycopersicum*, *SP3C*.

## Introduction

Domestication followed by breeding practices allowed the selection and improvement of traits of agricultural interest in most crops. A common set of traits were the target of human selection, including fruit size, colour and shape, along with changes in shoot architecture, alterations in flowering time, reduction of seed dispersal and dormancy and modifications in roots (Bai and Lindhout, 2007; Meyer and Purugganan, 2013). The ‘domestication syndrome’ comprises changes in these traits and underpins the adaptation of crops to agricultural settings, generally in monoculture, which demands large-scale production and cultivation under different environmental conditions (Denham *et al.*, 2020). A deeper understanding of the genetic control of domestication traits could allow the targeted manipulation of genomes using gene editing tools to rapidly create tailored agronomic traits (Zsögön *et al.*, 2017).

Some of the genes that control domestication and breeding-targeted identified in wheat, sugar beet, sunflower, barley, rice, potato and tomato include allelic variation in the coding or regulatory sequences in members of the *CETS* (*CENTRORADIALIS* in *Antirrhinum majus*; *TERMINAL FLOWER 1* in *Arabidopsis thaliana*; *SELF PRUNING* in *Solanum lycopersicum*) gene family (Bonnin *et al.*, 2008; Pin *et al.*, 2010; Blackman *et al.*, 2010; Comadran *et al.*, 2012; Ogiso-Tanaka *et al.*, 2013; Kloosterman *et al.*, 2013; Soyk *et al.*, 2017). *CETS* genes encode Phosphatidylethanolamine Binding Proteins (PEBPs) and play a central role in flowering control (Wickland and Hanzawa, 2015; Susila *et al.*, 2021). They are grouped into three clades, classified according to the sequence similarity with the founding members in *Arabidopsis*: *FLOWERING LOCUS T* (*FT*)-like flowering inducers, *TERMINAL FLOWER 1* (*TFL1*)-like flowering repressors, and *MOTHER OF FT AND TFL1* (*MFT*)-like, an ancestral clade without clear effects on flowering, but with other demonstrated effects on seed germination (Chardon and Damerval, 2005; Hedman *et al.*, 2009; Yu *et al.*, 2019).

The expansion of the *CETS* gene family in tomato, *Solanum lycopersicum*, resulted in 12 paralogues distributed into seven chromosomes, with five grouped as *FT*-like, five as *TFL1*-like and two as *MFT*-like (Lifschitz *et al.*, 2014; Cao *et al.*, 2016). A mutation in the coding region of the *SP* gene determines a conversion from indeterminate to determinate shoot growth via changes in auxin signaling (Silva *et al.*, 2018), with a greater number of inflorescences and highly synchronized fruit ripening. The *sp* mutation revolutionized tomato cultivation and is extensively used in breeding programs for the selection of industrial tomato cultivars, resulting in compact and productive plants that allow high density field cultivation and mechanized harvesting (Park *et al.*, 2014). Another *CETS* paralogue, *SELF-PRUNING 5G* (*SP5G*) acts as

a flowering repressor under long-day conditions, and a spontaneous mutation in the *cis*-regulatory region was selected to produce photoperiod neutral tomato plants (Soyk *et al.*, 2017; Zhang *et al.*, 2018). Changes in the balance of anti-florigenic *SP* and florigenic *SINGLE FLOWER TRUSS - SFT* alleles generate heterosis for yield (Krieger *et al.*, 2010) by altering the vegetative to reproductive growth balance (Vicente *et al.*, 2015). The manipulation of the *SP/SFT* balance using targeted gene editing tools allows the tailoring of shoot architecture to optimize plant growth and yield (Jiang *et al.*, 2013; Rodríguez-Leal *et al.*, 2017).

The manipulation of *CETS* genes in tomato and other crops has allowed improvement of elite varieties (McGarry and Ayre, 2021) or even domestication of wild species through directed genetic editing (Zsögön *et al.*, 2018). However, other traits besides those of the canonical ‘domestication syndrome’ may have been targeted, and may include altered responses to stress conditions (Milla *et al.*, 2015). Some roles for *CETS* members in responses to abiotic stress have been demonstrated. In *Arabidopsis*, *MFT* regulates seed germination under high salinity by modulating ABA signalling, and *BROTHER OF FT (BFT)* promotes delay in flowering under high salinity, providing an additional pathway to control flowering under salt stress (Xi *et al.*, 2010; Ryu *et al.*, 2011, 2014). The tomato florigen (*SFT*) is involved in control of stomatal opening (Kinoshita *et al.*, 2011) and regulates water-use efficiency (Robledo *et al.*, 2020).

Considering the large expansion of *CETS* in many species and evidence that these genes can regulate important physiological processes under stressful environmental conditions, the study of *CETS* genes is an important tool for improvement of crops in the face of challenging conditions for agriculture in the coming years (Zsögön *et al.*, 2021). Here, we characterized *SELF PRUNING 3C (SP3C)*, a member of the tomato *CETS* family whose function in flowering and plant development remained hitherto unexplored. We used reverse genetic approaches to create overexpressing and loss-of-function alleles for *SP3C* and evaluated its role on plant development. We demonstrated that *SP3C* acts as a flowering repressor. Transgenic plants overexpressing *SP3C* exhibited delayed flowering and increased number of leaves to anthesis. Additionally, we found that loss of *SP3C* function accelerated seed germination and increased root length but with lower root side branching, while *SP3C* overexpressing promoted the opposite effects. We also showed that *SP3C* expression is upregulated during a period of progressive drought, and the high levels of *SP3C* in overexpression lines conferred greater resistance to drought in the seedling stage.

## Materials and methods

### *Growth conditions and plant material*

Tomato (*Solanum lycopersicum* L.) plants, cultivar Micro-Tom (MT) (Martí *et al.*, 2006) were grown in a greenhouse at Viçosa (642 m above sea level, 20°45`S, 42°51`W), Minas Gerais, Brazil, under semi-controlled conditions: mean temperature of 26/18°C day/night, photoperiod 12-h/13-h (winter/summer), ~800  $\mu\text{mol m}^{-2} \text{s}^{-1}$  PAR, and irrigation to field capacity once per day. Seeds were germinated in polyethylene trays containing Tropostrato® commercial substrate until development of the first true leaf. At this stage, seedlings of each genotype were transplanted to pots with capacity of 350 ml, containing substrate described above. The basic fertilizing was done with NPK 2 g L<sup>-1</sup> (10-10-10) and limestone 4 g L<sup>-1</sup>. Two overexpressing MT lines (35S::SP3C#1 and 35S::SP3C#3) and three knockout lines (*sp3c*#12, *sp3c*#21 and *sp3c*#24) were selected at the T<sub>3</sub> or subsequent generations to perform the analyses.

### *Phylogenetic inference and sequence analyses*

The *CETS* gene sequences of domesticated tomato (cv. Heinz 1706) and its putative ancestral species *S. pimpinellifolium* (LA1589) we retrieved from the Sol Genomics Network (SGN, [solgenomics.net](http://solgenomics.net)) database. We aligned the sequences with MUSCLE and then submitted the alignment to trimA1 for alignment trimming and to FastTree for tree inference. The tree for the tomato genes was constructed using FigTree ([tree.bio.ed.ac.uk/software/figtree/](http://tree.bio.ed.ac.uk/software/figtree/)) and the clades were annotated using previously described phylogenetic information from the model species *Arabidopsis thaliana* (Kobayashi *et al.*, 1999; Mimida *et al.*, 2001; Hedman *et al.*, 2009). The positions of exons and introns were retrieved from SGN and the sequence differences between species were annotated manually, the gene alignments were performed and the figure drawn with Geneious R9.

Protein sequences were retrieved from Phytozome v12.1 database by searching the families containing putative CETS (CENTRORADIALIS/TERMINAL FLOWER 1/SELF-PRUNING)-like proteins described in Cao *et al.*, 2016. Most proteins were clustered in family 92441556 and a few in family 92426874 and 92423510. All sequences from *Arabidopsis thaliana*, *Solanum lycopersicum*, *Coccomyxa subellipsoidea* C-169 and *Dunaliella salina* were retrieved (Supplementary table S1) and aligned using the M-Coffee algorithm (Wallace *et al.*, 2006). The phylogenetic tree was reconstructed from the alignment using the PhyML 3.0 package (Guindon *et al.*, 2010), using the LG substitution model, tree optimization by tree

topology and branch length, tree improvement by subtree pruning and regrafting, and branch support by SH-like analysis ([www.hiv.lanl.gov/content/sequence/PHYML/interface.html](http://www.hiv.lanl.gov/content/sequence/PHYML/interface.html)). Chlorophyta species were used as outgroup. *S. lycopersicum* proteins were named accordingly to the chromosome and bin position described in (Eshed and Zamir, 1995).

For promoter analysis, the sequences of *SP*, *SFT* and *SP3C* genes (1000 bp upstream of the start codon) from *S. lycopersicum* cv. Heinz (1706) was extracted from Sol Genomics database ([solgenomics.net/](http://solgenomics.net/), ITAG release 2.40). Putative transcription factors binding sites and their role in biological processes were identified using the PlantCARE database (Lescot *et al.*, 2002).

### *Gene expression analyses*

Tomato plants, *S. lycopersicum* cv. M82 and *S. pennellii* LA716 were grown from seeds in square pots (7 × 7 × 8 cm) filled with a standard soil (Einheitserde Special, Typ Mini Tray, Einheitserde- und Humuswerke Bebr. Patzer, Sinntal Altengronau, Germany) mixed with fertilizer (Osmocote Start, 111-11-17 2MgO + TE, Everris Start) in ratio of 70 L of soil to 100 mL of the fertilizer. Plants were kept in a growth cabinet (Conviron) at 23/20°C, with 55/65% relative humidity and 16h/8h light/dark periods, and light intensity of 130 μmol m<sup>-2</sup> s<sup>-1</sup>. At the time of sampling *S. lycopersicum* seedlings had 2 compound leaves and *S. pennellii* had 3 small compound leaves and a leaf bud.

Throughout growth, the water content (WC) of the soil was maintained at 56% by daily rewatering. Added water was calculated by weighing of the plants. Drought was introduced after two weeks of growth by halting watering in the drought-treated plants. Plants were collected at the 5<sup>th</sup>, 8<sup>th</sup> and 11<sup>th</sup> day of drought treatment, which corresponds to decreasing soil water content in 40%, 30%, 20%, respectively. At each time point, compound leaves of 3 plants per condition and genotype were harvested and flash frozen in liquid nitrogen. Total RNA was independently extracted from 40 mg of grounded leaves of each plant using RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Samples were treated on column with DNase using RNase-Free DNase Set (Qiagen, Hilden Germany). Libraries were prepared according to the Illumina TruSeq RNA protocol and sequenced on an Illumina HiSeq2500 machine (Illumina, Inc., San Diego, CA) at the Genome Center of the Max Planck Institute for Plant Breeding Research Cologne.

We obtained a total of 423.91 million 100bp single end reads (average 11,8 million reads per sample, median 12.4 million reads per sample) that were aligned to the tomato genome reference sequence v4.0 using HISAT2 v 2.1.0 (Kim *et al.*, 2019) with default parameters



except for maximum intron length, that was set to 218033. An average of 90,5% of the reads were aligned to the reference (96,5% for M82 reads and 83,9% for *S. pennellii* reads). The number of reads per transcript was counted with the feature Counts function in the Rsubread package in R (Liao *et al.*, 2019). We surveyed the homogeneity of the samples with the PoissonDistance function in the Bioconductor PoiClu package (Witten, 2011). RPKMs were calculated for each gene and sample by dividing the number of reads by the length of the gene in kilobases and by the number of millions of reads in each sample. Short reads in this project have been deposited at SRA under project number PRJNA800740.

### *Molecular cloning and plant transformation*

The coding sequence of the *SP3C* gene (528 pb) was amplified by Polymerase Chain Reaction (PCR) from cDNA extracted from leaves of *Solanum lycopersicum* cv. Micro-Tom, purified and cloned in the plasmids pCR8/GW/TOPO<sup>®</sup> (Invitrogen) using the primers listed in Supplementary Table S2. Subsequently, the fragments were combined into the overexpression vector pK7WG2D by Gateway<sup>®</sup> technology (Karimi *et al.*, 2007). The correct assembly of the construct was confirmed by PCR and Sanger sequencing and used for *Agrobacterium tumefaciens* (EHA105 strain) transformation by thermal shock.

To generate a knockout in the *SP3C* gene, a CRISPR/Cas9 vector containing two guide RNAs (gRNAs) were designed to direct the Cas9 enzyme to two coding regions in the first and fourth exon. The gRNAs 5'-CAACCTCGCGTCGAAATTGG-3' and 5'-TGGATCCCAATTACCGGCTT-3' were artificially synthesized and cloned into the pMDC32 vector (Curtis and Grossniklaus, 2003) by a Golden Gate reaction (Weber *et al.*, 2011). The plasmids were confirmed by Sanger sequencing and then used for *Agrobacterium tumefaciens* (GV3101 strain) transformation by thermal shock. Both target sites contained restriction enzyme sites, which facilitated mutant screening by PCR followed by endonuclease digest assays. We ensured that the designed gRNAs would not target other genes by searching for potential off-targets using the Cas-OFFinder algorithm against the tomato genome assembly ([solgenomics.net/](http://solgenomics.net/), ITAG release 4.0) (Bae *et al.*, 2014).

For tomato *Agrobacterium tumefaciens*-mediated transformation, the EHA105 (overexpression vector) and GV3101 (CRISPR/Cas9 vector) strains were used according to Pino *et al.*, (2010), with some modifications. Briefly, seeds were sterilized and inoculated into flasks containing MS salts (Murashige and Skoog, 1962) half-strength, vitamins B5, 30 g L<sup>-1</sup> sucrose and 4 g L<sup>-1</sup> agar. The pH of the medium was adjusted to 5.7 ± 0.05 and autoclaved for sterilization. Micro-Tom (MT) cotyledons from 8-days-old seedlings were sectioned and used

for co-cultivation with *A. tumefaciens* in Root Inducing Medium (RIM) for two days,  $25 \pm 1$  °C in the dark. The explants were then transferred to Shoot Inducing Medium (SIM) and kept in  $25 \pm 1$  °C under 16/8 h photoperiod and  $10\text{--}20 \mu\text{mol m}^{-2} \text{s}^{-1}$  irradiance for four weeks. Well-developed shoots ( $\geq 5$  cm) were transferred into flasks containing MS medium without hormones, for shoot development and rooting. In all *in vitro* stages, the medium was supplemented with  $300 \text{ mg L}^{-1}$  timentin and  $100 \text{ mg L}^{-1}$  kanamycin or  $6 \text{ mg L}^{-1}$  hygromycin for selection of overexpression and CRISPR/Cas9 events, respectively.

### *Selection of overexpressing and mutant lines*

Regenerated overexpressing and knockout  $T_0$  plants were acclimatized under controlled conditions, 16/8 h photoperiod,  $150 \mu\text{mol m}^{-2} \text{s}^{-1}$  irradiance and  $25 \pm 1$  °C temperature, using autoclaved Tropostrato® substrate and vermiculite (1:1) for four weeks. Further,  $T_0$  transgenic plants were transferred to greenhouse conditions for genotyping and seed production.

To verify the transgene insertion, genomic DNA was extracted from young leaflets as described by Dellaporta *et al.* (1983). The presence of the antibiotic resistance genes, *NEOMYCIN PHOSPHOTRANSFERASE* for the overexpression and *HYGROMYCIN PHOSPHOTRANSFERASE* for CRISPR transformants, was confirmed by PCR. The positive events for overexpression construct were further confirmed by PCR using forward primer for the 35S promoter and reverse primer for *SP3C* gene. Finally, the expression level in *SP3C* overexpressing lines was confirmed by RT-qPCR in RNA extracted from leaves of  $T_3$  homozygous lines (see below for details). To screen CRISPR/Cas9-generated mutants, the *SP3C* sequence was amplified from DNA extracted of the hygromycin-resistant plants. The PCR products were purified and digested with BsrFI restriction enzyme (cutting site 5'RCCGGY3'), which cuts over the second gRNA sequence in the wild-type genotype, generating two DNA fragments of 605 bp and 347 bp. Transformation events showing only one band would have lost the restriction site and were considered potential mutants. Putative mutants were confirmed by Sanger sequencing in  $T_3$  homozygous plants. All the primers used are listed in Supplementary Table 2.

To obtain homozygous plants, overexpressing and mutant lines were self-pollinated producing  $T_1$ ,  $T_2$ ,  $T_3$  and  $T_4$  seeds. Progeny tests were carried out on seeds derived from individual plants through foliar spray of antibiotic ( $400 \text{ mg L}^{-1}$  kanamycin for five consecutive days or  $100 \text{ mg L}^{-1}$  hygromycin for three consecutive days). The progenies ( $n=30$ ) with 100% antibiotic-resistant plants (assessed visually as absence of leaf yellowing following application) were considered homozygous.



### *Real-time quantitative PCR (RT-qPCR)*

RNA was purified using ReliaPrep™ RNA Miniprep Systems (Promega), according to the manufacturer recommendations. The cDNA synthesis was performed with SuperScript™ III First-Strand Synthesis System (Invitrogen). Real-Time quantitative PCR reactions (qRT) were performed in a thermocycler Real-Time StepOnePlus PCR (Applied Biosystems) with a final volume of 14 µl using reagent SYBR Green Master Mix (Thermo Fisher Scientific). The melting curves were analyzed for nonspecific amplifications and dimerization of primers. Absolute fluorescence data were analyzed using the software LinRegPCR (Ruijter *et al.*, 2009) to obtain the values of quantification cycle (Cq) and calculate the primer efficiency. The abundance of transcripts was normalized against the geometric mean of two reference genes, *TIP4* and *EXPRESSED* and calculated according (Quadrana *et al.*, 2013). All the primers used are listed in Supplementary Table 2.

For *SP3C* expression profile the following tissues were used from Micro-Tom (MT) plants: germinated seeds, root tips ( $\leq 2$  cm), hypocotyls, leaf primordia ( $\leq 0.5$  cm), expanding leaf 1 ( $\leq 2$  cm), expanding leaf 2 ( $\leq 3$  cm), mature leaf, flowers (petals + anthers), fruit at immature green 3 stage (IG3) and fruit at red ripe stage (RR).

### *Seed germination assay*

For germination assays, 72 seeds of each genotype, previously sterilized, were selected and sown in Petri dishes containing 25 mL of MS salts (Murashige and Skoog, 1962) half-strength, vitamins B5, 30 g L<sup>-1</sup> sucrose and 4 g L<sup>-1</sup> agar. The pH of the medium was adjusted to  $5.7 \pm 0.05$  and autoclaved for sterilization. The seeds were kept in the dark at  $25 \pm 1^\circ\text{C}$  throughout the evaluation period. For each genotype four plates with 18 seeds per plate were set and the number of seeds with visible radicle was quantified daily.

### *Shoot growth analysis*

Flowering time was measured by counting the days from sowing to anthesis and the number of leaves was obtained from the plant base to the first inflorescence. The length and diameter of the internodes, and stem diameter were measured using a mechanical pachymeter (Mitutoyo Vernie®, Japan). The leaf angle was determined using a protractor, based on the insertion of the fourth and fifth leaf. The inflorescence morphology was evaluated by counting of flowers number per inflorescence, inflorescence length and number of sepals and petals per flower. At 60 days after germination the plants were harvested and the leaves digitized using a

scanner (HP Scanjet G2410) (Hewlett-Packard, Palo Alto, California, USA). Total leaf area was calculated using the software Image-Pro Plus version 4.5 (Media Cybernetics, Silver Spring, USA). Leaves and stems dry weight were obtained after oven drying at 70 °C for 72 h. Specific leaf area (SLA) was calculated through the relationship between leaf area (LA) and leaf dry weight (LDW), as described by the equation:  $SLA \text{ (cm}^2 \text{ g}^{-1}) = LA/LDW$ . All growth measurements were evaluated in 10 plants per genotype, in MT control, overexpression and knockout lines.

### *Root growth analysis*

To monitor root development, the plants were grown in vertical PVC cylinders with 30 cm of depth. The cylinders had a hollow bottom covered with a diameter  $1.8 \times 1.6$  mm mesh that allows a quantitative analysis of the root system development and the visualization of root emission over time. The number of secondary roots was quantified daily for thirteen days. The root length, volume and dry weight were evaluated in 55-day-old plants. The root volume was estimated with a graduated cylinder and root dry weight were calculated after oven drying at 70 °C for 72 h.

### *Water stress assay*

MT plants and either the overexpressing lines (*SP3C#3*) or mutants (*sp3c#21*) were grown paired in single pots. This is a way to ensure the same water supply for both genotypes, as all of them share the same soil water potential during the stress treatment (Bolaños *et al.*, 1993). The pots were filled with equal parts of substrate, sand and stones to allow better water drainage and ensure a faster and uniform drying. Watering was suspended after development of two pairs of true leaves in all genotypes, and withheld for eight days, while daily monitoring the water loss of pots. The pots of the well-watered control treatment were irrigated every day to field capacity. For both treatments, well-watered and drought, five replicates were conducted. The seedlings were photographed daily to visually monitor the stress symptoms. The chlorophyll content was estimated throughout the stress period, using a Soil Plant Analysis Development (SPAD).

On the last day of drought, the midday seedling water potential ( $\Psi$ ) was determined from 12:00 to 14:00, using a Scholander-type pressure chamber (model 1000, PMS Instruments, Albany, NY, USA). Measurements were made in four seedlings per genotype, on drought and control treatments. The relative water content (RWC) also was determined on the eighth day, in four seedlings per genotype. First, the seedlings were weighed to obtain fresh weight (FW)

and then submerged in distilled water for 24 h. After this time, the seedlings were weighed again to obtain turgid weight (TW) and dried at 70°C for 48 h to determine dry weight (DW). The RWC was determined by the equation:  $RWC = [(FW - DW) / (TW - DW)] \times 100$ . After this treatment, the remaining seedlings were rehydrated and the percentage of recovered seedlings evaluated. Total leaf area, specific leaf area and dry weight were evaluated in all surviving seedlings.

### Statistical analysis

The experimental design was completely randomized. Data were submitted to analysis of variance (ANOVA) and the means were compared by Tukey test at 5% level of significance ( $P \leq 0.05$ ) using Sisvar<sup>®</sup> software (version 5.6).

## Results

### Genomic characterization of the CETS gene family in tomato

The phylogenetic grouping of the CETS proteins in tomato and *Arabidopsis* shows that the CETS genes are distributed into three well-defined clades: *FT*-like, *TFL1*-like and *MFT*-like, based on their sequence similarity to the founding members in *Arabidopsis*. *TFL1*-like groups the flowering repressors and *FT*-like includes the flowering inducers (Fig. 1 and Supplementary Fig. S1). Sequence comparison between CETS genes of the wild ancestor of tomato *Solanum pimpinellifolium* and domesticated tomato (*S. lycopersicum*) revealed a high structural conservation between paralogues in both species (Fig. 1). All of them have four exons and three introns, and the main difference between paralogues within each species is the variation in intron length (Fig. 1). Previous work uncovered those polymorphisms leading to gene loss of function in some members of the family underlie important domestication and improvement traits. For instance, determinate growth habit is produced by a single nucleotide change in *SP* (Pnueli *et al.*, 1998), whereas deletions in the *cis*-regulatory region of *SP5G* and the fourth exon of *SP11B.1* synergistically reduce the short-day flowering response leading to day-neutral flowering (Song *et al.*, 2020) (Fig. 1).

We previously showed that *SP3D* (*SFT*) coordinates flowering with stomatal conductance and water loss (Robledo *et al.*, 2020), thus, here we decided to explore a potential role for other members of the CETS family in plant responses to drought.

The wild tomato relative *S. pennellii* is endemic of semi-arid and arid regions of the Andes Mountains and is therefore highly drought-tolerant (Taylor, 1986). We first analysed the effect

of progressive water deficit in the soil on the transcription of *CETS* genes in leaves of *S. pennellii* and *S. lycopersicum* (Fig. 2A). We found that *SP2G*, *SP3C* responded to the drought treatment by increasing their expression in both species. Interestingly, *SP11B.1* showed a decrease in *S. pennellii* in the drought treatment, but no expression in domesticated tomato, as previously described (Song *et al.* 2020). *SP3C* showed the highest expression increase during drought, so to gain further insight into its involvement in drought responses, we analyzed the promoter region of the *SP3C* gene, and compared it with that of the two main flowering regulators, *SP* and *SP3D* (*SFT*). Some motifs were common between two or three of these genes such as anaerobic induced motifs, light response and hormonal responses. Interestingly, MYB drought-responsive and seed-specific motifs were found exclusively in the *SP3C* promoter (Fig. 2B). The responsiveness to drought and the presence of the seed and drought responsive motifs in its promoter led us to produce a more in-depth functional characterization of *SP3C*.

### *SP3C* characterization and generation of overexpressing and mutant lines

*SP3C* encodes a putative PEBP-like protein of 175 amino acids and belongs to the *TFL1*-like clade (Fig. 1), suggesting a potential role as flowering repressor. This putative role is reinforced by the presence of a repressor amino acid signature, a histidine residue homologous to the *AtTFL1* His-88 (Supplementary Fig. S1) (Hanzawa *et al.*, 2005; Hedman *et al.*, 2009). We analyzed the expression levels of *SP3C* mRNA in different organs and developmental stages of tomato cv. Micro-Tom (MT). The most prominent levels of expression were found in ‘Immature Green 3’ stage fruits (IG3) and roots (Fig. 3A). We next generated overexpressing transgenic and loss-of-function mutant lines. Two independent transformation events, *35S::SP3C#1* and *35S::SP3C#3*, hereafter named *SP3C#1* and *SP3C#3*, showed *SP3C* expression levels between 700- and 800-fold higher than the MT control in the T<sub>3</sub> homozygous generation (Fig. 3B). For the loss-of-function experiment, three independent CRISPR-mediated *SP3C* homozygous events were obtained, *sp3c#12*, *sp3c#21* and *sp3c#24*, and the mutations were confirmed by Sanger sequencing (Fig. 3C). At the first gRNA site (Target 1) on the first exon, all three mutants had a single nucleotide insertion (+G) 172 bp downstream of the start codon, that resulted in disruption of the reading frame. At the site targeted by the second gRNA (Target 2), the line *sp3c#12* had a 3-bp deletion (CCG), whereas lines *sp3c#21* and *sp3c#24* showed a 55-bp deletion (Fig. 3C). These mutant alleles resulted in a premature stop codon and a predicted protein of 74 amino acids (Supplementary Fig. S2). With these

genotypes, we conducted a broad phenotypic characterization and investigated the effects of *SP3C* in phenology, growth and morphology of tomato plants (Fig. 3D).

### *SP3C is a repressor of seed germination*

While screening the overexpression lines for homozygous insertions we noticed a reduction in seed germination rate, especially in the *SP3C* overexpression line *SP3C#1* (Supplementary Fig. S3). These observations, together with the presence of seed-specific expression motifs in the *SP3C* promoter (Fig. 2B), led us to quantify seed germination rate in a controlled assay. Homozygous *SP3C* overexpressing lines displayed a significant delay in germination compared to wild-type MT plants (Fig. 4A). After 12 days, over 96% of the MT seeds and mutants germinated; however only 21% and 78 % of *SP3C#1* and *SP3C#3*, respectively, showed radicle protrusion through the seed testa. Moreover, a slight acceleration in mutant seeds germination between the 4th and the 5th day after sowing was observed as well for the loss-of-function lines *sp3c#12* and *sp3c#21*. Our results indicate *SP3C* repress seed germination.

### *SP3C is a flowering repressor*

Most vegetative growth parameters showed no evident alterations in both transgenic *35S::SP3C* and *sp3c* mutant lines compared to the MT control (Supplementary Table S3 and Fig. S4). However, *SP3C*-overexpressing plants showed delayed flowering and a greater number of days to anthesis, whereas *sp3c* mutants had a subtle but consistent acceleration of flowering (Fig. 4B, C). The number of leaves to the first inflorescence was higher in *35S::SP3C* lines, while it remained unchanged in *sp3c* mutants (Fig. 4D). These results suggest that *SP3C*, similarly to other members of the *CETS* family, contribute to flowering control; particularly, acting as a floral repressor in tomato.

In addition, we also analysed parameters related to inflorescence morphology and observed that although the floral structure was not altered, the inflorescence morphology was modified in the overexpressing lines. The *35S::SP3C* lines had a greater number of flowers per inflorescence unit and a longer inflorescence (Supplementary Table S3 and Fig. S5).

### *SP3C reduces root length and stimulates root branching*

The presence of drought-responsive motifs in the *SP3C* promoter and the high expression level in roots hinted at a putative role of *SP3C* in root development. To better explore this, we conducted an experiment under controlled conditions to assess root growth (Fig. 5A). We

observed that one of the *35S::SP3C* overexpressing events and two *sp3c* mutants showed reduced and increased root length in relation to MT control, respectively (Fig. 5B, C). A time-course of secondary root development revealed that *SP3C#3* plants had faster root branching than MT, whereas *SP3C#1* lagged behind MT, with all three genotypes eventually reaching ~100 secondary roots 32 days after transplanting. On the other hand, secondary root development was severely compromised in the mutant lines: all of them showed a slower rate of root branching and at the end of the evaluation period *sp3c#12* had 43% ( $p = 0.1236$ ), *sp3c#21*, 58% ( $p = 0.0028$ ), and *sp3c#24*, 65% ( $p = 0.0004$ ) lower number of secondary roots than MT (Fig. 5D). Root volume and dry mass were also significantly reduced in the mutants (Supplementary Fig. S6). Thus, these data revealed that *SP3C* represses principal root growth and increases root ramification.

### *SP3C improves the drought response in seedlings*

We showed that *SP3C* expression is increased under drought conditions, which, coupled to its impact on seed germination and root growth, could indicate a possible involvement in plant responses to water stress. To further investigate the relation between *SP3C* and drought, we conducted a water stress assay in 22-day-old seedlings of two selected lines: one overexpression (*SP3C#3*) and one mutant (*sp3c#21*), for eight days. The soil water loss was monitored daily during water stress. On the eighth day, the drought treatment pots had a loss of approximately 80% in their water content compared to the start of treatment (Supplementary Fig. S7).

During the time-course of progressive soil drying, *sp3c#21* mutant seedlings showed more intense leaf wilting symptoms, while *SP3C#3* visually appeared to be less affected by drought (Fig. 6A). On the seventh day of water stress, *sp3c#21* mutant seedlings were fully wilted, considerably more than MT control plants, and *SP3C#3* seedlings began to show noticeable symptoms of leaf rolling (Fig. 6A). We ensured that all plants from the drought treatment were under stress by measuring water potential ( $\Psi$ ), relative water content (RWC) and chlorophyll content (Fig. 6B-E). The  $\Psi$  and RWC of genotypes under drought was consistently lower than control plants.

Our results suggest that the greater tolerance of *SP3C#3* line to drought was associated with higher  $\Psi$  and higher water status at the end of the drought treatment. We also showed that the tolerance of this line is exclusively related to the increase in *SP3C* expression and not to other factors that could interfere in the genotype-specific response. Thus, we ensured that there



were no discrepancies in total leaf area, specific leaf area and biomass between *SP3C#3* lines and *sp3c#21* mutants, which could impact the stress response (Supplementary Fig. S7).

## Discussion

The transition from vegetative to reproductive growth is a key determinant of plant architecture and agronomic performance, as it conditions photosynthetic source-sink relations and productivity. Flowering time has therefore been one of the main breeding targets, allowing crops to perform better under specific environmental conditions and driving the expansion of agricultural areas to a wider range of latitudes and climates (Chen *et al.*, 2020). It had long been known that a flowering stimulus is produced in the leaves and translocated to the apical meristem, where it triggers the switch from vegetative to reproductive growth (Zeevaart, 2006). This signal, the ‘florigen’, was elusive for many decades but eventually shown to be small polypeptides (<200 amino acids) encoded by the *CETS* (*CENTRORADIALIS*, *TERMINAL FLOWER 1*, *SELF PRUNING*) gene family (Susila *et al.*, 2021). *CETS* family has undergone expansion via duplication and subfunctionalization in angiosperms (Wickland and Hanzawa, 2015). In spite of the partial functional redundancy of the orthologs within each sub-clade, in some species it has been shown that each one fulfils specific roles to fine-tune plant development in response to specific environmental cues.

Tomato is a valuable model to gain further functional insight into the *CETS* gene family (Carmel-Goren *et al.*, 2003). A large-scale genomic analysis demonstrated that three of the twelve *CETS* genes (*SP3C*, *SP5G* and *SP6A*) are located within genomic regions comprising putative domestication sweeps and six others (*SP1C.1*, *SP1C.2*, *SP3D*, *SP*, *SP11B.1* and *SP11B.2*) in putative improvement sweeps (Lin *et al.*, 2014). This suggests that selection of allelic variation on genes of the *CETS* family has been a key driver of the changes operated along the transitions between *S. pimpinellifolium* and *S. lycopersicum* var. *cerasiforme* (domestication), and between *S. lycopersicum* var. *cerasiforme* and modern fresh fruit and processing varieties (improvement) of *S. lycopersicum* (Razifard *et al.*, 2020). A single amino acid change in the *SP* gene (Fig. 1) leads to determinate growth, which includes alteration in the auxin signaling (Silva *et al.*, 2018), and underpins the development of processing tomato varieties that are suitable for mechanical harvest (Pnueli *et al.*, 1998). In their natural habitats wild tomatoes respond to photoperiodic cues and show a strong short-day response to flowering (Taylor, 1986). Mutations in the *SP5G* and *SP11B.1* loci (Fig. 1) were selected to suppress the photoperiodic response and are found in almost completely day neutral tomato varieties grown

in higher latitudes (Soyk *et al.*, 2017; Zhang *et al.*, 2018; Song *et al.*, 2020). Lastly, a loss-of-function mutation in the fourth exon of *SP6A* is ubiquitous in domesticated tomatoes (Fig. 1). Its functional significance is hitherto unknown, but *SP6A*, in conjunction with *SP5G*, is the main controller of tuberization in response to day length in potatoes (Abelenda *et al.*, 2016).

A deeper understanding of the expression patterns and phenotypic effects of the genes of the *CETS* family would provide potentially useful targets for breeding using the state-of-the-art gene-editing suite of tools (Gasparini *et al.*, 2021). A loss-of-function mutation in *SFT* (*SP3D*) improves the vegetative-to-reproductive growth balance when in heterozygosis and combined with the *SP* loss-of-function allele (*sp*) in homozygosis (Krieger *et al.*, 2010). Likewise, the tomato plant architecture can be fine-tuned and optimized by targeted mutations in the promoter region of the *SP* gene (Rodríguez-Leal *et al.*, 2017). This suggests that considerable room for improvement of plant growth and agronomic performance exists by manipulation of *CETS* family genes, not only in tomato but also in other crop species (McGarry and Ayre, 2012).

Here, we first showed that the expression of the *SP2G* and *SP3C*, orthologs of *Arabidopsis MFT* and *TFL1*, is strongly induced in the leaves of the drought-resistant wild species *S. pennellii* and domesticated tomato upon suspension of watering (Fig. 2A). The increased expression pattern in proportion to the time under drought was unique to both these genes and not observed in the other paralogues of the *CETS* family. We previously showed that the *SFT* (*SP3D*) is responsible for controlling water-use efficiency in tomato by altering stomatal responses upon flowering (Robledo *et al.*, 2020). This indicates that the *CETS* family genes have pleiotropic roles that go beyond the control of plant development and may include modulation of water relations.

We focused our analyses on dissecting the function of the *SP3C* gene, and demonstrated that it pleiotropically controls seed germination, flowering time and root development. Increased *SP3C* expression in transgenic lines led to delayed in seed germination and flowering. On the other hand, using CRISPR/Cas9 to create targeted mutations in the coding region of the *SP3C* gene accelerated seed germination and subtly shortened flowering time. Seed germination time is a fundamental domestication trait, as most wild species have genetic and physiological mechanisms to induce strong dormancy (Chahtane *et al.*, 2017; Penfield, 2017). Correct timing of germination to coincide with favorable environmental conditions can result in increased survival and fitness (Wang *et al.*, 2018). However, extended seed dormancy is not a desirable trait for crops, as it can lead to erratic and unpredictable germination rates (Rodríguez *et al.*, 2015), preclude the better use of fertilizers and extend the growth cycle. *MFT*

represses seed germination under far-red light by modulating ABA and hormone responses (Xi *et al.*, 2010; Vaistij *et al.*, 2018). Recent work in *Arabidopsis* has shown that ectopic *FT* expression in seeds increased dormancy by modulating gibberellin biosynthesis (Chen *et al.*, 2021). Further studies would clarify whether the modulation of seed germination by *SP3C* could involve hormonal alterations.

The transition to flowering in angiosperms is controlled through a mechanism that integrates environmental and endogenous signals that converge on the expression of *CETS* genes (Cao *et al.*, 2021). *CETS* genes encode flowering repressors (similar to *TFL1*) and flowering inducers (similar to *FT*), which can differ between them by as little as a single amino acid (Hanzawa *et al.*, 2005), and compete for the binding site of *FLOWERING LOCUS D* (*FD*), a bZIP transcription factor with localized expression in the meristem (Collani *et al.*, 2019). Binding of *FD* to a *CETS* protein and adaptor proteins of the 14-3-3 family leads to formation of the Flowering Activator Complex (FAC) (Ho and Weigel, 2014), which drives the reprogramming of the meristem and its conversion from vegetative to floral (Périlleux *et al.*, 2019). Here, we showed that *SP3C* is a negative regulator of flowering, which is consistent with its placing in the *TFL1*-like clade of proteins that in tomato includes the anti-florigenic signals *SP* and *SP9D*. Moreover, the high expression of *SP3C* in roots and immature fruits immediately points to functions beyond the control of flowering time.

Root growth patterns were also affected by alteration of *SP3C* function. High *SP3C* expression led to reduced root length but increased root branching and secondary roots. Loss of *SP3C* function had the opposite effect of increasing root length but decreasing secondary root formation. Previous work showed that *TFL1* genes are expressed in the roots of many species, including *Arabidopsis*, where they can act as repressors of root growth, leading to reduced root length (Bouché *et al.*, 2016). The expression of *CETS* family genes in roots suggests a potential mechanism whereby the coordination of flowering time and root growth is achieved. In many species (*e.g.* sugar beet and potato) the transition to flowering triggers a halt in root growth that is fundamental to coordinate resource allocation and guarantee reproductive success (Pin *et al.*, 2010; Navarro *et al.*, 2011; Abelenda *et al.*, 2016). Further investigation of the specific roles of diverse *CETS* paralogues on this phenomenon could provide a novel target to create root systems that are tailored to specific environments based on the design of ideotypes and their creation using gene editing tools (Lynch, 2019).

The pleiotropic effects described for many members of the *CETS* family, on which we have expanded here, would make them important drivers of domestication and breeding, as relatively simple genetic changes could produce multiple beneficial traits simultaneously. The

discovery and characterization of a flowering repressor with pleiotropic effects on seed germination and root growth represents an advance in the effort to link genotypes to agronomically relevant traits in horticultural crops. Further work is required to shed more light on the effect of drought on *SP3C* expression and function. In summary, our study provides a new target for simultaneous modulation of seed germination, flowering and root growth that could be fine-tuned for agronomic gain.

## Supplementary data

The following supplementary data are available at *JXB* online.

*Table S1.* Sequences used for phylogenetic analysis.

*Table S2.* Oligonucleotides DNA sequence for PCR primers used in this study.

*Table S3.* Growth parameters measured in Micro-Tom (MT), overexpression lines and mutants.

*Fig. S1.* Characterization of the *SP3C* gene.

*Fig. S2.* Amino acid sequences of the *sp3c* mutant alleles.

*Fig. S3.* *SP3C* acts as a repressor of seed germination.

*Fig. S4.* Tomato plants harboring allelic variations in the *SP3C* gene.

*Fig. S5.* High levels of *SP3C* increase the flowers number and the inflorescence length.

*Fig. S6.* Impacts of *SP3C* on root grown and development.

*Fig. S7.* *SP3C* improves tomato seedling response to drought.

## Acknowledgements

We thank the support and contributions of the UFV Plant Physiology Graduate Program.

## Author contributions

L.E.P.P., M.R. and A.Z. designed study. J.d.R.M., A.Q., B.S.L., J.M.R., S.J.C., M.H.V. and M.Ry. performed experiments. M.Ry. and J.M.J-G. analyzed data. D.M.R., J.M.G-G., L.E.P.P., M.R. and A.Z. acquired funding and managed research personnel. J.d.R.M. and A.Z. wrote the manuscript. All authors commented, edited and approved the final manuscript.

## Conflict of interest

The authors declare no conflict of interest.

## Funding

This work was funded in part by a grant (RED-00053-16) from the Foundation for Research Assistance of the Minas Gerais State (FAPEMIG, Brazil) and by funding from the German Research Foundation under the German-Israeli Project Cooperation program (DFG DIP project number FE552/12-1 awarded to J.M.J.-G.).

Disclaimer: The views and opinions expressed in this work do not necessarily reflect the official policy or position of the authors' respective employer or government." Mention of any trade name, proprietary product or specific equipment does not constitute a guarantee or warranty by USDA-ARS and does not imply its approval to the exclusion of other products that may also be suitable. The USDA-ARS is an equal opportunity and affirmative action employer and all agency services are available without discrimination.

## Data availability

All data supporting the findings of this study are available within the paper and within its supplementary data published online. No restrictions are placed on the materials, such as material transfer agreements.

## References

- Abelenda JA, Cruz-Oró E, Franco-Zorrilla JM, Prat S.** 2016. Potato StCONSTANS-like1 suppresses storage organ formation by directly activating the FT-like StSP5G repressor. *Current Biology* **26**, 872–881.
- Bae S, Park J, Kim J-S.** 2014. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* **30**, 1473–1475.
- Bai Y, Lindhout P.** 2007. Domestication and Breeding of Tomatoes: What have We Gained and What Can We Gain in the Future? *Annals of Botany* **100**, 1085–1094.
- Blackman BK, Strasburg JL, Raduski AR, Michaels SD, Rieseberg LH.** 2010. The role of recently derived FT paralogs in sunflower domestication. *Current biology* **20**, 629–635.
- Bolaños J, Edmeades GO, Martinez L.** 1993. Eight cycles of selection for drought tolerance in lowland tropical maize. III. Responses in drought-adaptive physiological and morphological traits. *Field Crops Research* **31**, 269–286.
- Bonnin I, Rousset M, Madur D, Sourdille P, Dupuits C, Brunel D, Goldringer I.** 2008. FT genome A and D polymorphisms are associated with the variation of earliness components in hexaploid wheat. *Theoretical and Applied Genetics* **116**, 383–394.
- Bouché F, D'Aloia M, Tocquin P, Lobet G, Detry N, Périlleux C.** 2016. Integrating roots into a whole plant network of flowering time genes in *Arabidopsis thaliana*. *Scientific Reports* **6**, 29042.
- Cao K, Cui L, Zhou X, Ye L, Zou Z, Deng S.** 2016. Four tomato *FLOWERING LOCUS T*-like proteins act antagonistically to regulate floral initiation. *Frontiers in plant science* **6**, 1213.

- Cao S, Luo X, Xu D, Tian X, Song J, Xia X, Chu C, He Z.** 2021. Genetic architecture underlying light and temperature mediated flowering in *Arabidopsis*, rice, and temperate cereals. *New Phytologist* **230**, 1731–1745.
- Carmel-Goren L, Liu YS, Lifschitz E, Zamir D.** 2003. The *SELF-PRUNING* gene family in tomato. *Plant Molecular Biology* **52**, 1215–1222.
- Chahtane H, Kim W, Lopez-Molina L.** 2017. Primary seed dormancy: a temporally multilayered riddle waiting to be unlocked. *Journal of Experimental Botany* **68**, 857–869.
- Chardon F, Damerval C.** 2005. Phylogenomic analysis of the PEBP gene family in cereals. *Journal of molecular evolution* **61**, 579–590.
- Chen L, Cai Y, Qu M, et al.** 2020. Soybean adaption to high-latitude regions is associated with natural variations of *GmFT2b*, an ortholog of *FLOWERING LOCUS T*. *Plant, Cell & Environment* **43**, 934–944.
- Chen F, Li Y, Li X, Li W, Xu J, Cao H, Wang Z, Li Y, Soppe WJJ, Liu Y.** 2021. Ectopic expression of the *Arabidopsis* florigen gene *FLOWERING LOCUS T* in seeds enhances seed dormancy via the GA and DOG1 pathways. *The Plant Journal* **107**, 909–924.
- Collani S, Neumann M, Yant L, Schmid M.** 2019. FT Modulates Genome-Wide DNA-Binding of the bZIP Transcription Factor FD1. *Plant Physiology* **180**, 367–380.
- Comadran J, Kilian B, Russell J, et al.** 2012. Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nature genetics* **44**, 1388–1392.
- Curtis MD, Grossniklaus U.** 2003. A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant physiology* **133**, 462–469.
- Dellaporta SL, Wood J, Hicks JB.** 1983. A plant DNA miniprep: version II. *Plant molecular biology reporter* **1**, 19–21.
- Denham T, Barton H, Castillo C, Crowther A, Dotte-Sarout E, Florin SA, Pritchard J, Barron A, Zhang Y, Fuller DQ.** 2020. The domestication syndrome in vegetatively propagated field crops. *Annals of Botany* **125**, 581–597.
- Eshed Y, Zamir D.** 1995. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. *Genetics* **141**, 1147–1162.
- Gasparini K, Moreira J dos R, Peres LEP, Zsögön A.** 2021. De novo domestication of wild species to create crops with increased resilience and nutritional value. *Current Opinion in Plant Biology* **60**, 102006.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O.** 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology* **59**, 307–321.
- Hanzawa Y, Money T, Bradley D.** 2005. A single amino acid converts a repressor to an activator of flowering. *Proceedings of the National Academy of Sciences* **102**, 7748–7753.
- Hedman H, Källman T, Lagercrantz U.** 2009. Early evolution of the MFT-like gene family in plants. *Plant molecular biology* **70**, 359–369.
- Ho WWH, Weigel D.** 2014. Structural features determining flower-promoting activity of *Arabidopsis FLOWERING LOCUS T*. *The Plant cell* **26**, 552–64.



- Jiang K, Liberatore KL, Park SJ, Alvarez JP, Lippman ZB.** 2013. Tomato yield heterosis is triggered by a dosage sensitivity of the florigen pathway that fine-tunes shoot architecture. *PLoS Genetics* **9**, e1004043.
- Karimi M, Depicker A, Hilson P.** 2007. Recombinational cloning with plant gateway vectors. *Plant physiology* **145**, 1144–1154.
- Kim D, Paggi JM, Park C, Bennett C, Salzberg SL.** 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature biotechnology* **37**, 907–915.
- Kinoshita T, Ono N, Hayashi Y, Morimoto S, Nakamura S, Soda M, Kato Y, Ohnishi M, Nakano T, Inoue S.** 2011. *FLOWERING LOCUS T* regulates stomatal opening. *Current Biology* **21**, 1232–1238.
- Kloosterman B, Abelenda JA, Gomez M del MC, et al.** 2013. Naturally occurring allele diversity allows potato cultivation in northern latitudes. *Nature* **495**, 246–250.
- Kobayashi Y, Kaya H, Goto K, Iwabuchi M, Araki T.** 1999. A pair of related genes with antagonistic roles in mediating flowering signals. *Science* **286**, 1960–1962.
- Krieger U, Lippman ZB, Zamir D.** 2010. The flowering gene *SINGLE FLOWER TRUSS* drives heterosis for yield in tomato. *Nature genetics* **42**, 459–463.
- Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S.** 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research* **30**, 325–327.
- Liao Y, Smyth GK, Shi W.** 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic acids research* **47**, e47–e47.
- Lifschitz E, Ayre BG, Eshed Y.** 2014. Florigen and anti-florigen—a systemic mechanism for coordinating growth and termination in flowering plants. *Frontiers in Plant Science* **5**, 465.
- Lin T, Zhu G, Zhang J, et al.** 2014. Genomic analyses provide insights into the history of tomato breeding. *Nature Genetics* **46**, 1220–1226.
- Lynch JP.** 2019. Root phenotypes for improved nutrient capture: an underexploited opportunity for global agriculture. *New Phytologist* **223**, 548–564.
- Martí E, Gisbert C, Bishop GJ, Dixon MS, García-Martínez JL.** 2006. Genetic and physiological characterization of tomato cv. Micro-Tom. *Journal of experimental botany* **57**, 2037–2047.
- McGarry RC, Ayre BG.** 2012. Manipulating plant architecture with members of the CETS gene family. *Plant Science* **188–189**, 71–81.
- McGarry RC, Ayre BG.** 2021. Cotton architecture: examining the roles of *SINGLE FLOWER TRUSS* and *SELF-PRUNING* in regulating growth habits of a woody perennial crop. *Current Opinion in Plant Biology* **59**, 101968.
- Meyer RS, Purugganan MD.** 2013. Evolution of crop species: genetics of domestication and diversification. *Nature reviews genetics* **14**, 840–852.
- Milla R, Osborne CP, Turcotte MM, Violle C.** 2015. Plant domestication through an ecological lens. *Trends in ecology & evolution* **30**, 463–469.
- Mimida N, Goto K, Kobayashi Y, Araki T, Ahn JH, Weigel D, Murata M, Motoyoshi F, Sakamoto, W.** 2001. Functional divergence of the *TFLI*-like gene family in *Arabidopsis* revealed by characterization of a novel homologue. *Genes to Cells* **6**, 327–336.

- Murashige T, Skoog F.** 1962. A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* **15**, 473–497.
- Navarro C, Abelenda JA, Cruz-Oró E, Cuellar CA, Tamaki S, Silva J, Shimamoto K, Prat S.** 2011. Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. *Nature* **478**, 119–122.
- Ogiso-Tanaka E, Matsubara K, Yamamoto S, et al.** 2013. Natural variation of the *RICE FLOWERING LOCUS T1* contributes to flowering time divergence in rice. *PloS one* **8**, e75959.
- Park SJ, Jiang K, Tal L, Yichie Y, Gar O, Zamir D, Eshed Y, Lippman ZB.** 2014. Optimization of crop productivity in tomato using induced mutations in the florigen pathway. *Nature genetics* **46**, 1337–1342.
- Penfield S.** 2017. Seed dormancy and germination. *Current Biology* **27**, R874–R878.
- Périlleux C, Bouché F, Randoux M, Orman-Ligeza B.** 2019. Turning Meristems into Fortresses. *Trends in Plant Science* **24**, 431–442.
- Pin PA, Benlloch R, Bonnet D, Wremerth-Weich E, Kraft T, Gielen JJ, Nilsson O.** 2010. An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science* **330**, 1397–1400.
- Pino LE, Lombardi-Crestana S, Azevedo MS, Scotton DC, Borgo L, Quecini V, Figueira A, Peres LE.** 2010. The Rg1 allele as a valuable tool for genetic transformation of the tomato 'Micro-Tom' model system. *Plant methods* **6**, 1–11.
- Pnueli L, Carmel-Goren L, Hareven D, Gutfinger T, Alvarez J, Ganai M, Zamir D, Lifschitz E.** 1998. The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. *Development* **125**, 1979–1989.
- Quadrana L, Almeida J, Otaiza SN, et al.** 2013. Transcriptional regulation of tocopherol biosynthesis in tomato. *Plant Molecular Biology* **81**, 309–325.
- Razifard H, Ramos A, Della Valle AL, et al.** 2020. Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution* **37**, 1118–1132.
- Robledo JM, Medeiros D, Vicente MH, Azevedo AA, Thompson AJ, Peres LE, Ribeiro DM, Araújo WL, Zsögön A.** 2020. Control of water-use efficiency by florigen. *Plant, cell & environment* **43**, 76–86.
- Rodríguez MV, Barrero JM, Corbineau F, Gubler F, Benech-Arnold RL.** 2015. Dormancy in cereals (not too much, not so little): about the mechanisms behind this trait. *Seed Science Research* **25**, 99–119.
- Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB.** 2017. Engineering quantitative trait variation for crop improvement by genome editing. *Cell* **171**, 470–480.
- Ruijter JM, Ramakers C, Hoogaars WMH, Karlen Y, Bakker O, Van den Hoff MJB, Moorman A.** 2009. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic acids research* **37**, e45–e45.
- Ryu JY, Lee H-J, Seo PJ, Jung J-H, Ahn JH, Park C-M.** 2014. The *Arabidopsis* floral repressor BFT delays flowering by competing with FT for FD binding under high salinity. *Molecular plant* **7**, 377–387.

- Ryu JY, Park C-M, Seo PJ.** 2011. The floral repressor *BROTHER OF FT AND TFL1* (*BFT*) modulates flowering initiation under high salinity in *Arabidopsis*. *Molecules and cells* **32**, 295–303.
- Silva WB, Vicente MH, Robledo JM, Reartes DS, Ferrari RC, Bianchetti R, Araújo WL, Freschi L, Peres LE, Zsögön A.** 2018. SELF-PRUNING acts synergistically with DIAGEOTROPICA to guide auxin responses and proper growth form. *Plant physiology* **176**, 2904–2916.
- Song J, Zhang S, Wang X, et al.** 2020. Variations in Both *FTL1* and *SP5G*, Two Tomato *FT* Paralogs, Control Day-Neutral Flowering. *Molecular Plant* **13**, 939–942.
- Soyk S, Müller NA, Park SJ, Schmalenbach I, Jiang K, Hayama R, Zhang L, Van Eck J, Jimenez-Gomez JM, Lippman ZB.** 2017. Variation in the flowering gene SELF PRUNING 5G promotes day-neutrality and early yield in tomato. *Nature Genetics* **49**, 162–168.
- Susila H, Jurić S, Liu L, et al.** 2021. Florigen sequestration in cellular membranes modulates temperature-responsive flowering. *Science* **373**, 1137–1142.
- Taylor IB.** 1986. Biosystematics of the tomato. In: Atherton JG., In: Rudich J, eds. The tomato crop: a scientific basis for improvement. London: Chapman & Hall, 1–34.
- Vaistij FE, Barros-Galvão T, Cole AF, Gilday AD, He Z, Li Y, Harvey D, Larson TR, Graham IA.** 2018. *MOTHER-OF-FT-AND-TFL1* represses seed germination under far-red light by modulating phytohormone responses in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **115**, 8442–8447.
- Vicente MH, Zsögön A, de Sá AFL, Ribeiro RV, Peres LEP.** 2015. Semi-determinate growth habit adjusts the vegetative-to-reproductive balance and increases productivity and water-use efficiency in tomato (*Solanum lycopersicum*). *Journal of Plant Physiology* **177**, 11–19.
- Wallace IM, O'sullivan O, Higgins DG, Notredame C.** 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic acids research* **34**, 1692–1699.
- Wang M, Li W, Fang C, et al.** 2018. Parallel selection on a dormancy gene during domestication of crops from multiple families. *Nature Genetics* **50**, 1435–1441.
- Weber E, Gruetznier R, Werner S, Engler C, Marillonnet S.** 2011. Assembly of designer TAL effectors by Golden Gate cloning. *PloS one* **6**, e19722.
- Wickland DP, Hanzawa Y.** 2015. The *FLOWERING LOCUS T/TERMINAL FLOWER 1* gene family: functional evolution and molecular mechanisms. *Molecular Plant* **8**, 983–997.
- Witten DM.** 2011. Classification and clustering of sequencing data using a Poisson model. *The Annals of Applied Statistics* **5**, 2493–2518.
- Xi W, Liu C, Hou X, Yu H.** 2010. *MOTHER OF FT AND TFL1* regulates seed germination through a negative feedback loop modulating ABA signaling in *Arabidopsis*. *The Plant Cell* **22**, 1733–1748.
- Yu X, Liu H, Sang N, Li Y, Zhang T, Sun J, Huang X.** 2019. Identification of cotton *MOTHER OF FT AND TFL1* homologs, *GhMFT1* and *GhMFT2*, involved in seed germination. *Plos one* **14**, e0215771.
- Zeevaart JAD.** 2006. Florigen Coming of Age after 70 Years. *THE PLANT CELL ONLINE* **18**, 1783–1789.

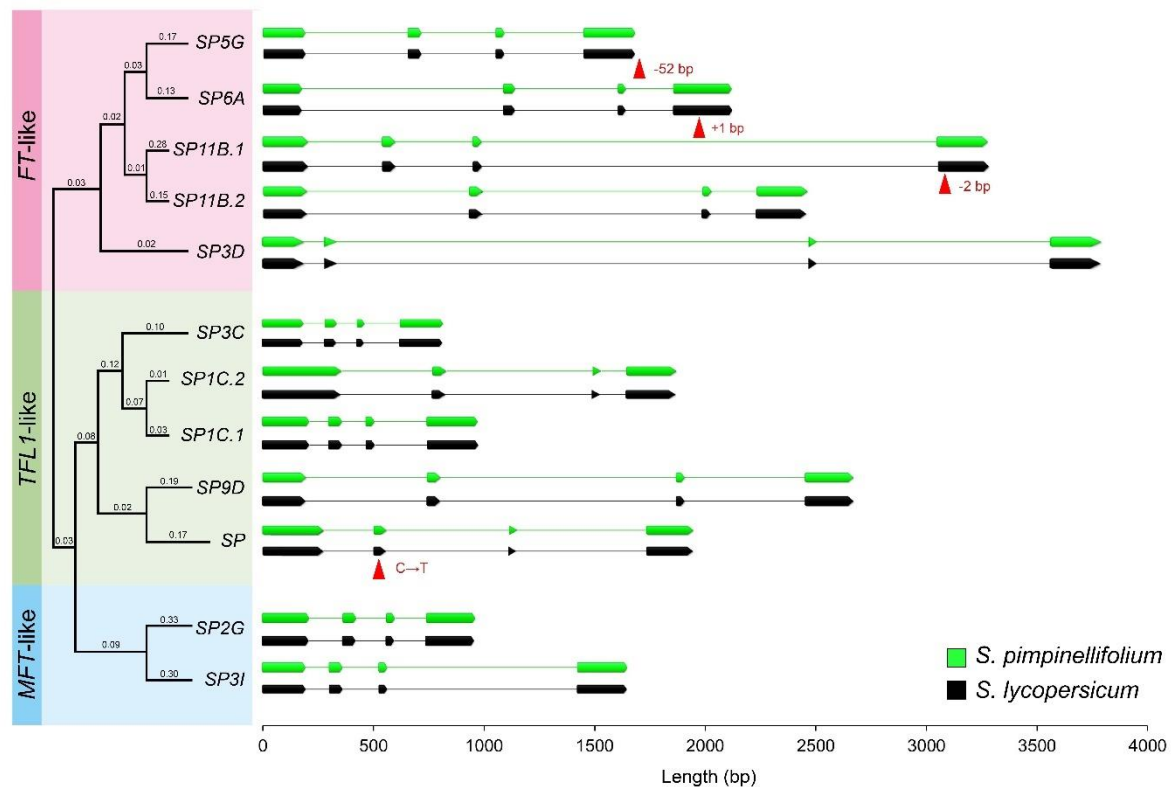
**Zhang S, Jiao Z, Liu L, Wang K, Zhong D, Li S, Zhao T, Xu X, Cui X.** 2018. Enhancer-promoter interaction of *SELF PRUNING 5G* shapes photoperiod adaptation. *Plant physiology* **178**, 1631–1642.

**Zsögön A, Čermák T, Naves ER, Notini MM, Edel KH, Weinl S, Freschi L, Voytas DF, Kudla J, Peres LEP.** 2018. De novo domestication of wild tomato using genome editing. *Nature biotechnology* **36**, 1211–1216.

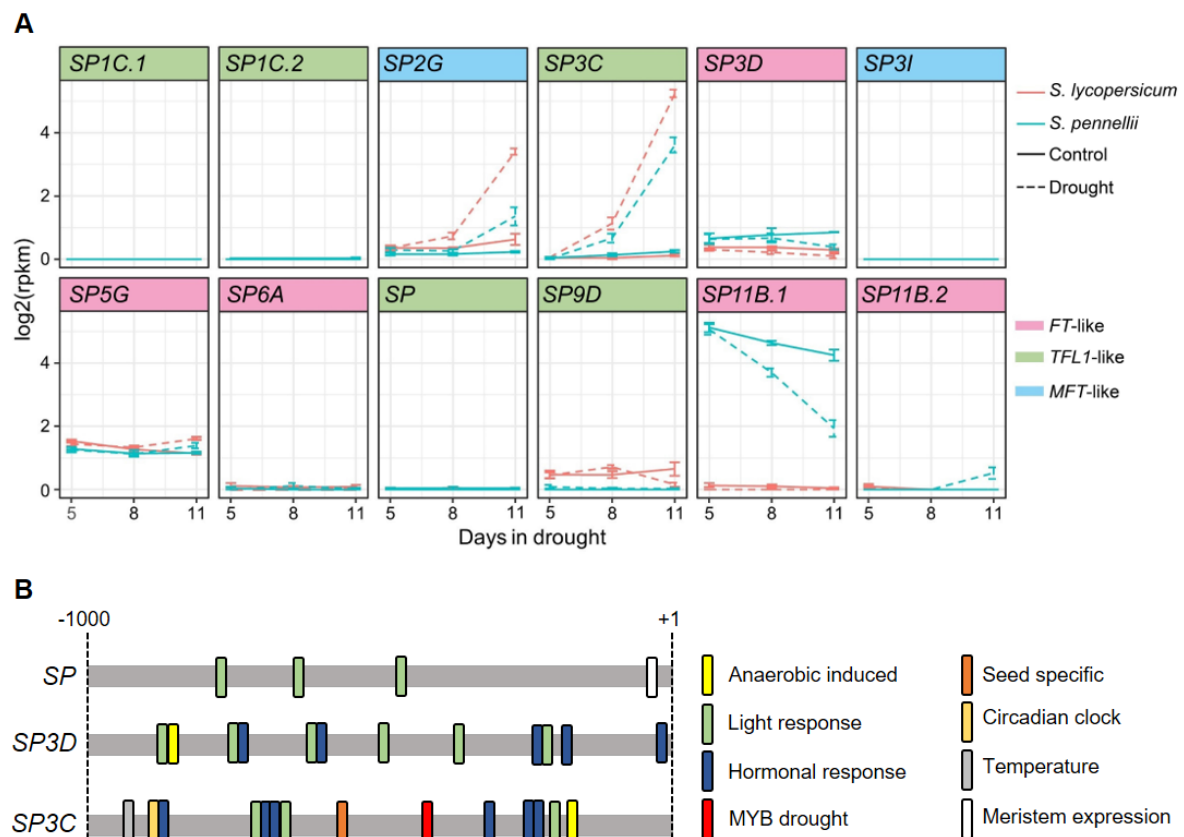
**Zsögön A, Cermak T, Voytas D, Peres LEP.** 2017. Genome editing as a tool to achieve the crop ideotype and de novo domestication of wild relatives: Case study in tomato. *Plant Science* **256**, 120–130.

**Zsögön A, Peres LE, Xiao Y, Yan J, Fernie AR.** 2021. Enhancing crop diversity for food security in the face of climate uncertainty. *The Plant Journal* doi.org/10.1111/tpj.15626

## Figures and legends

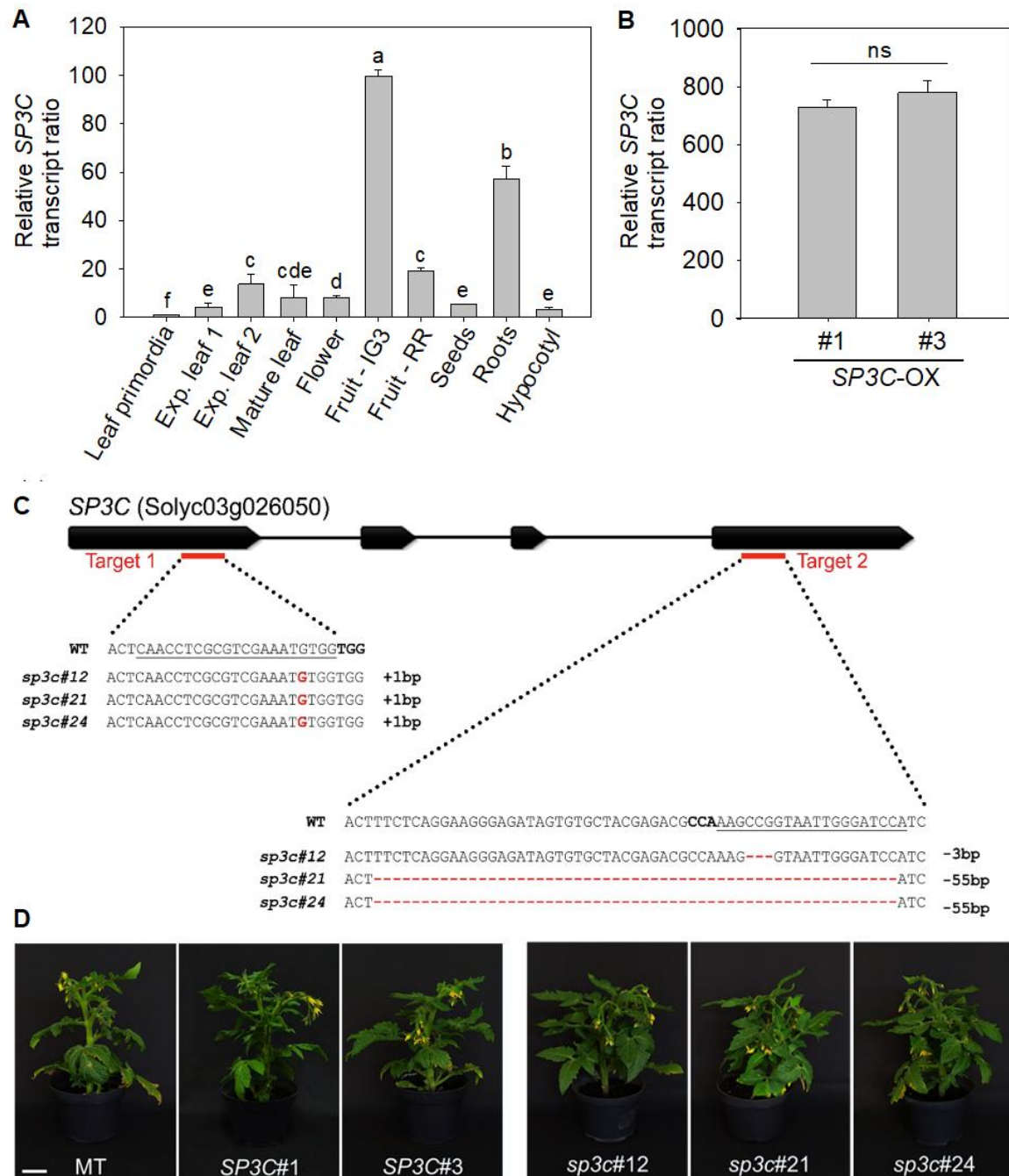


**Fig. 1. Allelic variation in the *CETS* gene family underpins tomato domestication and improvement.** Left: Phylogenetic reconstruction obtained from alignment of tomato *CETS* (*CENTRORADIALIS/TERMINAL FLOWER 1/SELF-PRUNING*) gene sequences. The three clades, based on the *Arabidopsis thaliana* homologues, are: *MOTHER OF FLOWERING LOCUS T* (MFT)-like genes (blue), *TERMINAL FLOWER 1* (TFL1)-like genes (green) and *FLOWERING LOCUS T* (FT)-like genes (pink). Values show the mean substitution rate per nucleotide. Right: Schematic alignment comparing the *CETS* genes structure in the putative tomato ancestral species *Solanum pimpinellifolium* (green) and domesticated tomato *Solanum lycopersicum* cv. Heinz 1706 (black). Diagrams show exons (boxes) and introns (lines). The previously described polymorphisms relevant to tomato domestication and improvement are indicated by red arrows (see text for details). The full sequences are listed in Supplementary Table S1.

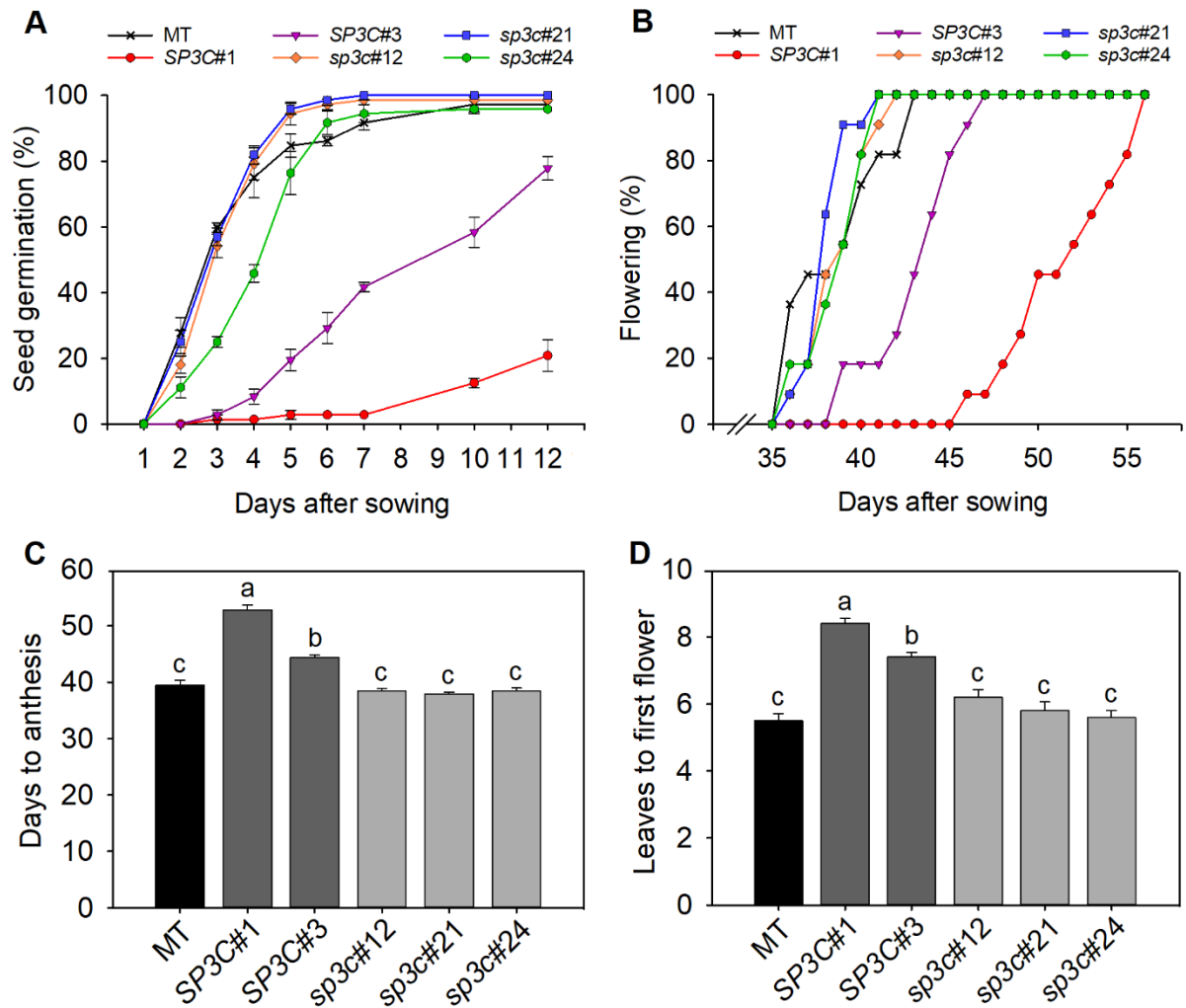


**Fig. 2. *SP3C* expression is induced by drought in tomato.** (A) Relative expression of *CETS* genes in leaves of *S. pennellii* (turquoise) and *S. lycopersicum* (magenta), collected at three points after suspension of watering (drought) compared to well-watered (control) plants. *SP2G* and *SP3C* were upregulated in both species along the time-course of drought, *SP3C* showing a more expressive response, and *SP11B.1* was downregulated as the drought progressed, but only in wild specie *S. pennellii*. (B) Putative transcription factor binding sites found in the promoter region (1000 bp upstream of the start codon) of the *SP*, *SP3D* (*SFT*) and *SP3C* genes from *S. lycopersicum*. Anaerobic induced, light response and hormonal response motifs are shared in two or three of these genes. While, MYB drought responsive and seed-specific motifs were exclusively found in the *SP3C* promoter.

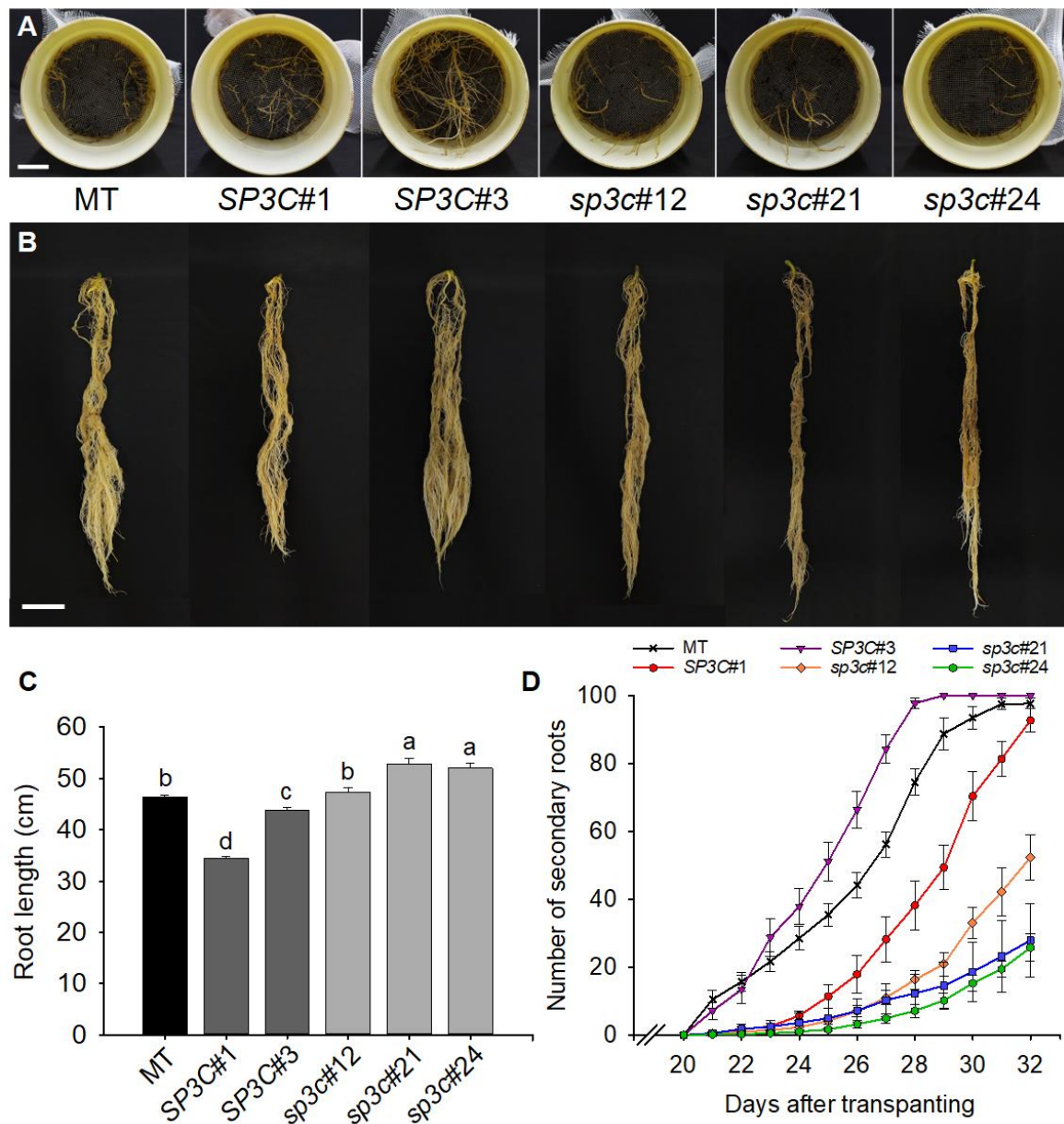




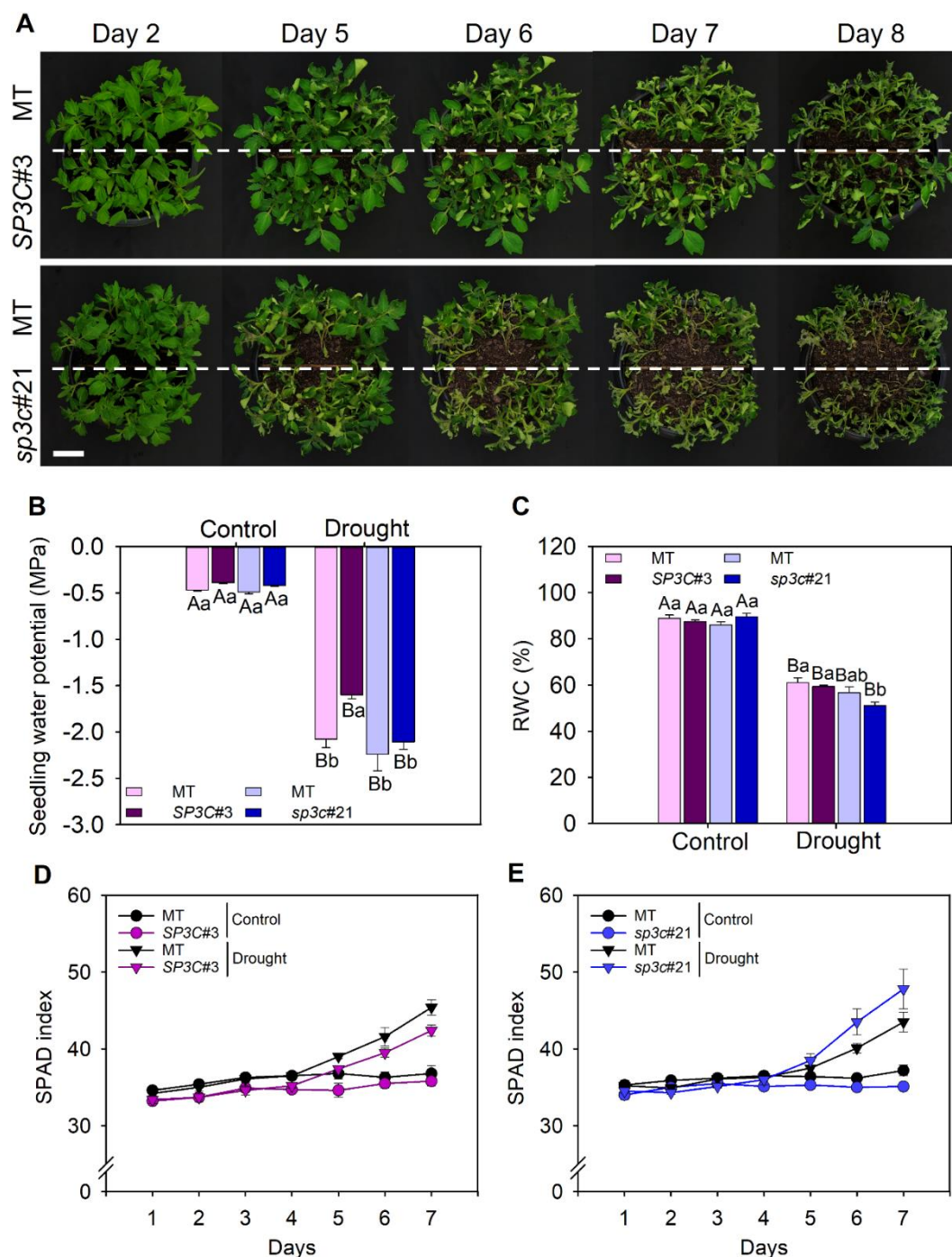
**Fig. 3. Generation of new allelic variants for *SP3C* in tomato cv. Micro-Tom (MT).** (A) Expression profile of *SP3C* in leaf primordia, expanding leaf 1, expanding leaf 2, mature leaf, flower, fruit at immature green 3 stage (IG3), fruit at red ripe stage (RR), seeds, roots and hypocotyl. Data were normalized against the leaf primordia sample. Values are means  $\pm$  SE (n=3) and different letters indicate significant differences between each analysed sample ( $P < 0.05$ ). (B) Relative *SP3C* transcript levels in young leaves of 35S::*SP3C*#1 (*SP3C*#1) and 35S::*SP3C*#3 (*SP3C*#3) overexpressing lines. Data were normalized against the MT wild-type. Values are means  $\pm$  SE (n=3). (C) CRISPR/Cas targeting to the first and fourth exon of *SP3C* using two guide RNAs (gRNAs); target 1 and target 2, red lines. The sequences below show the *sp3c*#12, *sp3c*#21 and *sp3c*#24 alleles identified from three independent transgenic plants, T<sub>3</sub> generation. gRNAs sequences are shown underlined, protospacer-adjacent motifs (PAMs) are highlighted in bold, insertions in red and deletions in red dashes. The sequence length differences compared to MT are shown to the right (D) Representative plants of MT, *SP3C* overexpressing lines (*SP3C*#1; *SP3C*#3) and *sp3c* loss-of-function mutant lines (*sp3c*#12; *sp3c*#21; *sp3c*#24) with 60 days after sowing. Scale bars, 5 cm.



**Fig. 4. *SP3C* represses seed germination and delays flowering.** (A) Germination time in tomato cultivar Micro-Tom (MT), *35S::SP3C#1* (*SP3C#1*) and *35S::SP3C#3* (*SP3C#3*) overexpression lines and *sp3c#12*, *sp3c#21* and *sp3c#24* mutant lines. The number of germinated seeds (in percentage) was evaluated daily for 12 days. Data are means  $\pm$  SE of four biological replicates with 18 seeds each. (B) Flowering time in MT, overexpression lines and mutants. (C) Mean number of days from sowing to anthesis. Data are means  $\pm$  SE ( $n = 10$  plants). Different letters indicate statistically significant differences (ANOVA + Tukey's test,  $P < 0.05$ ). (D) Number of leaves to first inflorescence. Data are means  $\pm$  SE ( $n = 12$  plants). Different letters indicate statistically significant differences (ANOVA + Tukey's test,  $P < 0.05$ ).



**Fig. 5. *SP3C* regulates root architecture.** (A) Representative photo of secondary roots emission in tomato cultivar Micro-Tom (MT), *SP3C* overexpressing lines (*SP3C*#1; *SP3C*#3) and *sp3c* loss-of-function mutant lines (*sp3c*#12; *sp3c*#21 and *sp3c*#24) 32 days after transplanting for vertical PVC cylinders. Scale bars, 2 cm. (B) Root morphology at 50 days after transplanting (dat). Scale bars, 2 cm. (C) Root length measurements at 50 dat. Data are means  $\pm$  SE (n = 10 plants). Different letters indicate statistically significant differences (ANOVA + Tukey's test, P < 0.05). (D) Time course of secondary root emission. The number of visible roots was measured at 15 dat for a period of 13 days.



**Fig. 6. SP3C improves tomato seedlings response during the drought.** (A) Representative photo of genotypes Micro-Tom (MT) grown together with overexpressing line (MT/SP3C#3) and Micro-Tom (MT) grown together with mutant line (MT/sp3c#21) throughout the assay. The leaf wilt symptoms were recorded daily, during eight days of drought. Visually, the seedlings began to show symptoms of water deficit on 5<sup>th</sup> day and were progressively intensified until 8<sup>th</sup> day of drought. The overexpression line SP3C#3 showed more subtle symptoms of winding and leaf wilt, while in mutant sp3c#21 the symptoms were clearly more intense. Scale bars, 2 cm. (B) Seedling water potential measured at 12:00 on the 8<sup>th</sup> day after suspension of watering, in MT/SP3C#3 (purple) or MT/sp3c#21 (blue). (C) Relative water content (RWC) obtained by the relation between fresh, turgid and dry weight of seedlings, on the 8<sup>th</sup> day after stress. Data are means  $\pm$  SE n = 5 plants. Different letters indicate statistically significant differences (ANOVA + Tukey's test, P < 0.05). (D) Chlorophyll content estimated by SPAD along the drought in MT/SP3C#3 and (E) MT/sp3c#21.

