CRANFIELD UNIVERSITY

DEMETRYUS SILVA FERREIRA

UNDERSTANDING THE GENETIC AND MORPHOLOGICAL BASIS OF *BUSHY ROOT* AND *BIFURICATE*, TWO MUTATIONS AFFECTING PLANT ARCHITECTURE IN *SOLANUM LYCOPERSICUM* L.

SCHOOL OF WATER, ENERGY and ENVIRONMENT

PhD in ENVIRONMENT and AGRIFOOD
Academic Year: 2013 - 2017

Supervisors:

Prof Andrew Thompson
Dr Fady Mohareb

March, 2017

CRANFIELD UNIVERSITY

SCHOOL OF WATER, ENERGY and ENVIRONMENT

PhD in Environment and AgriFood

Academic Year 2013 - 2017

DEMETRYUS SILVA FERREIRA

Understanding the genetic and morphological basis of *bushy root* and *bifuricate*, two mutations affecting plant architecture in *Solanum lycopersicum* L.

Supervisors:

Prof Andrew Thompson
Dr Fady Mohareb

March, 2017

ABSTRACT

The classical ethyl methanesulfonate (EMS) tomato mutant *bushy root* (*brt*) was studied using a homozygous near isogenic line (*brt*NIL) in the Micro-Tom (MT) genetic background. The mutation has a pleiotropic phenotype comprising slow seedling development, which may be a consequence of a maternally-inherited small seed phenotype, and a more compact, smaller but not bushier, root phenotype. The number of lateral roots, total root length and taproot size are all smaller in *brt*NIL than the WT. The *BRT* locus was mapped to a 137 kbp region containing 9 candidate genes on chr 12; an InDel in the promoter region of *Solyc12g014590* – containing two highly conserved pirin domains (Pirin_C and Pirin), was detected. Different expression patterns were confirmed by transcriptomic results, supporting *Solyc12g014590* as the gene responsible for the *brt* phenotype.

A naturally occurring recessive mutant named *bifuricate* (*bif*) shows an increase in inflorescence (truss) branching in comparison to the wild type (WT) control line, LAM183. In addition, the number of flowers per truss was 235% higher in *bif* plants than WT. Low temperature is known to increase truss branching, and so a four day low temperature treatment was applied and it was demonstrated that flowering increased significantly more in *bif* than in LAM183. The *BIF* locus was mapped to a 2.01 Mbp interval of chromosome 12 containing 53 genes. All coding region polymorphisms in the interval were surveyed, and two genes *Solyc12g019420* (a BTB/TAZ transcription factor) and *Solyc12g019460* (a MAP kinase) contained one stop codon predicted to disrupt gene function; both genes are excellent candidates for inflorescence branching control based on literature evidence. A newly developed introgression browser was used to demonstrate that the origin of the *bif* mutant haplotype is *Solanum galapagense*.

Keywords: branching; BTB-POZ BTB-TAZ; bushiness; EMS-mutant; low temperature; MAP kinase; root architecture; root branching; seed area; *Solanum galapagense;* transcriptomics; truss branching; Ulp1.

In memoriam of my dear mother

Mãe, muito obrigado por acreditar em mim.

Sinto sua falta todos os dias.

ACKNOWLEDGEMENTS

Firstly, I would like to express my profound gratitude to Prof Andrew Thompson, for all the advice and help during my journey. Thank you, Prof Thompson. Your support meant more than you will ever know, and I will never be able to express how grateful I am for this opportunity.

I would also like to thank Dr Fady Mohareb, as my second supervisor. Your support and kind words (especially during some panicking episodes in these last few months) were invaluable.

I am constantly complaining about the weather, in Cranfield, but I had a great time on the last 42 months. Surely, this would not be possible without the amazing people that I have known and learnt so much from.

In our research group, I would like to specially thank Zoltan, who started his journey in Cranfield with me, and taught me about molecular biology and so much more (e.g. good words in English and French) and Pedro, for helping me through one of the hardest moments of my life. Guys, in the last few months, we have shared much more than language, lab work, and weekend rotas. I am very grateful for your friendship and for your support. I would also like to thanks Tom, for always being there for me – especially when I had endless problems with bioinformatics. My heartfelt thanks to Jane, Jake, Melody and Emmanuel; I am not sure if I would be able to present this thesis if not by your support and love.

I am very grateful for the support provided by the staff during my time here. I would like to specially thank Zoe, Angel, Carol (and Dominick), Maria Carmen, Emma, S., Elsa, Sharon, Debbie, Adrienne and Senita.

Senita, I am not only thankful for your help with my IT problems, I am very grateful for your friendship and support – especially in these last final crazy weeks. You are the best and I am sure you will love Brazil.

During my time in Cranfield, I met special people. Some have become dear friends and will be with me for the rest of my life.

Alicia Rodriguez Sixtos Higuera Simoes…, I am so glad that we started our PhD at the same day/time. Your friendship is special and means the world for me. Alicinha, thanks for your unconditionally love and friendship through the brightest and darkest time. I would also like to thank you for loving my *Baileys* and *brigadeiros*, you are a hero.

Daniela, Pieter, Celine and Jolien, Thanks for the amazing time and trips. And, as promised, I would like to especially thank Daniela e Pieter for bringing me Kebab during my long nights in the glasshouse, and helping me when I was drowning tomatoes seedlings. You are the best and will be always be part of me.

I would also like to thank the *Portuguese gang* Joana: Francisco, Andre, Anish and Pablo; my friends from the water department; The *Spanish gang,* in special Esther, Fernando, Maria (ballena), Antonio, Juani, Francisco and Angela. Also would like to thank the support and love from the Mexicans that I met Cranfield, in particular, Citlali, Omar, Octavio, Victoria, Kazuo, Ali, Francis, Juan Carlos, Georgina, Silvia and Ricardo. Thanks guys!

I would also like to thank my fellow Brazilians in Cranfield, Tiago Matos, Ingrid, Guilherme, Vinicius, Renan and in special, Thiago Castro. *Obrigado*.

Thiago, thanks for the long chats over the phone, your everlasting support and your warm friendship.

Back in Brazil, I want to thank my friends and family; who did everything to support my decisions, no matter how crazy they were. Thank you all very much.

Specially, Silvia and Joyce who are much more than friends; are my sisters and endured me for hours on Skype and here in England! Silvia, thanks for the helping with my endless genotyping experiments. Joyce, thanks for sharing with me your experience and beers in Chester, I will always root for your success. You are amazing, and I am grateful for having you both in my life.

Tati and Glenda, your personality and advice were essential during the last years. Thank you for every laugh and for making every moment special.

Esther and Leonardo, Thanks for your support and help with the project. It was great to work with you both.

Judia, Sonia, Gaby and Mayara thanks for your support. I love you girls.

I am very grateful for siblings; Regiane, Renato, Junior, Fernanda and Thiago. I am especially grateful for my loving and caring sister-in-law Sandra and my cute nephew Bruno, who contributed so much in my PhD journey. Junior, thanks for your love.

Thiago, you are a part of me. Thank you for everything, including your astonishing math skills and your support. I love you very much.

It was a long and very arduous journey; I would not be able to conclude the PhD without the support of my family. In special, My dad, dinda, Andressa, Amanda, Conceicao, Josiane, Mariana, Carolina, Ilaura, Durvalino, Percilia, Lindomar, Laine, Wagner, Regiane and Jefferson.

I would like to thank and name every member of my family and friends who helped me through my time here. But, I would require 50 extra pages – so I hope you know how much you have contributed for my growth as a human being and as a researcher. What I am today is the result of my contact with you. So, thanks for your help, comfort and support.

Lastly, I would like to thank The Queen, Elizabeth II. Thank you for letting me in the country. It is very much appreciated.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

ABA-insensitive4	*abi4*
ABERRANT SPIKELET AND PANICLE1	*ASP1*
Abscisic acid	ABA
albescent	*alb*
Allele-specific PCR	AS-PCR
Aluminium	Al
ANANTHA	*AN*
APETALA1	*AP1*
Apical Meristem	AM
Auxin indole-3-acetic	AUX/IAA
Auxin indole-3-butyric	AUX/IBA
Beta carotene	*B*
bifuricate	*bif*
BLIND	*bl*
BRANCHED1	*BRC1*
Branching Meristem	BM
BRANCHING SILKLESS1	*BD1*
Brassinosteroids	BRs
Bric-A-Brac/Tramtrack/Broad complex	BTB
Bulk Segregant Analysis	BSA
Burrows-Wheeler Aligner	BWA
bushy root	*brt*
bushy root- Near Isogenic Line	*brt*NIL
Charge-Couple Device	CCD
Chromosome	Chr

Cleaved Amplified Polymorphic Sequence	CAPS
CLEVATA	CLE
Clustered Regularly Interspaced Short Palindromic Repeats	CRISPR
Combinatorial Fluorescence Energy Transfer	CFET
Common Primer Extension	CPE
COMPOUND INFLORESCENCE	*s*
C-TERMINALLY ENCODED PEPTIDE	CEP
CUPSHAPED COTYLEDONS	*CUC*
CyclinD2	*CYCD2*
Cytokinin	CK
Days After Germination	DAG
Days After Sowing	DAS
decreased apical dominance	*dad*
Denaturing high-performance liquid chromatography	DHPLC
DENSE AND ERECT PANICLE1	*DEP1*
Deoxynucleotide Triphosphates	dNTPs
Dicotyledonous	dicot
dwarf	*d*
dwarf root	*drt*
Ethyl Methanesulfonate	EMS
FALSIFLORA	*FA*
FAR-RED ELONGATED HYPOCOTYL3	*FHY3*
Floral Meristem	FM
Fold Change	FC
Food and Agriculture Organization of the United	FAO

Nations	
FRIZZY PANICLE	*FZP*
Fruit weight	*FW*
Genetic Modified Organisms	GMOs
Genome Analysis Tool Kit	GATK
Genotyping-By-Sequencing	GBS
Gibberellic acid	GA
GOBLET	*GOB*
High Affinity Transport System	HATS
High-Throughput	HT
Homologous Recombination	HR
HT-Next Generation Sequencing	HT-NGS
INDETERMINATE FLORAL APEX1	*IFA1*
INDETERMINATE SPIKELE1	*IDS1*
Inflorescence Meristem	IM
Inorganic phosphate	P_i
Insertion and Deletions	InDel
Iron	Fe
ISOPENTENYLTRANSFERASE	*IPT*
JOINTLESS	*j*
Kompetitive Allele Specific PCR	KASP
LANCELOTA	*LA*
LATERAL ORGAN BOUNDARIES1	*LOB1*
LATERAL SUPRESSOR	*Ls*
LEAFY	*LFY*
LEAFY HULL STERILE1	*LHS1*

MALDI- time of flight mass spectrometry	MALDI-TOF MS
Matrix-assisted laser desorption/ionization	MALDI
MicroRNA	miR
Micro-Tom	MT
miniature	*mnt*
mitogen-activated protein kinase	MAP kinase
Molecular inversion probe	MIP
Monocotyledons	monocot
more axillary branching1	*max1*
Mutator-like element	MULE
Next generation sequencing	NGS
Phosphite	PO_3^-
NITRATE-REGULATED1	ANR1
Nitrogen	N
Non-homologous end joining	NHEJ
P-glycoprotein	*PGP*
Phosphorus	P
PhytochromeB	phyB
PIN-FORMED1	PIN1
PINOID	PID
Polymerase chain reaction	PCR
ramosus	*rms*
Reduced-representation sequencing	RRS
REVERSED GERM ORIENTATION1	*RGO1*
RNA-sequencing	RNA-seq
Rolling circle amplification	RCA

self-pruning	*sp*
Shoot apical meristem	SAM
Single Molecule Real Time	SMRT
single-base extension	SBE
Single-nucleotide polymorphism	SNP
SISTER OF INDETERMINATE SPIKELET1	*SID1*
Spikelet meristem	SM
Strigolactones	SL
SUPERNUMERARY BRACT	*SNB*
tasselssed4	*ts4*
Tasselssed6	*Ts6*
TEOSINTE BRANCHED1	*TB1*
TERMINAL FLOWER1	*TFL1*
TILLER ANGLE CONTROL	*TAC1*
transcriptional adapter zinc finger	TAZ
True single molecule sequencing	tSMS
UNIFLORA	*UF*
UNUSUAL FLOWER ORGAN	*UFO*
Variant Call Format	VCF
Variant Effect Predictor	VEP
Whole genome resequencing	WGR
Wild type	WT
Zero-mode wave-guides	ZMWs

1 GENERAL INTRODUCTION

Tomato is a crop with high agricultural value, and the tomato world production has increased about 40% during the last ten years. Tomato cultivation takes place under protection in glasshouses or plastic tunnels and in open fields (Schwarz et al. 2013; FAOSTAT 2014). The cultivation method has a big impact on production efficiency and energy use (Muigai et al. 2003; Kirda et al. 2004; Hatirli et al. 2006; Ozkan et al. 2011). One reason for this improvement in world production is the results of research and breeding, which led to increased fruit quality and yield. According to the Food and Agriculture Organization of the United Nations (FAO), in 2013 the world production was around 163 million tonnes (FAOSTAT 2014). China, India and USA are the biggest tomatoes producers. Brazil is the eighth biggest tomato producer (fresh market) worldwide, with an average production of 4.3 million tonnes per year (FAOSTAT 2014). Around 42% of the world's production comes from countries with favourable conditions of solar radiation, air temperature and precipitation (Jędrszczyk et al. 2012). Food demand is sharply increasing; in 2100 the world population is predicted to reach 10 billion (DESA 2014), thus the pressure to increase food production is exponentially rising (Schmautz et al. 2016; Suhl et al. 2016; Bernstad et al. 2017).

Solanaceae plants, e.g. tomato, potato, pepper, eggplant and physalis, are some of the crops with the biggest agricultural value for the fresh or industrial markets (Rigano et al. 2013; Vélez et al. 2016; Gebhardt 2016; Garofalo et al. 2017). Due to a highly conserved genome throughout the Solanaceae family (Fernandez et al. 2009), tomato is used as a genetic model fruit crop to study host-pathogen interactions (Ji, J. Scott, et al. 2007), abiotic stress responses (Bahmani & Maali-Amiri 2015), yield (Marincs et al. 2017) and other key biological processes. This research has the potential to be converted to other fruit crops in tropical (banana, avocado, mango and papaya) and temperate (apple, peach and pear) fruits (Pino-Nunes et al. 2009; Lombardi-Crestana et al. 2012; Quinet et al. 2006; J. Wang et al. 2014; Belović et al. 2016). Tomato is also a good model organism due to other important

characteristics such as the short life cycle (12 weeks, Ichihashi & Sinha 2014), easy growth in open fields and greenhouses, relatively small and diploid genome, self-compatibility, availability of a diverse range of sexually compatible wild species (Emmanuel & Levy 2002), suitability for mutagenesis experiments employing DNA damaging agents such as ethyl methanesulfonate EMS (Just et al. 2013; Sikder et al. 2013), X-ray or fast-neutron (Meissner et al. 1997; Huther et al. 2013; Larbat et al. 2014; Kobayashi et al. 2014). In addition, tomato scientists have access to an extensive collection of single-gene mutants, a frequently updated reference genome, a database with thousands of validated single nucleotide polymorphisms (SNPs) and an active research community connecting a network of more than 50 countries (Fernandez et al. 2009; Shirasawa et al. 2010; The Tomato Genome Consortium 2012; Kim et al. 2014; Kobayashi et al. 2014). These resources allowed the publication of a substantial number of studies in different topics, for instance, the genetic basis of tomato. Breeders and geneticists rely on forward and reverse genetics strategies to uncover genes responsible for a particular phenotype. Forward genetics approaches begin with the description of a phenotype determined by Mendelian inheritance of a single gene. On the contrary, reverse genetics is looking for a phenotype in plants carrying mutation(s) for a defined gene(s).

1.1 Plant plasticity

Different and sometimes hostile environments have contributed to the capacity of plants to adapt their development and fitness (Santos et al. 2015; Mohammed et al. 2016). Phenotypic plasticity – differentiation among cells and tissues developed in response to external influences (Givnish 2002), has been studied in relation to the mechanisms of plant development (López-Bucio et al. 2003; Gandour et al. 2013). Phenotypic plasticity has evolved, but occurs regardless of the plant genotype in response to environmental interactions or under stochastic processes (Van Kleunen et al. 2005). This plasticity consists of changes in plant growth and metabolism (Mohammed et al. 2016). For instance, stem elongation represents an advantage to plants in a dense plant stand, with

competition for light, but early elongation also reduces structural stability and requires diversion of resources away from organs with the fundamental functional roles of photosynthesis (leaves) and water and nutrient capture (roots) (Weinig & Delph 2001; Dechaine et al. 2007).

Plants need to be fully able to acquire light, water and ions from their habitat, thus, plant architecture (see section 1.2.1, page 23) and developmental vigour are essential for the adaptation and exploration of available resources in the surrounding environment and exploitation of available resources. In order to best adapt to the surrounding environment, plants need to regulate their development at different levels, e.g. biomass investment in organs during vegetative growth and reproductive development (Vidyadhar et al. 2015; Poorter 2015). The morphology of plants is due to some important allometric relationships that control size and shape of plant organs (Niklas 2006; Busov et al. 2008). Plant size and architecture are complex traits to study because of coordinated growth between distinct organs. Plant development requires a balance between the fixing and use of carbon containing molecules, in other words, a balance is necessary between the “sink” and the “source” of assimilate in the plant.

1.1.1 Sink and source association

A balance between vegetative shoot growth (source) and the growth of sink organs such as truss and root is required for plant fitness. There must also be a balance between shoot demand for water and ions, and the ability of the root system to provide these resources. Therefore, both root and shoot development are intimately linked (Wissuwa et al. 2005; Vercruyssen et al. 2011; Osorio et al. 2014). Plant architecture is a complex trait that depends on several physiological processes regulating growth and development, for instance, nutrient uptake and carbon partitioning (Barthélémy & Caraglio 2007).

The soil nutrient profile influences plant growth and development. Nutrients as nitrogen and phosphorus are considered major limiting factors to

plant growth (Umehara et al. 2008; Yoneyama et al. 2012; Bianco et al. 2015; Mohammed et al. 2016). Nitrogen (N) is extensively used in fertilizers to ensure high yields, which has caused environmental problems such as eutrophication. The role of N in plant metabolism has been extensively studied and it well characterized (Nagel et al. 2001; Bénard et al. 2009; Yoneyama et al. 2012; Luo et al. 2015) in order to make its use more efficient.

Phosphorus (P) is also crucial for plant development, but is not readily available in many farming sites (Vejchasarn et al. 2016). Thus, P is also used as a fertilizer to increase P availability. It has been reported that lack of regulation and inspection can lead to severe excess of P fertilization, which may contribute to environmental degradation (Cordell et al. 2009; Vejchasarn et al. 2016). Plants with P deficiency are common. In addition, the interaction between P, iron (Fe) and aluminium (Al) makes the ions absorption by the roots harder (Zhou et al. 2009). Over the last decades, different studies were conducted to characterise the plant response to P starvation (Lambers et al. 2015). Plants exposed to environments with low P availability show accumulation of sucrose in the phloem (Hammond & White 2011; Smet et al. 2012). This strategy is commonly adopted to increase carbon resources in the root, which will lead to an enlargement of the root to better explore the surroundings – generating phenotyping plasticity. It has been reported (although, it is not a consensus) that some specific traits in the root system like taproot diameter and length, and root branching are affected by low levels of P (Hammond & White 2011; Niu et al. 2013; Nadira et al. 2014; Li-xiang & Dan 2014).

Underlying phenotype plasticity, i.e. phenotype variation due to external factors, are many mechanistic pathways which rely on gene products whose activity is modulated by the environment to create the variation. Branching of plant roots and shoots is variable and related to genes affecting hormone metabolism, regulatory factors and the cell cycle (Kaul et al., 2000; Chae et al., 2013). All aspects in plant physiology and development are associated with hormone metabolism, for example elongation growth is controlled by auxin

transport and this is affected by several different genes including *P-glycoprotein (PGP)* (Higashide et al. 2014); branch outgrowth is affected by strigolactones that act through pathways influenced by genes such as *more axillary branching1 (max1)* (Martín-Trillo et al. 2011) – see section 1.3.1.2, page 29; also dwarfism and round leaves in plant with deficiency in the production/signalling of the brassinosteroid (BRs) hormone (Shannon & Meeks-Wagner 1991; Busov et al. 2008; Wang & Wu 2015) – see section 1.2.2.2, page 26.

Transcription factors also control plant organ, growth and development – e.g. negative regulation of fruit size affected by the *FW2.2* gene (Cong & Tanksley 2006)*;* reduced branching of lateral root regulated by GRAS and the *LATERAL SUPRESSOR (Ls)* gene (Lozano et al. 2009; Busch et al. 2011); and relation with TCP-domain affected by the *LANCELOTA* (*LA*) gene (Kaul et al., 2000; Busov et al., 2008). The amount of cells is proportional to the plant tissue volume and consequently to the organ size, but the genetic control, as mentioned before, is not final. As an example, the *CyclinD2 (CYCD2*) gene is reported to regulate the transition between G1 and S (i.e. phases of the cell cycle - (Calegari 2012). Studies on *CYCD2* overexpression led to an increased growth rate, but the plant final size was the same as the WT (Busov et al., 2008).

Studies on root architecture are started by visual characterization and comparison to established models. Research conducted with roots is not straightforward; roots are mostly underground and cannot be easily observed without disturbing the surrounding soil. Nevertheless, root architecture is extensively studied and investigated due to its important functions – uptake of water (Pop et al., 2011) and ions (Jung and McCouch, 2013), anchorage (Tao et al., 2010) and storage (Ubeda-Tomás et al., 2012).

1.2 Root architecture and root classification

Historically, it was arduous to find agreement between plant morphologists regarding root classification and description (Esau, 1977; Zobel and Waisel, 2010; Gratani, 2014). To improve research involving root architecture classification, the international society for root research (ISRR) recommended a standard nomenclature system (Zobel and Waisel, 2010). Currently, four different root categories are commonly used in tomato root classification – taproot, lateral root (LR), shoot-borne root and basal root. The taproot, also called “primary root”, is the first root to develop, from the radicle. LRs, shoot-borne and basal roots are secondary roots. Lateral root is the term adopted to describe any root branching from another. To assist the classification, branching should be considered and described in association with the “original root” (e.g. *third-order lateral of the taproot* or *second-order lateral of the basal root* (Zobel and Waisel, 2010; Esau, 1977). Shoot-borne are originated from the shoot, these roots are also considered as adventious; however, the term “adventious roots” is only commonly used by plant anatomists when describing monocotyledonous plants (e.g. maize). Basal roots are originated from the hypocotyl (Zobel and Waisel, 2010). The number of lateral, shoot-borne and basal roots is related with the efficiency of water and ion absorption as more roots surface increase the root area/soil volume ratio (Péret et al., 2009a).

However, the mechanisms controlling root initiation are still not clear due to a large number of factors that can influence the morphogenesis. Root initiation has been studied and reviewed in the last two decades (Laplaze et al., 2007; Laskowski et al., 1995; Lavenus et al., 2013; Péret et al., 2009b; Charlton, 1991; Hochholdinger et al., 2004; Guo et al., 2008). Root architecture, i.e. the result of different developmental processes, is ultimately controlled by genetic and environmental factors.

1.2.1 Classification systems in roots

Even with the changes in root architecture caused by wide ecological niche diversity, an overall scheme was proposed by Fitter (1987) to classify the form of the root system by considering: (i) the balance of primary and adventitious

roots: a large number of adventitious roots all connected to the stem base are produced by some species particularly monocotyledons, e.g. sugar-cane; these systems lack a single dominant axis, differently from dicotyledonous plants with a primary root system; (ii) the degree of branching in primary or adventitious roots systems, as described by several models (Tisdall & Oades 1982; Coutts 1983; Fitter et al. 1991); (iii) the plasticity of branching, for example as influenced by nutrient availability (Fitter 1987), (Fig 1).

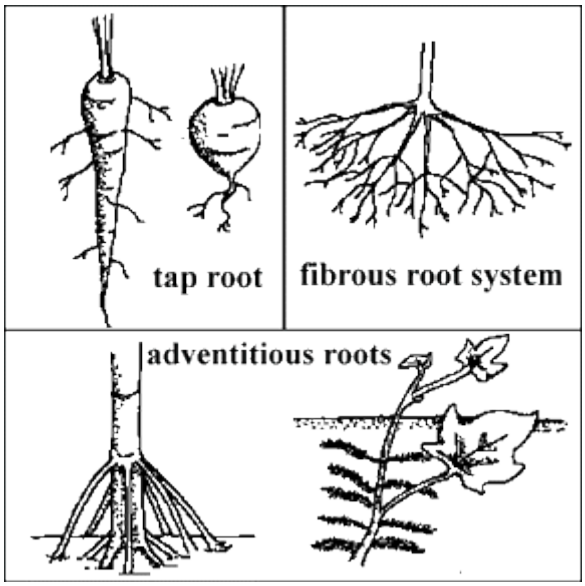


Figure 1: Examples of different root systems and architecture (available at: http://botany.csdl.tamu.edu/FLORA/Wilson/tfp/veg/roots.gif).

Alternative classification systems were suggested by Cannon (1949) based on the root activity and overall morphology. For example, root systems could be classified as very compact and very shallow or very extensive and very deep (Wahid 2000); or classification was based on root origin and development (Cannon 1949), e.g. primary root system type III, which has a long taproot with short laterals, whereas adventitious root system type VII has adventitious roots clustered and originating from an axis of the shoot. The latter classic system is the most used for root anatomists and soil scientists, because it considers different aspects of the plant plasticity, diversity and environment.

1.2.1.1 Root architecture: development and functions

As briefly mentioned before, dicotyledonous (e.g. beetroot) root systems develop from a primary root, thus, dicot root systems have a primary root system. The model of such system (dicot) is constructed based on the number of basal roots in older plants, and length of the laterals (extension covered) relative to the taproot; the root can be classified as taprooted or diffuse (Zobel 2005). On the case of tomato, most of the root system is developed from basal roots (Stofella 1983).

The root system in monocotyledons (monocots) is developed from the short-lived primary root and nodal roots, i.e. adventitious roots, forming a fibrous root system or adventitious system – are called “fasciculate root system” (Chun et al. 2005). In certain cereals, during early stages of the development, the primary root system is the main source of water and ions; until the adventitious root system is formed and established to assume this role (Klepper 1992). Developmental processes, including production of a new main axis, axial growth, radial growth and root senescence and decay (reviewed by Hodge et al. 2009) also impact the root system architecture.

As already mentioned in this study, the root has three main functions: anchorage, water and ion uptake and storage. But, in fact, it was reported that only 10% and 30% of the root length is involved with the assimilation of ions and water, respectively. Different functions between different classes of roots were reported, for example, seminal roots (early roots developed from the radical) have an important role in water absorption, but assimilate less P than the nodal roots – (Mistrik & Mistrikova 1995).

1.2.2 The molecular control of the root architecture

The root system is affected by a wide range of exogenous factors, including biotic factors such as helminths and pathogenic microorganisms, and abiotic factors such as high salinity and the mechanic impedance imposed by soils (Hodge et al. 2009). These factors cause extreme phenotypic differences in

individuals with the same genotype. Different approaches have been used to understand which genes are involved in environmental responses.

Although the phenotypic plasticity in roots is well established, the genetics (and molecular mechanisms) behind it are still obscure. Lately, published works on *Arabidopsis* is changing this scene by helping to elucidate the mechanisms of lateral root development (Casimiro et al. 2003), inflorescence regulation (Hanano & Goto 2011) and shoot-root interaction (Widman et al. 2014), for example.

1.2.2.1 Responses to nitrogen and phosphorus

Studies carried out in *Arabidopsis* proved that high concentrations of N in soil reduce LR and taproot length. On the other hand, plants suffering with nitrogen starvation produce longer LRs (Linkohr et al. 2002), but the LR density is decreased, i.e. less LRs are produced. This response is partly controlled by the MADS box transcription factor NITRATE-REGULATED1 (ANR1) as shown by the observation that *anr1* mutants have lower LR density in response to nitrate-rich zones (Zhang & Forde 1998). Root branching, and therefore root system architecture, is affected by N transport by the low- and high-affinity transport system (HATS) (Little et al. 2005). The low affinity transport system works when N is abundant, in contrast to HATS, which works when N is limiting (Glass et al. 2002). In addition, the *NRT2.1* gene from the NRT2 family of transporter-like proteins that is required for HATS, is reported to be involved with lateral root initiation (Little et al. 2005; Remans et al. 2006).

Experiments where plants were exposed to high levels (10 mM PO_3^-) of inorganic phosphate (P_i) showed an increase in taproot length, but decrease LR density and suppress LR elongation. Nevertheless, under low/moderate levels (1 mM – 0.5 mM PO_3^-), P_i increases the growth of the primary root (Linkohr et al. 2002). Extremely low levels (0.05 mM PO_3^-) of P_i cause a severe inhibition to taproot and LR development (López-Bucio et al. 2003). These effects were caused by a stop in meristematic function due to the lack of response to auxin

(López-Bucio et al. 2002); and were not displayed by plants on the control group, treated without P manipulation into the growth medium. Internal P_i concentration is important to the root architecture (Williamson et al. 2001). *PHO2* and miR-399 (induced under starvation), play a dominant role in P_i signalling pathway, they regulate the action of each other (Bari et al. 2006) – Figure 2.

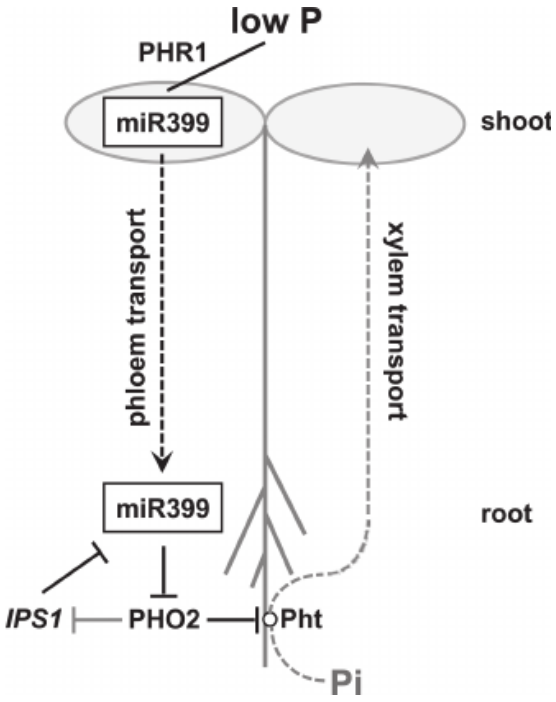


Figure 2: Phosphate homeostasis regulated by miR399 and PHO2. Low P concentration in the shoot results in the action of the transcription factor PHR1. PHR1 is reported to be involved with the induction of miR399 (Bari et al. 2006). Transported by the phloem, miR399 acts at the root, inhibiting the action of PHO2; resulting in the upregulation of Pi transporter genes (Pht), increasing the Pi uptake capacity and translocation via the xylem to the shoot (Adapted from: Scheible et al. 2011).

1.2.2.2 Phytohormones

The exogenous factors mentioned above affect the root architecture mainly by changing the plant hormonal profile, thus, its homeostasis. For instance, mutants in the auxin transport mediator *AXR4* show no response to N-rich-zones (Zhang et al. 1999), and the abscisic acid (ABA)-insensitive *abi4* mutant

does not show any reduction in LR density due to increased levels of N (Signora et al. 2001). This suggests that local responses to N are regulated by auxin and ABA would mediate systemic responses (Osmont et al. 2007).

As already described, the transport and signalling of auxin affect the root system architecture directly. Auxin indole-3-butyric acid (IBA) promotes the formation of adventitious roots, and for this reason, its signalling pathway has been extensively studied and discussed in monocotyledonous plants (Laskowski et al. 2006). Auxin indole-3-acetic acid (IAA) induces LR development when used in the growth media (Laskowski et al. 1995). Auxin is involved in all phases of LR development (i.e. initiation, emergence and elongation) (Liu et al. 2013). Acropetal auxin transporters, in the taproot, promote LR formation by stimulating cell division in the perycycle. Basipetal auxin transporters, in the root tip, also affect the LR formation (Casimiro et al. 2003). In addition, studies with mutants and transformed plants confirmed the function of auxin in LR development – mutants with increased auxin biosynthesis had more branched root systems (Boerjan et al. 1995). Accordingly, mutants with lower auxin production or with an impaired auxin signalling present a small root system, with less branching (Hobbie & Estelle 1995).

The well-established antagonism between cytokinin (CK) and auxin is also present in the root. CK regulates root growth/branching negatively – role confirmed by mutants with impaired cytokinin receptors AHK2 and AHK3, which show a quicker root development and more branched roots (Riefler et al. 2006).

Many auxin signalling mutants are also ethylene-insensitive, thus, ethylene is reported to be involved in some auxin responses (Swarup et al. 2002). In maize and *Arabidopsis*, moderate levels of ethylene are reported to constrain root growth, affecting the quiescent centre and root cap. It was suggested that, in rice, ethylene might regulate LR emergence due to the breakdown of cortical cells (Laskowski et al. 2006). Also in rice, gibberellic acid (GA) is reported to act with ethylene to stimulate adventitious root development

(Steffens et al. 2006). In general, GA regulates primary root develop and it is required for its growth (Kaneko et al. 2003; Gerald et al. 2006).

Brassinosteroids are reported to stimulate the taproot development at low concentrations, but are inhibitory at higher levels (Osmont et al. 2007). BRs will also affect the LR growth, by regulating auxin transport. Studies on BRs (Goda et al. 2002; Müssig et al. 2002; Nakamura et al. 2003; Nemhauser et al. 2004; Kim et al. 2006; Mouchel et al. 2006) suggests an important interaction between auxin and BR hormone signalling; e.g. the function of several genes involved in root development and architecture are regulated by both BR and auxin (Nakamura et al. 2003; Kim et al. 2006). The most reported function of BRs in root is related with its development (Mouchel et al. 2006); mutants with an impaired BR biosynthesis and signalling show a dwarf root phenotype (Noguchi et al. 1999).

In normal conditions, i.e. when the plant is not under stress, exogenous ABA and CK are reported to restrain root development; ABA reduces LR development and primary root growth (Casimiro et al. 2001), cytokinins reduce LR density (Werner et al. 2003). ABA and AUX/IAA interactions have been also suggested for similar alterations, i.e. reduced LR density, in the root system (Beaudoin et al. 2000; De Smet et al. 2003). However, under hydric stress, ABA has been reported to stimulate root growth by preventing the excess of ethylene production (Sharp & LeNoble 2002). ABA was also reported to be responsible for the arrest of shoot growth and development in drying soil (Sharp & LeNoble 2002) – Figure 3.

1.3 Shoot development and architecture

The shoot systems in angiosperms also display enormous phenotypic variation. This is mainly due to alterations to the truss architecture and branching or differences in the “growth habit” (Bell 1992).

Monopodial plants (e.g. *Arabidopsis*) present a simple architecture. The apical meristem (AM) is indeterminate during the plant growth and

development; all organs are laterally originated (e.g. leaves). Due to the strict flowering control, monopodial plants show distinct vegetative and reproductive phases (Pnueli et al. 2001).

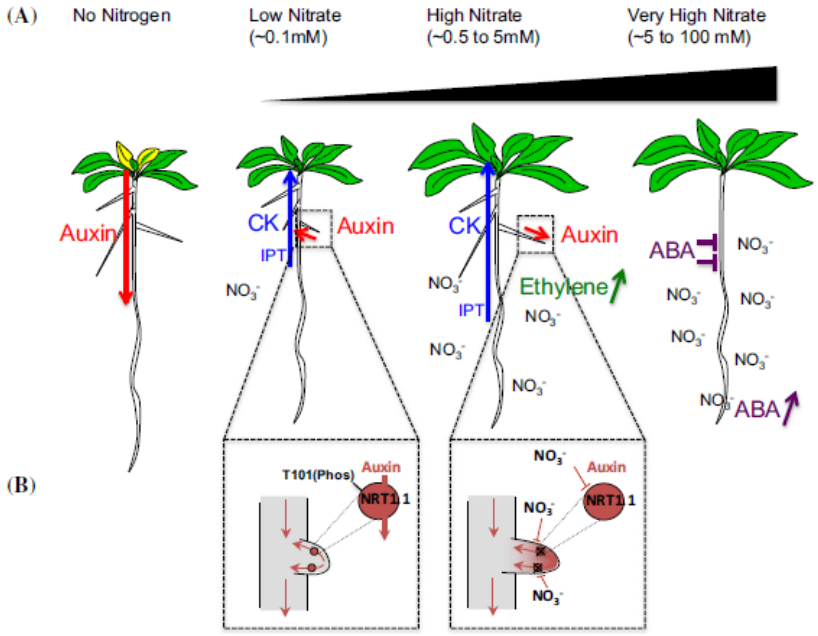


Figure 3: Hormonal control of the NO_3^- assimilation pathway and transport. (a) Summary of the concentration and effects of the phytohormones (Auxin, CK = Cytokinin, Ethylene and ABA = Abscisic acid), in different concentrations of NO_3^-. (b) Molecular mechanisms of auxin transport involving the NRT1.1 protein. In low NO_3^- concentration conditions, the phosphorylated form of NRT1.1 protein acts as an auxin transporter promoting auxin flux out of the root primordium. In higher NO_3^- conditions, NRT1.1 mediated auxin transport is bypassed which promotes auxin accumulation in the root primordium resulting in lateral root emergence (Adapted from Krouk, 2016).

In sympodial plants (e.g. tomato), the shoot is composed of different axes; the first inflorescence is produced at the termination of the primary vegetative apex (initial segment) after 8 to 12 leaves are produced (depending on the species or cultivar (Pnueli et al. 1998). The axillary bud, below the inflorescence, forms a new vegetative shoot (sympodial segment), which will give continuity to plant growth and development. The shoot architecture of sympodial species alternates after the termination of the initial segment, i.e. the sympodial segment will produce 3 leaves and then terminate with an inflorescence, another sympodial segment will be produced by the axillary bud

and the process will be repeated indefinitely (Sussex & Kerk 2001; Janssen et al. 2014) – Figure 4.

1.3.1 The molecular basis of the shoot control

Shoot- and root-specific organs start their formation in the postembryonic development (Domagalska & Leyser 2011), and regulation of this process can be exerted at this point. At the same time, secondary meristems are formed, e.g. cambium stripes, and the secondary shoot apical meristem (SAM) activity starts to increase, usually throughout development of leaves and then reduces its growth to form a dormant bud (Nordström et al. 2004). Depending on the source of the dormancy signal, the physiological state of the bud is classified into three different classes: (i) paradormancy, is induced by signals carried from outside the bud, e.g. phytohormones – Cline & Deppong 1999; Anderson et al. 2012); (ii) ecodormancy, is caused by environmental factor on the bud – (Shimizu-Sato & Mori 2001); (iii) endodormancy or true dormancy, is a strong arrest of bud growth and develop induced by internal physiological factors, maintained even in favourable conditions (Bilavcik et al. 2015). The molecular regulation of the bud growth and development is associated with several different mechanisms, outlined below and previously described (Yang & Jiao 2016).

1.3.1.1 Axillary meristem initiation

The axillary meristem is formed at the centre of the boundary zone; this region is essential for the meristem and organ maintenance development (Žádníková & Simon 2014). During the establishment of boundary zones, the transcription factor *LATERAL ORGAN BOUNDARIES1 (LOB1)* stimulates “phyB (phytochrome B) activation-tagged suppressor1” (*BAS1*) (Kebrom et al. 2010). *BAS1* is reported to negatively regulate brassinosteroid signal transduction (Neff et al. 1999; Bell et al. 2012).

During the initial growth of the leaf primordium, the auxin efflux carrier PIN-FORMED1 (PIN1) is oriented by the leaf primordium (Gendron et al. 2012). After the development of the boundary zone, PIN1 is reoriented towards the SAM (Q. Wang et al. 2014; Y. Wang et al. 2014). The PIN1 synchronisation to SAM happens through the action of the kinase PINOID (PID); which is responsible for the position of PIN1 in the plant (Furutani et al. 2004). An increase of auxin caused by the expression of the transgenic gene *iaaM* in the boundary zone, inhibited axillary meristem development; in contrast, when AUX/IAA protein BODENLOS is used to decrease auxin signalling in this area, the formation of axillary meristem is stimulated (Q. Wang et al. 2014).

Some genes have been reported to affect lateral organ development. For instance, *RPS10B* that encodes the S10e protein (Stirnberg, Liu, et al. 2012) and *FAR-RED ELONGATED HYPOCOTYL3* (*FHY3*), both essential for auxin homeostasis (Stirnberg, Zhao, et al. 2012). In addition, the *CUC* transcription factors, *CUPSHAPED COTYLEDONS 1, 2* and *3*, were reported to have a dominant, and yet redundant, role in meristem formation and development in *Arabidopsis* (Spinelli et al. 2011). In tomato, *GOBLET (GOB)* is an orthologue of the *CUC* genes (Busch et al. 2011). The development of SAM and formation of boundary zones are linked to the expression of these genes (Teichmann & Muhr 2015). *GOB* is down-regulated by BRs; therefore, low BR activity allows the higher expression of *GOB* genes (Bell et al. 2012; Gendron et al. 2012).

1.3.1.2 Axillary meristem control

The physiological response of the axillary buds is very dynamic, thus is an important trait for the shoot architecture. The axillary activity is negatively regulated by the shoot apex, i.e. apical dominance (Teichmann & Muhr 2015). The maintenance of apical dominance regulated by an internal signal, i.e. auxin (reported by Thimann & Skoog 1933), from a dominant shoot apex (Snow 1925). The biosynthesis of auxin occurs mainly in young leaves (Ljung et al. 2001), and it is transported downwards in the stem. Apical dominance is interrupted by removing the auxin source. On the other hand, exogenous

application of auxin on the apex can restore apical dominance (Thimann & Skoog 1933).

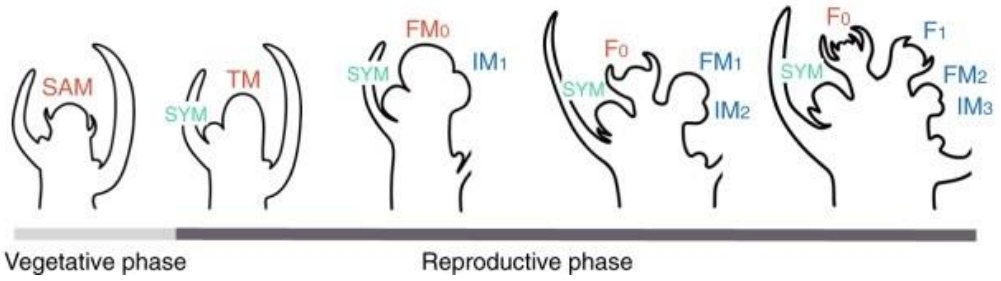


Figure 4: Inflorescence development in tomato. In the vegetative phase, the shoot apical meristem (SAM) initiates a leaf and an axillary meristem. Between the transition, the SAM takes an intermediate, transitional meristem (TM) fate whereas the last vegetative axillary meristem called the sympodial (SYM) takes over shoot growth (Perilli et al. 2012). The TM initiates a new phytomere composed by the inflorescence meristem (IM). TM and IM maturate toward floral meristem (FM) fate and become flowers (F) (Thouet et al. 2012). Each IM initiates another IM in the meantime of maturating to FM (Adapted from Périlleux, Lobet, & Tocquin, 2014).

The negative regulation, i.e. growth inhibition, of the bud is mediated by a signal transported upwards (Thimann & Skoog 1933). Both cytokinins and strigolactones (SL) are suggested to be involved in this negative regulation (Teichmann & Muhr 2015). Studies carried out in plants with different cytokinin levels confirmed the dominant role of CK in the bud control. Bud dormancy was broken when cytokinin was applied to the axillary buds, even with the intact plant apex (Sachs & Thimann 1967). Considering that the bud growth is controlled by auxin; it was suggested that the action of CK in the bud is regulated by auxin (Nordström et al. 2004). Auxin is transported downwards, from plant apex, to the stem reducing the expression of *ISOPENTENYLTRANSFERASE (IPT)* – gene responsible for the biosynthesis of CK (Tanaka et al. 2006).

Similar work in peas shows that auxin induces the CK oxidase gene *PsCKX2* (Shimizu-Sato et al. 2008). CK oxidases incapacitate CK action by converting active CK into an inactive form, and as a result, bud dormancy is maintained (Werner et al. 2001).

Root tips are a major site of CK biosynthesis (Miyawaki et al. 2004). However, studies where overexpressing CK mutant roots were grafted to WT scions showed that bud outgrowth was not stimulated (Faiss et al. 1997). These results suggested that cytokinin was not the only hormone involved in bud outgrowth regulation. Mutants showing a different branching pattern in *Arabidopsis*, *more axillary branching* (*max*), (Bennett et al. 2006); pea, *ramosus* (*rms*), (Beveridge et al. 1994; Beveridge et al. 1996; Sorefan et al. 2003); petunia, *decreased apical dominance* (*dad*), (Simons et al. 2007) and rice, *dwarf* (*d*), (Mori et al. 2002; Gao et al. 2009; Jiang et al. 2013) resulted in the identification of the dominant role in branching played by strigolactones. SLs are transported from roots through the xylem (Kohlen et al. 2011) to shoot (Beveridge et al. 1996; Sorefan et al. 2003; Bennett et al. 2006; Beveridge 2006). Biosynthesis of SLs is stimulated by auxin (Sorefan et al. 2003).

The biosynthesis of SL is reported to take place in the root and in the shoot (Auldridge et al. 2006; Umehara et al. 2008; MASHIGUCHI et al. 2009). Three possible repressors of SL signalling were recently identified: DELLA proteins (in *Arabidopsis*, Nakamura et al. 2013), *BES1* gene (in *Arabidopsis*, Zhou et al. 2013), and the D53 protein (in rice, Zhou et al. 2013). The D53 protein is reported to act by preventing SL degradation (Zhou et al. 2013). The *SUPPRESSOR OF MORE AXILLARY GROWTH2 LIKE* 7 in *Arabidopsis*, was suggested to have a similar function as D53 (Cardoso et al. 2014). Although some key precursors of SL are still being characterized and described (Abe et al. 2014), it is known that the SL signalling process involves the hormone binding to a receptor, activating an F-box protein-containing the CSF/E3 ligase complex (Skaar et al. 2013). This process is similar to the signalling process of auxin, jasmonic acid and GA (McSteen & Zhao 2008). The latter regulates the ubiquitination and subsequent degradation of a transcriptional repressor, leading to changes in gene transcription (Hagen & Guilfoyle 2002; Hartweck 2008; Memelink 2009).

After the evidence that apical dominance is regulated by auxin, there is still the transport velocity of auxin in the shoot to be considered: bud

development is initiated before auxin levels in the stem diminished when the apical auxin source was removed (Morris et al. 2005). For this reason, a different regulator could be involved with the bud control (Mason et al. 2014), for example, after the removal of the auxin source, there is escalation of sucrose levels in axillary buds. Additionally, buds broke dormancy after receiving a sucrose treatment (Teichmann & Muhr 2015) and measures of sucrose transport are consistent with the suggestion that sucrose is the acropetal signal regulating bud growth (Phillip 1975). Furthermore, sucrose treatment regulates negatively the branching suppressor *BRANCHED1* (*BRC1*) gene (Phillip 1975).

1.3.1.3 The meristematic control on inflorescence architecture

The shoot architecture of mature plants is determined by the axillary meristem initiation, bud growth and branch patterning (Cline 1997). Changes in theses aspects are responsible for the phenotypic variation observed in different species (Leyser 2005). The genotype is deeply responsible for this phenotypic variation, but, as already mentioned, responses to environment are essential and also have a considerable impact on the plant phenotype (Casal et al. 1986; López-Bucio et al. 2002; Yoneyama et al. 2013). Besides providing diversity to plants, truss architecture is an important trait for the reproductive success in angiosperms; agronomically speaking it is a trait that can strongly influence yield potential.

Arabidopsis, tomato (Figure 5), rice and maize are commonly used as models in the literature to describe the molecular mechanisms involved in inflorescence branching (Wang & Li 2008). Rice and maize inflorescence architecture is determinate before the floral meristem (FM) production (Harrop et al. 2016). But, in comparison to rice and tomato, maize shows a different and more complex inflorescence – the architecture development of both male and female in maize inflorescence, tassel and ear, respectively, is different (Lunde & Hake 2005) .

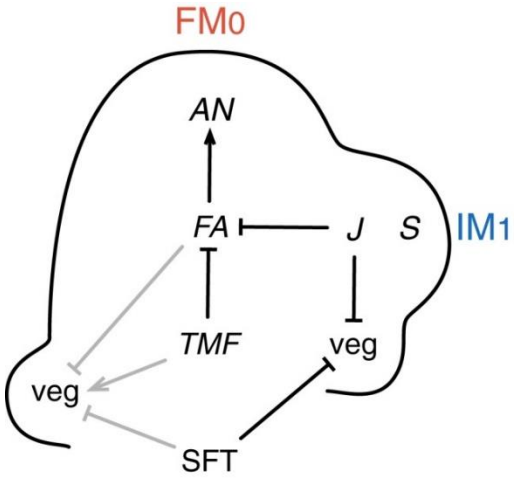


Figure 5: Network of genes involved with meristem fate in tomato inflorescences. Interactions between vegetative and reproductive phases are represented by the grey line. FALSIFLORA (FA) is upregulated by a SINGLE FLOWER TRUSS (SFT) signal, both repressing the vegetative growth (veg). TERMINATING FLOWER (TMF) has an antagonistic role to SFT, thus promoting veg by repressing FA. ANANTHA (AN) and FA are required for the maturation of inflorescence meristem (IM) to floral meristem (FM); AN is also reported to determine the FM identity. Different from AN, SINGLE FLOWER TRUSS (SFT) and JOINTLESS (J) are not required for FM identity, but are reported to be involved with the inhibition of the veg in IM. Both act preventing premature maturation of IM toward FM, possibly by repressing FA. COMPOUND INFLORESCENCE (S), on the other hand, accelerates IM maturation (adapted from Périlleux et al., 2014).

Non-orthologue genes have been suggested to influence shoot architecture in a redundant way by controlling the same regulatory pathways. (Kellogg 2007). In *Arabidopsis* and tomato, the floral meristem (FM) is directly produced by the inflorescence meristem (IM) – see chapter 3, page 80; (Schmitz & Theres 1999). In rice and maize, two intermediate meristems are formed between FM and IM (Tanaka et al. 2013). In grass inflorescences, after the SAM changes fate (i.e. transition between vegetative and reproductive phases), SAM is converted to IM; which will initiate primary, secondary or tertiary branching meristem (pBM, sBM or tBM – depending on the species; e.g. rice produces all types). The BM will induce initiation of the spikelet meristems (SM), which initiate the FMs. Finally, the floral meristem produces florets and

outer organs (pelea). Several genes have been reported to be involved in the initiation and regulation of the intermediate meristems.

The transition between the BM to SM is controlled by *BRANCHED SILKLESS1* (*BD1*) and *FRIZZY PANICLE* (*FZP*), in maize and rice, respectively (Chuck et al. 2002; Komatsu et al. 2003). Indeterminate branches are produced in *bd1* and *fzp* mutants (Tanaka et al. 2013), suggesting that these two genes regulate the BM establishment and the SM identity.

Studies with maize have shown that the transition between SM and FM is mediated by *INDETERMINATE SPIKELET1* (*IDS1*) and *SISTER OF INDETERMINATE SPIKELET1* (*SID1*) genes (Chuck et al. 1998; Chuck et al. 2008). Chuck et al. (2008) showed that in *ids1* and *sid1* double mutants, glumes are uninterruptedly formed, suggesting that the transition between the SM to FM is impaired. It was also suggested that *IDS1* and *SID1* have redundant functions in the control of inflorescence architecture, and also encode AP2/ERF transcription factors. Regarding the sex determination of the inflorescences, the mutants *tasselseed4 (ts4)* and *Tasselseed6 (Ts6)*, were used to determine the role of the miR172 microRNA family which inhibits the AP2 domain transcription factors function (Chuck et al. 2007). The *INDETERMINATE FLORAL APEX1 (IFA1)* and *REVERSED GERM ORIENTATION1 (RGO1)* genes are reported to be involved with the regulation of the SM identity in maize (Laudencia-Chingcuanco & Hake 2002; Kaplinsky & Freeling 2003). Hence, *IDS1*, *IFA1* and *RGO1* have redundant functions controlling the determinacy of the meristems in the reproductive phase of maize (Tanaka et al. 2013).

The *SUPERNUMERARY BRACT* (*SNB*) and *OsIDS1* genes have orthologous function to *IDS1* and *SID1* in rice (Lee et al. 2007; Lee & An 2012), suggesting that even with the different inflorescence architectures, the molecular mechanisms controlling the transition from the SM to the FM is conserved between maize and rice. The *LEAFY HULL STERILE1 (LHS1)* MADS box gene in rice is involved in SM identity – *lhs* mutants do not show the C terminus of the MADS box protein, resulting in an indeterminate SM as

phenotype (Jeon et al. 2000). Mutations in *ABERRANT SPIKELET AND PANICLE1 (ASP1*) are reported to impair the auxin signalling, thus affecting inflorescence architecture and spikelet development (Yoshida et al. 2012).

1.3.2 Shoot architecture in plant domestication

Some traits were prioritized during crop domestication, and it is likely that inflorescence architecture is one of those traits (Zhang & Yuang 2014). Different branching profiles in monocot plants are very evident – e.g. panicoids were selected to show a less tillering and reduced axillary development (Teichmann & Muhr 2015).

In cultivated maize crops, ideally, one single female inflorescence (ear) is produced and the apical axis is terminated in a male inflorescence (tassel). In contrast, wild maize species (e.g. teosinte) show multiple axillary branches developed at the same time; each new branch terminates with a tassel (Doebley et al. 1997). One of the most important genes during the maize domestication is *TEOSINTE BRANCHED1* (*TB1*). Doebey (1997) showed that even some minor changes in the expression of *TB1* are enough to cause substantial alterations in the shoot phenotype of maize and/or teosinte (Doebley et al. 1997). The *DENSE AND ERECT PANICLE1* (*DEP1*) gene is responsible for the increase in height and number of seeds (grains) per panicle observed in modern rice crops (Huang et al. 2009).

Genes involved with the signalling of SL were also selected during domestication (Teichmann & Muhr 2015). Reduced levels of strigolactones cause high tillering, suggesting that SL play a bigger role than already described in plant architecture (Cardoso et al. 2014).

The limited physical space for farming led to the selection of *TILLER ANGLE CONTROL1, TAC1* (Yu et al. 2007). Tillers of the wild rice (*Oryza rufipogon*) grow horizontally to supress competition with other plants, but also considerably reduce the density of plants per m^2 (Galli & Gallavotti 2016). After analysing 152 different accession, Yu et al. (2007) showed that all lines with

mutation on TAC1 show a decreased tiller angle, and concluded that the TAC1 mutation was under heavy selection (Yu et al. 2007).

1.4 Plant domestication in tomato: wild species as germplasm source

Due to morphological trait analysis (e.g. anther anatomy), taxonomists previously considered wild species of tomato as part of genus *Lycopersicon* (Grandillo et al. 2011). However, the latest molecular and phylogenetic results determined that tomato should be classified within the genus *Solanum* with the other important crops, as potato and eggplant (Knapp et al. 2004; Fischer et al. 2011; Grandillo et al. 2011). Regardless of the classification, wild species (WS) have been widely used by breeders and geneticists to use genetic variation into tomato breeding projects.

The in-breeding and selection for agronomic traits in small populations of tomato cultivars resulted in a narrow genetic base between these cultivars (Foolad 2007). On the other hand, wild species show extensive genetic variation. As a result, wild species are used to identify quantitative trait loci (QTLs) carrying polymorphisms related with high value traits for crops such as disease resistance, insect resistance, abiotic stress tolerance, yield, fruit quality and yield, for example (Grandillo et al. 2011).

Research centres in the United States of America started breeding programs using wild tomato species in the beginning of the 20th century, mainly to introduce disease resistance traits into cultivar species (Bai & Lindhout 2007). Afterwards, F_1 hybrid tomatoes became very popular and started to dominate the classic cultivars used for fresh market tomatoes. Initially (until the 50's), breeders were selecting and compiling traits for both the industrial processing and fresh markets, but due to different demands, crops started to be bred for different proposes. For fresh market, breeders initially selected for bigger yield, and then post-harvest shelf-life traits were considered very important. Today, the quality and nutritional profiles of fruits are the priority for

consumers. On the other hand, cultivars for industrial processing were bred to carry important qualities for the industry. For instance, tomato varieties were bred for mechanical harvesting by changing the growth habit (*sp* genes) and the fruit physiology (*jointless*) (Fridman et al. 2001).

Tomato can be affected by more than two-hundred diseases caused by a large number of pathogens (e.g. viruses and bacteria). All described resistance genes arose from wild species (WS) – i.e. *S. chilense*, *S. peruvianum*, *S. habrochaites* and mainly from *S. pimpinellifolium* (Foolad et al. 2014). Around 20 resistance genes were introgressed into different cultivars' genomes, providing resistance against around 42 major diseases. Research conducted in disease resistance represents the biggest use of the WS germplasm (Foolad 2007; Grandillo et al. 2011).

Insect resistance represents a significant advantage for crops and different cultivars. WS are a rich source of resistance against the most relevant insect pests of tomato. This is particularly true for *S. habrochaites,* which is resistant to at least 18 pest species (Ji, J. W. Scott, et al. 2007). Resistance genes were also found in *S. pennellii, S. pimpinellifolium, S. cheesmaniae, S. chmielewshii, S. peruvianum, S. corneliomulleri, S. arcanum* and *S. chilense* (Farrar & Kennedy 1991; Muigai et al. 2003). Several factors (e.g. trichrome characteristics) are related with arthropods resistance, but due to the rich natural diversity of the aggressor organisms, the selection and breeding of this trait has proven to be particularly complicated (Grandillo et al. 2011).

The diversity of environments where WS have been found is also very relevant because it has allowed the evolution of resistance to abiotic stresses such as drought and salinity. Breeding cultivars with such traits is an objective with large economic impact and it has been extensively studied in tomatoes (Cuartero & Fernández-Muñoz 1999; Santa-Cruz et al. 2002; Chaves et al. 2003; Fischer et al. 2011; Sandhu et al. 2014; Rogers & Benfey 2015). As an example, genes related with drought tolerance were identified in *S. pimpenellifolium, S. pennellii* and *S. sitiens*; and high tolerance to low

temperatures was demonstrated in *S. habrochaites* (Fischer et al. 2011; Arms et al. 2015).

The market for fresh tomato is very diverse regarding customer demand for traits (e.g. fruit colour, shape, sugar content) (Grandillo et al. 2011). Wild species show a very diverse range of fruit characteristics, for example *S. chmielewskii* present an increase of 40% of the total soluble solids, without significant yield loss. In some regions the red fruit colour is considered as an important trait (Labate et al. 2007). Studies with the *Beta carotene (B*) gene on chromosome 6, found in several WS, were essential to understand the fruit colour control in tomato. High expression of *B* increases considerably (15x) the level of pro-vitamin A (β-carotene) in the fruit (Labate et al. 2007; Grandillo et al. 2011).

1.5 Technologies for genomics

The analysis of distinctive DNA polymorphisms in individuals and populations is considered the foundation of modern molecular genetics. Such studies allow genotypes to be identified and connected to possible phenotypes. Single nucleotide polymorphisms (SNPs), especially, can be used in phylogenetic studies (McCormack et al. 2013). Improvements and advances in high-throughput (HT) DNA sequencing has allowed more efficient breeding and genetic mapping protocols using SNP markers; these polymorphisms can be directly linked to traits with economic importance, and can be readily fine-mapped to identify candidate genes and causative polymorphisms.

1.5.1 Sequencing tools and approaches

The rapid advance in technology has drastically changed the way complete genome sequencing may be accomplished. Illumina platforms of second- HT-next generation sequencing (HT-NGS) are able to generate billions of bases in a single lane run, which revolutionized the studies of DNA sequences (Schadt

et al. 2010). These second generation of HT-NGS are based on emulsion PCR amplification of DNA fragments. The amplification step was required to amplify the light signal for reliable base detection, but could also result into the addition of wrong bases (Schadt et al. 2010); which was resolved by eliminating this step on the third generation HT-NGS – i.e. based in the sequencing of a single DNA molecule.

Heliscope™ was released as the first single-molecule DNA sequencing platforms. Its sequencer is based on “true single molecule sequencing” (tSMS). In tSMS technology the DNA fragments are hybridized to oligonucleotides after the library preparation (Ozsolak et al. 2010). As a result, up to 28 Gb of data can be generated in a single sequencing run (Fatih Ozsolak & Milos 2011; F Ozsolak & Milos 2011).

Single Molecule Real Time (SMRT™) is a different platform based on real time sequencing in zero-mode wave-guides (ZMWs). The DNA fragment is extended using deoxynucleotide triphosphates (dNTPs) tagged with fluorescent labels, which are detected on a charge-coupled device (CCD) – (Levene et al. 2003; Eid et al. 2009). Pacific Biosciences, responsible for the development of the SMRT™ platform, claims that their sequencing system is capable of generating ~100 Gb/hour, with long reads (>10000 bp) in a single run (Pareek et al. 2011).

Real time single molecule DNA sequencers are currently being produced and improved. VisiGen Biotechnologies has designed a ‘real-time sensor’ for individual nucleotides tagged with specific fluorescent dyes. Although the platform is still under development, it was reported to generate 4 Gb of data per day, with reads 1 kb longer than any current platform (Selvin 2000).

The Nanopore DNA sequencer is different from the platforms previously described, because it is not based on fluorescent nucleotide and detection. The Nanopore DNA sequencer works by passing a DNA fragment through a nanopore while attached to a covalent charged molecule; this process identifies the characteristics (e.g. diameter) of the DNA molecule for base calling (Liang et al. 2017; Leung et al. 2017; Cornelis et al. 2017).

In general, the third generation platforms perform better than the second HT-NGS technologies. The accuracy and volume of generated data is higher, for *de novo* assembles longer reads can be obtained. Unfortunately, the third generation HT sequencing is still expensive, but with the technology fast developing, it has the potential to be cheaper and more efficient than second-NGS platforms, after proper establishment.

1.5.2 Genome sequencing and genotyping techniques

The traditional genotype identification and detection is accomplished with the use of markers (e.g. Cleaved Amplified Polymorphic Sequence, CAPS) based on allele-specific SNPs. Currently, almost all genotyping methods require PCR to amplify the target region containing the polymorphism, followed by an allelic discrimination step.

The allelic discrimination is based on allele-specific reactions. There are four classical methods: primer extension, hybridization, ligation, and enzymatic cleavage. In the primer extension approach, nucleotides in a primer are incorporated to a DNA template, by a specific enzyme to detect both alleles or with distinct primers for each allele. The common primer extension (CPE) reaction is based in a primer selected/designed to anneal in a sequence juxtaposed to a SNP site, which is amplified by a polymerase enzyme (Sokolov 1990). The primer can be designed to detect a single or multiple SNPs, which increased the commercialization of different CPE systems for SNP genotyping.

Allele detection by fluorescence: systems based in fluorescent signals are commonly used by genotyping platforms, due to its easy implementation and high sensitivity. Besides the usual applications (Kim & Misra 2007), fluorescent methods are employed in a single-base extension, SBE, (Hoogendoorn et al. 1999; Premstaller et al. 2002) and denaturing high-performance liquid chromatography (DHPLC) (Liu et al. 1998; Xiao & Oefner 2001).

Allele detection by chemiluminescence: a method with rapid detection, easy automation and high signal-to-noise ratio of SNP genotyping based in a cascade of enzyme reactions involving inorganic pyrophosphate (Ronaghi et al. 1996). SNP genotyping assays associate the PCR-amplified template DNA with a primer designed based on the SNP site (100). Chemiluminescency is also used for DNA sequencing methods, i.e. Pyrosequencing™ (Kim & Misra 2007).

Allele detection by mass: systems using matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), are the PinPoint assay (Haff & Smirnov 1997; Ross et al. 1998), MassEXTEND™ (Braun et al. 1997; Cashman et al. 2001), SPC-SBE (Kim et al. 2002), and GOOD assay (Sauer et al. 2002). Basically, the allelic discrimination on these platforms takes place by amplifying PCR products as a template. Each extension will have a different mass and correspond to one of the alleles of SNP (Le Hellard et al. 2002; Sanchez et al. 2003).

These platforms are based in SBE detection with primers tagged with fluorescent ddNTPs (Fan et al. 2000; Hirschhorn et al. 2000). Examples of these systems are the SNaPshot® (Kim & Misra 2007), the SNPstream™ assay (Pastinen et al. 1997; Nikiforov et al. 1994; Bell et al. 2002), allele-specific PCR (AS-PCR) (Gibbs et al. 1989; Medintz et al. 2001; Latorra et al. 2003) and Tag-It™ approach (Bortolin et al. 2004).

Hybridization techniques perform the allelic discrimination by detecting the mismatches in the target-probe. The efficiency of the differentiation relies on the length, sequence and SNP site in the probe. Some HT SNP assays were designed to use microarrays, and they have the advantage of enzyme exclusion to perform the allelic distinction – e.g. GeneChip® array (Kennedy et al. 2003; Matsuzaki et al. 2004) and TaqMan® genotyping assay (Holland et al. 1991; Livak 1999).

Ligation approaches uses ligase enzymes to conduct allelic discrimination. Two oligonucleotides are hybridized to a single-stranded DNA and then ligase enzymes assembling them to a single oligonucleotide. Three probes are used by a traditional ligation platforms; 2 allelic-specific and 1

common to merge the adjacent template to the SNP specific probe at the 3' end due to the higher enzymatic sensitivity (Landegren et al. 1988). Different methods can detect the SNPs in the ligation production – e.g. combinatorial fluorescence energy transfer - CFET (Nilsson et al. 1994; Tong et al. 2001), rolling circle amplification - RCA (Lizardi et al. 1998), molecular inversion probe - MIP (Hardenbol et al. 2003; Hardenbol et al. 2005).

Enzymatic cleavage can also be used for allele discrimination. This method works by the action of enzymes capable of cleaving the DNA, when recognizing specific sequences and structures. For instance, CAPS (Dillon et al. 2006; Foolad 2007; Tezuka et al. 2009). This method does not require any probes. On the other hand, genotyping assays based on restriction enzyme site have limited applicability due to the limited number of SNPs that can be used – e.g. the Invader® assay (Lyamichev et al. 1999).

As described above, many different assays can be used to screen SNPs in different species. These assays need to be robust, rapid and cost-effective (Ragoussis 2009). In my opinion, currently, two methods are particularly efficient for genotyping during genetic mapping, and these have different applications depending on numbers of SNPs and samples.

1.5.2.1 Genotyping-by-sequencing

A recently employed approach used for HT SNP detection and genotyping is genotyping-by-sequencing (GBS) (Andrews et al. 2016; Voss-Fels & Snowdon 2016). Compared the other methods based on NGS sequencing, GBS is significantly cheaper. GBS was successfully used as a breeding tool in important crops – e.g. maize (Elshire et al. 2011) and potato (Uitdewilligen et al. 2013). In association with phenotyping data, the GBS approach *can* increase the rapidity of mapping and detection of genes regulating important agronomic traits, which can be later introduced into cultivated crops (Edwards et al. 2013).

1.5.2.2 Bulk segregant analysis

The bulk segregant analysis (BSA) technique can dramatically decrease the costs of genotyping, especially when conducted to define an initial mapping position (Salunkhe et al. 2011). The BSA method is based on the sequencing/genotyping of two pools of DNA from individuals with the same phenotype (Venuprasad et al. 2009; Sun et al. 2010). In other words, a segregating population (F_2) is divided into two different pools (i.e. bulks), one for each phenotype (Michelmore et al. 1991). These two pools are sequenced by NGS and then screened compared for sequence differences (del Viso et al. 2012).

The number of individuals in each pool is important; enough genetic recombination events represented in both bulks will exclude non-causative SNPs. Furthermore, SNPs linked to a recessive causative mutation (by linkage) will be scored as homozygous in the bulk with the target trait, but will remain heterozygous in WT pools (del Viso et al. 2012). BSA applied to NILs using large pools are very effective (del Viso et al. 2012), especially in the case of mutations caused by ethyl methanesulfonate (EMS) because the exposure to EMS causes a relatively low number of mutations which subsequently need to be resolved (Ansari et al. 2013); this represents an efficient and faster genotyping method.

BSA was successfully used to map diverse traits in different organisms; wheat (Hu et al. 2016), maize (Farooqi et al. 2016), rice (Salunkhe et al. 2011; Vikram et al. 2012; Rani et al. 2013), melon fruit (Chayut et al. 2015), barley (Poulsen et al. 1995), yeast (Dunham 2012), and tomato (Lin et al. 2006; Elsayed et al. 2012; Truong et al. 2015).

1.5.3 CRISPR-Cas9

As an alternative to the traditional breeding, genome editing techniques can be used – e.g. meganucleases; zinc-finger nucleases, (Ito et al. 2015); transcription activator-like effector nucleases, TALEN, (Khlestkina & Shumny

2016) and clustered regularly interspaced short palindromic repeat-associated nuclease protein, CRISPR-CAS (Zsögön et al. 2017). So far, the newest and most efficient system is CRISPR-CAS9 (Shi et al. 2017).

In the CRISPR-Cas9 method, the target locus is broken by the action of Cas9 (CRISPR associated protein 9), and the DNA repair is carried out by non-homologous end joining (NHEJ), or by homologous recombination (HR)(Shi et al. 2017). NHEJ is responsible for some small InDels at the repair junction (Puchta 2017); HR, on the other hand, may occur in a conservative or a non-conservative pathway (Steinert et al. 2016). Although somatic cells mainly use NHEJ pathways (Shi et al. 2017; Puchta 2017), when homologous fragments are near to the target site (i.e. tandem duplications), the induced double-strand break is repaired by HR (Steinert et al. 2016).

Most genome engineering in plants is now performed with the CRISPR-Cas9 system, through the NHEJ pathway. The required nuclease is integrated into the modified plant, and later eliminated from the offspring by genetic segregation to produce plants with no transgenic DNA (Fauser et al. 2014). However, according to the regulations in some countries, these plants are still defined as genetic modified organisms (GMOs), due to the involvement of recombinant DNA technology in their production (Sprink et al. 2016).

CRISPR-Cas9 was used recently as an efficient tool in the genome editing of important crops like rice (Srivastava et al. 2017), maize (Char et al. 2017; Shi et al. 2017) and tomato (Ito et al. 2015; Xu et al. 2016; Soyk et al. 2017) – Figure 6.

1.5.4 Bioinformatics' tools for data analysis

The analysis and management of genomic data is as important as the chosen HT-SNP genotyping or mapping method. Initially, the raw data is submitted to a quality control tool; e.g. FastQC (Kroll et al. 2014), PRINSEQ (Schmieder & Edwards 2011), MultiQC (Ewels et al. 2016) and Qualimap 2 (Okonechnikov et al. 2015). In some cases, possible contamination may lead to low-quality reads that impact directly the mapping and SNP calling in later steps. For this reason,

filtering and trimming tools – e.g. FASTX-Toolkit (Qi et al. 2015), Trimmomatic (Bolger et al. 2014) and AdapterRemoval 2 (Schubert et al. 2016), are used to prepare the reads in the FASTQ file for the next step.

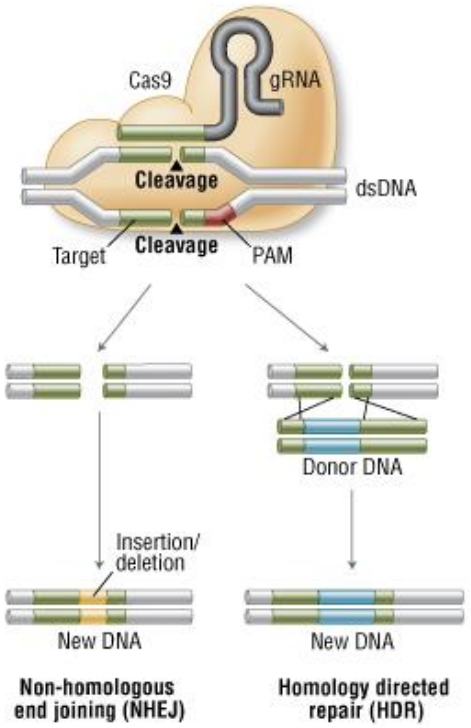


Figure 6: Genome editing using CRISPR-Cas9. With the help of the guide-RNA (gRNA), the nuclease (Cas9, CRISPR-associate protein) complex binds to the target site and cleaves the double-stranded DNA; activating the double-strand break repair machinery (PAM = protospacer-adjacent motif). In the absence of a homologous repair template, non-homologous end joining can result in InDels disrupting the target sequence (Puchta 2017). Alternatively, precise mutations can be made by providing a homologous repair template (Soyk et al. 2017) (Adapted from: https://www.neb.com/~/media/NebUs/Files/Feature%20Articles/Images/FA_Cas9_Fig2_Cas9forGenomeEditing.png).

Most of the important crops count with available reference genomes (Scheben et al. 2017). Thus, the reads are simply aligned to the reference using a read aligner – e.g. BWA-MEM (Li 2013); Bowtie2 (Langmead & Salzberg 2012); Tmap (Zhu et al. 2014) or Novoalign (Hwang et al. 2015). Once the reads are aligned, the steps to detect polymorphisms (i.e. SNPs and InDels) are

conducted by a variant callers (Nielsen et al. 2011) – e.g. the Genome Analysis Tool Kit, GATK (McKenna et al. 2010); SAMtools (H. Li et al. 2009); Freebayes (Garrison & Marth 2012); SOAPsnp (R. Li et al. 2009) and Platypus (Rimmer et al. 2014). Diverse studies were carried out comparing the efficacy of different aligners/variant callers, and they have reported different conclusions (Li 2013; O'Rawe et al. 2013; Pabinger et al. 2014; Zook et al. 2014; Clevenger et al. 2015), but it is clear that the detection method used by each variant caller contributes to these differences (Scheben et al. 2017). The methods used by SOAPsnp, SAMtools and GATK UnifiedGenotyper look for polymorphisms base-by-base. On the other hand, GATK HaplotypeCaller, FreeBayes and Platypus assemble local haplotypes to search for polymorphisms (Scheben et al. 2017). The methods for modelling and estimating sequencing errors are also different; some (e.g. FreeBayes and GATK) use Bayesian methods and others apply the Markov model to estimate error (Garrison & Marth 2012; Li 2013). Thus, due to the different data sets, methods applied and expected outcomes it is virtually impossible to designate a generally superior tool.

For GBS data, there are some common pipelines based on tools such as TASSEL-GBS (Glaubitz et al. 2014), Stacks (Catchen et al. 2011) and UNEAK (Lu et al. 2013). Again, each is used for different types of data sets: TASSEL-GBS is used to with large quantities of low-coverage reads, whereas UNEAK and Stacks are used in *de novo* methods for species without reference genomes (Scheben et al. 2017).

Once the variant is detected, an annotation step is applied to the data set. The objective is to identify variants in important regions such as exons and gene promoters. The annotation tools Annovar (Wang et al. 2010), SnpEff (Cingolani et al. 2012), Variant Effect Predictor, VEP (McLaren et al. 2010) and VariantAnnotation (Obenchain et al. 2014) require an annotated reference genome or transcript set. There is also a difference in a variant's predicted impact depending on the tools used for annotation (Scheben et al. 2017), due to the different methodologies applied by the tool. For instance, SnpEff (5kb) and

Annovar (1kb) use different metrics, i.e. distance between gene and variant, to characterise upstream and downstream regions (Hwang et al. 2015).

Data management is also something that needs to be considered when working with genomics. Currently, storage and integration are the biggest challenge faced by bioinformaticians and scientist in the field (Batley & Edwards 2009; Lee et al. 2012). One of the strategies adopted to deal with the vast amount of data is based on data compression using different file formats. For instance, NGS data is encoded in FASTQ files; using SAMtools FASTQ files can be transformed into binary (BAM) files, 3-4 times smaller. Another strategy accomplished with SAMtools (BCFtools) is to store only the detected variants in a variant call format (VCF) file (Danecek et al. 2011); and later compress it using tabix (Li 2011). Unfortunately, it is not uncommon for data to be lost during the analysis pipeline thus, raw FASTQ files are kept for security reasons, which still represents a considerable expenditure with storage hardware. It is being suggested that cloud-based storage systems can offer a solution to that problem (O'Driscoll et al. 2013). Public databases can also be part of the solution of this problem. Broad access to annotation files and other resources can help and serve as a centralised point of information to the research community and the public.

1.5.5 Reference genomes

Making use of second and third generation sequencing and high density SNP arrays, studies relating genotype and phenotype are becoming faster, cheaper and more straightforward. The tomato reference genome (cv. Heinz 1706) is publicly available and extensively used for SNP identification between different *Solanum* species genomes (Kim et al. 2014). As already described, NGS technology has become more accessible and cost effective, the quality of the outcome improves constantly. This explains why genomic resources are quickly increasing which promotes the whole-genome analyses, transcriptome characterization and detection of variations across various genomes (Sim, Durstewitz, et al. 2012).

The tomato genome reference was published by the Tomato Genome Consortium, in 2012. Since then, it has changed and improved on several occasions. In the beginning of February of 2017, the latest version SL3.0 was announced. The release includes new scaffolding arrangements, updated gene structures and ontology, very important integration of bacterial artificial chromosome (BAC) data, and new RNA-sequencing data (SGN 2017).

1.6 Aim

The aim of this project is to improve the quality and yield of *Solanum lycopersicum* L. by understanding the genetic control of plant architecture. The project will focus on branching of roots and flower trusses.

1.6.1 Objectives

Root architecture

1. To develop methodology for root phenotyping using qualitative imaging, and to use this to phenotype recombinants.
2. To study the environmental interaction of *bushy root* mutant.
3. To obtain genetic data from the *bushy root* mutant through deeper sequencing.
4. To develop additional SNP markers within the mapping interval for *brt* and to use them to obtain a map position for *brt.*
5. To fine map *brt* by searching for additional recombination event within the mapping interval.
6. To search for recombination in a F_2 population derived from a cross between Micro-Tom and near isogenic line containing the *brt* mutation.
7. To identify candidate genes in the mapping interval and test their function using transgenic experiments.
8. To identify the organs where the mutant gene is expressed.

Trusses architecture

1. To characterise the mutant phenotype.
2. To study the penetrance and environmental interaction of *bifuricate* and *FAT*.
3. To obtain genetic data from the LAM183 and *bifuricate* by sequencing both lines.
4. To delimit the introgression on the *bifuricate* material and determinate if its phenotype is the result of compound inflorescence (*s*) mutation.
5. To identify single nucleotide polymorphisms (SNP) linked to the *bifuricate/flower augmented (FAT) per truss* gene for use in mapping.
6. To search for recombination in a F_2 population derived from a cross between [(*LAM183* for *bifuricate*) and (Micro-tom for NIL-*FAT*)].
7. To genetically map *bifuricate and* NIL-*FAT* in a segregating population to a small genomic region.
8. To identify the genetic locus responsible for the *bifuricate* and *FAT* mutants.
9. To identify and characterise additional genetic sources of truss branching genes.
10. To breed the *bifuricate* trait into the MT genetic backcross.
11. To identify the organs where the mutant gene is expressed.

2 FINE MAPPING AND PHENOTYPIC CHARACTERISATION OF THE TOMATO *BUSHY ROOT* MUTATION ON CHROMOSOME 12

Silva Ferreira, D. K.[1]; Kevei, Z. L.[1]; Kurowski, T.[1]; Lazaro E. P. Peres[2]; Awan, S.Z.[3], Sergeant, M.J.[3], Mohareb, F.[1] and Thompson, A. J.[1]

[1]Cranfield Soil and AgriFood Institute, College road, Cranfield University, Bedfordshire, MK43 0AL, UK.

[2]ESALQ, Universidade de São Paulo (USP), Av. Pádua Dias, 11 - São Dimas, Piracicaba - SP, Brazil.

[3]University of Warwick, Coventry, CV4 7AL, UK.

Correspondence: Andrew Thompson. E-mail: a.j.thompson@cranfield.ac.uk

Highlight

Fine mapping of the classical root mutant *bushy root*, phenotype characterization and detection of insertion in the promoting region of an orthologue of *PR1*, composed by the highly conserved pirin domains.

Abstract

The tomato mutant *bushy root* (*brt*) arose from a cross between tomato accessions “Stock No. 1” and cultivar “Red Cherry” following EMS mutagenesis of pollen (Zobel 1972). Micro-Tom (MT) was used to produce a homozygous near isogenic line (*brt*NIL). The mutation has a pleiotropic phenotype: small seed size was observed, and it was determined that this phenotype is controlled by the maternal tissue; the small seed size is likely to explain the slow early development exhibited by *brt*NIL. The mutant root system was characterised by image analysis and can be described as a compact root system; i.e. Fewer and smaller lateral roots and shorter taproot. The *brt*NIL genome was re-sequenced and assembled in comparison to MT, against the reference *S. lycopersicum* Heinz 1706 (version SL2.50); this shows that the introgression spans 64 Mbp of the 67 Mbp on chr 12. After fine mapping, in the current interval, i.e. 137 kbp, a frameshift insertion was detected in the promoting region of a gene with highly conserved cupin domains. The *Pirin* gene (*Solyc12g014590*) is homologue to *PRN1* in *Arabidopsis* and the transcriptomics data supports Solyc12g014590 as a strong candidate for the *brt*NIL phenotype.

Key words

EMS-mutant, genome resequencing, root architecture, root mutant, pirin, quercetinase, *Solanum lycopersicum*, seed area, transcriptomics

2.1 INTRODUCTION

Plant architecture and developmental vigour is essential to the exploration of the environment and exploitation of available resources. Root architecture is well known for being plastic and essential for water and ion absorption, anchorage and storage (Nibau et al. 2008). Changes in root function and architecture have resulted in recent enhancements for crop production (Hammer et al. 2009) and plants with improved root systems are of great interest for vegetable production where elite scion genotypes with favourable aboveground traits are grafted as rootstocks that can be bred to address more specialised traits such as resistance to soil borne diseases and scion vigour (Schwarz et al. 2013; Asins et al. 2015). Rootstocks have been selected for diverse conditions, e.g. low nutrients availability (Schwarz et al., 2013), hydric stress (Wang et al., 2012), high salinity (Santa-Cruz et al., 2002) and pest control (Gregory et al., 2013). A lot has been done to understand root system architecture and development, particularly in *Arabidopsis* and cereal crops, but the knowledge of the underlying genetic mechanisms remains rudimentary in grafted vegetable crops such as tomato.

Recently, making use of second generation sequencing and high density single-nucleotide polymorphism (SNP) arrays, studies relating genotype and phenotype are becoming faster, cheaper and more straightforward. The tomato reference genome (cv. Heinz 1706) is public and extensively used for SNP identification between different genomes (Kim et al. 2014). Array-based genotyping techniques—able to score thousands of markers simultaneously—simplified the use of variants (SNPs) for genetic map construction and genome-wide investigation (Foolad 2007; Sim, van Deynze, et al. 2012). Use of NGS for variant discovery and high-throughput genotyping technologies now provides the opportunity for some classical root mutants (Table 1) to be easily investigated to search for causative genes (Kevei et al. 2015). Study of such mutants will enable a better understanding of the genetic control of root system architecture and development. The tomato mutant collection of the Tomato Genetics Resource Centre established by Dr Charles M. Rick (Davis, California)

(Emmanuel & Levy 2002) includes a small number of 17 classical tomato root mutants.

Table 1: Classical Mendelian mutations affecting tomato's root system.

Mutant	Phenotype	Chr	Reference
Aerial roots (aer)	High number of shoot-borne roots produced.	*	(Rick & Butler 1956)
Bushy root (brt)	Increases number of lateral roots, radical branches early	Chr 12	(Zobel 1972)
Bushy root 2 (brt2)	Slow initial development with a kinked root phenotype.	Chr 4	(Voland 1987)
Diageotropica (dgt)	Low geotropic response, changing the root architecture.	Chr 1	(Zobel 1972)
Dwarf root (drt)	Reduced hypocotyl and internodes, compact root phenotype.	Chr 2	(Rick & Butler 1956)
Rosette (ro)	Reduced root branching; single and short taproot.	Chr 2	(Rick & Butler 1956)
Root suppressed (Rs)	Reduced or non-existent root system.	Chr 4	(Rick & Butler 1956)
Twisted root (twr)	Abnormal trichomes development. Approximately, 38 days after planting the roots start to twist.	Chr 8	(Rick & Butler 1956)

The *bushy root* (*brt*) mutant was recovered in the F_2 progeny from a cross between ethyl methanesulfonate (EMS)-treated pollen from cv. California Red Cherry and a plant from Charles Rick's collection called "*Stock No. 1*". The cross was made to provide a preliminary linkage map using classical

phenotype-based genetic markers and to eliminate any accidental self-pollination by including a gene for male sterility in the parental line that acted as pollen recipient (Zobel 1972). Zobel was not able to provide a definitive *brt* phenotypic description due to the complicated genetic background that resulted from this cross. Nevertheless, the mutant was named "*bushy root*" due to early branching of lateral roots compared with the wild type (WT): the lateral root (LR) emergence in *brt* happened between 7 to 8 days after sowing (das) compared with the WT, which branched at 9 das (Zobel 1972). In addition to the early lateral root emergence, the *brt* root system was dwarf which gave the visual impression of being "bushier" than the WT, and the *brt* seedling leaves were reported to develop in a dark green and rugose appearance in early development (Zobel 1972). However, when the taproot had reached approximately 20 cm, the mutant plants grew normally until they became indistinguishable from the WT (Zobel 1972). The dwarf root growth habit and the "bushy" root system of the *brt* mutant are unique characteristics which can be distinguished from other dwarfs in classical linkage tests, including the *dwarf* (*D*) tomato mutant. After histological investigation, an intense starch accumulation was detected in the cortical parenchyma of the basal portion of *brt*, compared with both parents (i.e. cv. California Red Cherry and "*Stock No. 1*"). In young plants this pattern exhibited in the lower portion on the hypocotyl (Zobel 1973). Linkage results showed that *brt* is linked to *albescent (alb),* a classical gene located at the top (short arm) of chromosome 12 (Solanaceae project 2017). The tomato reference sequence (Asamizu et al. 2012) now facilitates the precise mapping of the *brt* gene. Additionally, to the availability of the reference genome of *S. lycopersicum* cv. Heinz 1706, two large studies have provided systematic resequencing of the genomes of 452 cultivars and wild species of tomato (Aflitos et al. 2014; Lin et al. 2014); several other smaller studies have also added to this list, and the MT (Kevei et al. 2015) genome resequencing data is available (SGN 2017).

Micro-Tom (MT) has been extensively used since it was first established as a model cultivar for genetic studies due to its small size. Micro-Tom (MT) has a strong dwarf phenotype and is the result of a cross between *Florida Basket*

and *Ohio 4013-3*, initially conceived as a novelty variety (Pons et al. 2014). At least three genes are responsible for the MT phenotype: the *self-pruning* (*sp*) gene (*Solyc06g074350*), which belongs to the *CELTS* family of regulatory genes (Pnueli et al. 1998; Carmel-Goren et al. 2003; Kobayashi et al. 2014), controls the interchange between vegetative and reproductive cycles in the tomato shoot and is responsible for the determinate growth habit that is suitable for field production with mechanical harvesting (Kaul et al. 2000; Saleem et al. 2013); the *dwarf* (*D*) gene (*Solyc02g089160*) is responsible for short internodes, and rugose, dark-green leaves (Meissner et al. 1997). This gene encodes *BR-6-oxidase* required for the biosynthesis of brassinolide – the most active brassinosteroid (BR) which was identified from pollen in *Brassica napus* (Shimada et al. 2001). Thus, MT contains a deficiency in BR biosynthesis – a phytohormone required for cell division and elongation growth (Martí et al. 2006). Finally MT carries the *miniature* (*mnt*) mutation (long arm of chromosome 11), that has yet not been well characterized, but it was suggested to be involved with the signalling and metabolism of gibberellin (GA) (Saito et al. 2009) and to contribute to the MT dwarf phenotype.

To facilitate genetic analysis, *brt* was introgressed into the MT genetic background to create a *brt* near isogenic line (*brt*NIL) which was characterized for growth and architecture traits and for its profile of phytohormones in comparison to MT. The results of the *brt*NIL characterization were somewhat different to Zobel's previous description – the colour of *brt* leaves was lighter, its early development was slower than the WT (MT) (Pino-Nunes & Peres 2005) and it was suggested that *brt* might have a low sensitivity to cytokinin which might cause some aspects of the *brt* phenotype such as smaller leaves and different leaf edge patter, as observed in plants overexpressing cytokinin oxidase (Werner et al. 2001; Bartrina et al. 2011). Auxin and ethylene were also tested, but both the *brt*NIL and WT showed similar sensitivity to these hormones (Pino-Nunes & Peres 2005; Pino-Nunes & Peres 2009). The *brt*NIL seeds were characterized as smaller than WT's seeds, it was suggested that the seeds size was maternally inherited (Pino-Nunes & Peres 2005; Pino-Nunes & Peres 2009). In this paper, we have characterised and analysed the *brt* phenotype in

the MT background and present a strong candidate gene likely to be responsible for the mutant phenotype.

2.2 MATERIALS AND METHODS

2.2.1 Plant material

Seeds were given accession numbers either prefixed with “DMF” or “WSS” and held in the Cranfield University seed collection. Seeds of MT (accession number WSS1254) and *brtNIL* (WSS1252) were provided by LEP (University of Sao Paulo - ESALQ); these seeds were sown and self-pollinated to bulk the seed stock (Figure 7).

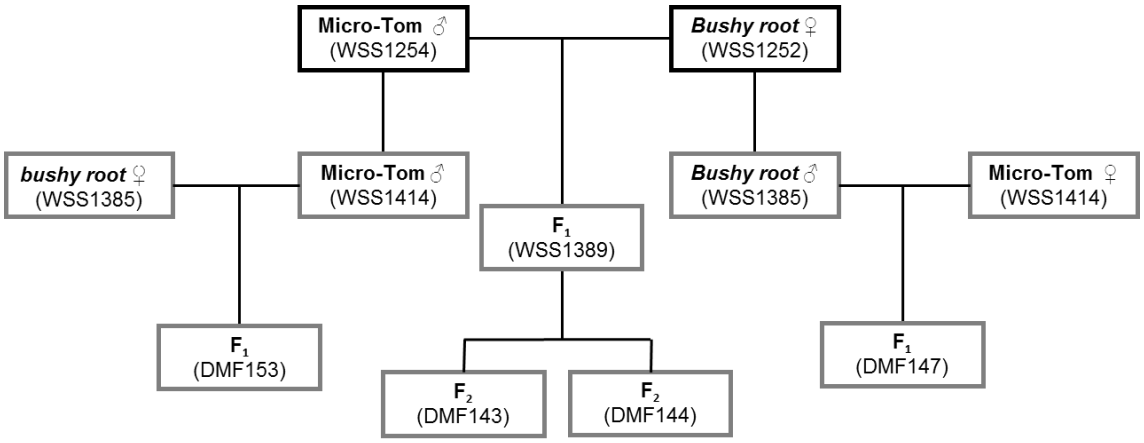


Figure 7: Seeds pedigree. Seeds produced in Brazil are represented by the black outline, while seeds produced in Cranfield (UK) are in grey. Genotyping experiments were carried using seeds from seeds package DMF143.

2.2.2 Seed extraction

Red ripe tomato fruits were cut in half and the recovered seeds were submerged in 1 gL^{-1} brewer’s pectolase (Ritchie, Burton-upon-Trent, UK) in 0.12 M HCl overnight at room temperature. The seeds were then washed thoroughly in tap water, and left to dry at room temperature for one week.

2.2.3 Plant growth

To avoid viral transmission, all seeds were sterilized by immersion for 10 minutes in an aqueous solution of 50% v/v Domestos (Unilever, UK) commercial bleach, equivalent to 2.5% w/v hypochlorite, and then thoroughly washed with tap water to remove residual bleach. After sterilization, seeds were placed in 9 cm Petri dishes containing two layers of Whatman No. 1 filter paper (85 mm diameter) soaked with 3 ml of tap water. Petri dishes were placed in the dark at 25°C for 3 days, and the chitted seeds were sown into 0.54 L round pots, 11 cm diameter x 8.6 cm depth, in a glasshouse with either 70% perlite (Sinclair, LBS Horticulture Ltd, Colne, Lancashire, UK) and 30% horticultural sand (Sporting Surface Supplies Ltd, Surrey, UK.) for phenotyping experiments or with compost (Sinclair multi-purpose Compost, LBS Horticulture Ltd, UK) for seed production. Pots were irrigated according to demand, and were fed twice a week with Hoagland solution (5 mM K_2SO_4; 1 mM H_3PO_4; 5 mM $Ca(NO_3)_2$; 2 mM $MgSO_4$; 100 µM EDTA Fe-Na; 42.2 µM H_3BO_3; 9.1 µM $MnCl_2$; 0.76 µM $ZnSO_4$ and 0.32 µM $CuSO_4$. pH = 5.8, adjusted with H_3PO_4) at half strength before flowering and full strength after flowering.

2.2.4 Root phenotyping

After the plant was removed from the pot and the growth medium was carefully washed away in running tap water, it was placed into a square Petri dish to be photographed. Images were captured using a Nikon Coolpix s8000, 14.2-megapixel digital camera and were analysed by ImageJ (Nakano et al. 2012; Tajima & Kato 2013). Root length was determined by setting a scale, and then drawing over the target roots with a straight or segmented line; length was provided through Image J menu commands: “analyse” > option “measure”). “Bushiness” was the term defined to describe the branching profile in the root system using three parameters obtained from the analysis at *ImageJ*: tap root length, total root length and number of laterals (Equation 1).

Equation 1: Mathematical expression used to calculate bushiness.

$$bushiness = \frac{Total\ root\ lenght\ \ x\ Number\ of\ lateral\ roots}{Taproot\ length}$$

2.2.5 Seed phenotyping

The seeds accessions *brt*NIL (WSS1385, n = 2714 seeds), MT (WSS1414, n = 826) and F_1 (MT x *brt*NIL, n = 719 and *brt*NIL x MT, n = 788), were scanned at 1800 dpi on a SE A3 USB 600 Pro flatbed scanner (Mustek, Hsin Chu, Taiwan). The seed area was determined from the scanned image using ImageJ (menu "analyse" > option "analyse particles").

2.2.6 NGS genomic data generation and analysis

Genomic DNA from MT and *brt*NIL plants was extracted using the DNeasy plant mini kit (Qiagen; Manchester, UK), according to the manufacturer's instructions. The MT genome was resequenced and made available by Kevei et al. (2015); the *brt*NIL genome, on the other hand, was resequenced on one lane of Illumina HiSeq 2500 platform generating 132 Million 126 bp paired-end reads. After the quality control performed by FastQC (Schmieder & Edwards 2011), to ensure that the reads used had high quality (on average above 33x); the raw data were mapped to the reference genome SL2.50 and genome annotation ITAG2.4, and analysed on Cranfield University's GRID, i.e. 32 CPUs 8GB/core, running a Linux environment on the Cranfield University GRID. Firstly, the reads were aligned by the Burrows-Wheeler aligner (BWA, version 0.7.4), using default specifications; secondly, the aligned reads were compressed into a binary (bam) format (Picard tools) and then sorted and indexed by Samtools (version 0.1.19); thirdly, the GATK package (Genome Analysis Tool Kit, Broad Institute, Cambridge, USA, version 3.3.0) was used to realign the Insertions and Deletions (InDels) and variant calling (HaplotypeCaller, using default settings).

This pipeline produced variant call format (VCF) files. The annotation of the VCF files was completed by SnpEff (version 4.0) using ITAG2.40 (associated with tomato genome reference version SL2.50) available on the Sol website (http://solgenomics.net/). Finally, the variants were filtered using GATK's *variant filtration* tool (Quality Depth < 2, Fisher Strand > 60, Mapping Quality < 40, Haplotype Score > 13 and Mapping Quality RankSum < 12.5) (Kevei et al. 2015). **Tom Kurowski** filtered unique polymorphisms using a custom Bash unpublished script which excluded polymorphisms shared between the data sets, similar to the mechanism of bedtools (Quinlan & Hall 2010). The filtered VCF file was uploaded to the GenoVerse Genome Browser (Bragin et al. 2012) which can be accessed on: http://elvis.misc.cranfield.ac.uk/GenoverseDemetryus/

2.2.7 DNA Extraction for genotyping

Genomic DNA was extracted from young leaves using a protocol based on the use of Chelex-100 (Bio-Rad, Hemel Hempstead, UK) with modifications (Wang et al. 2012; Casquet et al. 2012; Phillips et al. 2012; Walsh 1991; Turan et al. 2015; Pirllea et al. 2016). In brief, ~20-30 mg of leaf tissue from expanding leaves of < 2 cm length was disrupted in an extraction buffer containing 5% w/v Chelex 100 and soda glass balls (Smith Scientific Ltd, Kent, UK) by a mixer mill (VWR, Lutterworth, UK); the samples were subsequently incubated at 100^{o}C for 5 min in a PCR machine with a hot-lid (MJ Research, Canada). Lastly, the samples were centrifuged for 4 minutes at 5,509 × g and stored at 4^{o}C for short-term or at -20^{o}C for long-term (96-well format protocol detailed in Appendix A).

2.2.8 Genotyping by polymerase chain reaction (PCR)-based markers

Fine mapping was performed using additional Kompetitive Allele Specific PCR KASP markers designed in the mapping interval. The reactions were performed using 1 µl of purified genomic DNA (~ 0.1 µg) in a 10 µl reaction volume

containing 1x KASP master mix buffer (LGC, Teddington, UK). Using a CFX96 (Bio-Rad, Hemel Hempstead, UK) real-time PCR machine, thermal cycling was initiated at 94°C for 15 minutes, followed by 9 cycles of 94°C for 20 secs, 61 to 55°C for 1 minute (0.6°C drop per cycle), and then 25 cycles of 94°C for 20 sec, 55°C for 1 minute, then the temperature was decreased to 37°C for 1 minute for the final step of fluorescent plate-reading. KASP assays used fluorophores FAM and HEX for distinguishing genotypes; results were analysed in the Allelic Discrimination feature of CFX manager software (BioRad). All KASP assay (primers) were developed by LGC (UK) based on polymorphism and flanking sequence data provided; see Table 2).

Table 2: SNPs used to develop KASP markers on chromosome 12

Marker	Position (bp)	SNP
Alfa3.1	9060801	G/A
Alfa4.2	15142101	T/G
Alfa13-13 β	17280701	A/G
Alfa12 β	56204801	A/C
Alfa 1	5897101	G/A
Beta 1	2587701	C/T
Beta 1B	2588327	C/T
Beta 2	3012101	C/T
Beta 3	6464101	T/C
Beta 4	7053001	T/A
Beta 5	7561001	T/A
Beta 6	8083101	T/A
Beta 7	10581601	A/G
Beta 8	12330301	G/C
Beta 9	64145301	T/C
Gamma 1	2793261	G/A
Gamma 2	2899377	A/T

Gamma 3	3132238	A/T
Gamma 4	4193216	T/A
Gamma 5	4711140	T/A
Gamma 6	5104880	G/A
Gamma 7	3024605	A/G
Gamma 8	3766247	A/T
Gamma 9	2919476	G/A
Gamma 10	2559167	G/A
Gamma 11	7025456	C/T
Delta 1	4721022	T/G
Delta 2	5162086	C/A
Delta 3	5599054	T/C
Delta 4	5989307	G/T
Delta 5	6682212	G/T
Epsilon 1	5266967	G/C
Epsilon 2	5309547	C/T
Epsilon 3	5472682	G/C
Epsilon 4	5518140	A/T
Epsilon 5	5611274	A/G
Epsilon 6	5733516	C/A
Epsilon 7	5811912	T/C
Zeta 1	5609035	G/A
Zeta 2	5489626	G/T
Zeta 3	5498168	A/T
Zeta 4	5508534	C/T

2.2.9 Statistical analysis

Sample standard deviation and standard error were calculated using SigmaPlot. ANOVA was performed to evaluate the effects of the genotypes on the phenotype using SigmaPlot with significant difference claimed if $p < 0.05$ in a Tukey and Duun's post hoc test.

2.2.10 RNA-seq data generation and analysis

Chitted seeds from *brt*NIL recombinant 4 (DMF163) and MT (WSS1414) were allocated in rockwool blocks and covered by fine vermiculite. Using a randomised design, each rockwool block was placed into a round net basket, i.e. 5.5 cm diameter and 5 cm depth, in the low pressure aeroponics system in the glasshouse. This system works by constantly exposing the rookwool and roots to fine droplets of water (day 0 to 7 after transplanting), or half-strength Hoagland's nutrient solution (days 8 to 26 after transplanting). 27 days after transplanting the distal part of the root system (i.e. a portion of the taproot and laterals) was sampled from the plants, snap-frozen in liquid nitrogen and stored at −80°C. Total RNA was extracted from the samples using the Spectrum plant total RNA kit (Sigma), according to manufacturer's instructions. 12 RNA samples, i.e. 3 root- and 3 shoot samples per genotype, were sequenced in 2 lanes, using 100 bp paired-end reads by Illumina HiSeq 2000. The quality control on the RNA-seq data was performed using FastQC (Kroll et al. 2014). Then, the data was trimmed and filtered accordingly using Trimmomatic (Bolger et al. 2014), based on the following parameters: LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15 and MINLEN:40. The gene expression levels were calculated by the RSEM package (Li & Dewey 2011) using the STAR aligner version 2.4.a (Dobin et al. 2013). The expected counts for all transcripts were combined into a single matrix using the rsem-generate-data-matrix command. Differential gene expression analysis was performed in R using the limma package version 3.26.9. Normalisation was performed using the 'Voom With Quality Weights' method, i.e. which combine the voom observational-level weights with sample-specific quality weights in a designed experiment, with quantile normalization.

2.3 RESULTS

2.3.1 Root development and characterisation

Different traits were analysed based on the data available in the literature and visual inspection throughout the development of the *brt*NIL and MT. One of the most evident differences between *brt*NIL and MT was the taproot length. Initial experiments showed no length data overlap between both genotypes. The experiment was repeated using different (seven) time points during the plant development (Figure 8). In each point, 48 (24/genotype) root systems were analysed. The taproot of *brt*NIL plants is smaller than WT (MT) in all time points ($p < 0.005$).

2.3.2 Bushiness

Sixty plants (30 per genotype) were sown (perlite and sand mix), and 25 days after germination their roots were washed and photographed. The results show that *brt*NIL is 62.2% less bushy than the WT (MT). The *brt*NIL mutant presents a more compact (taproot length: *brt*NIL = 6.7 ± 0.9 and MT = 12.9 ± 0.3) root system with less (number of laterals: *brt*NIL = 38.6 ± 1.6 and MT = 95.5 ± 3.29) and smaller lateral roots, $p < 0.005$ (Figure 9).

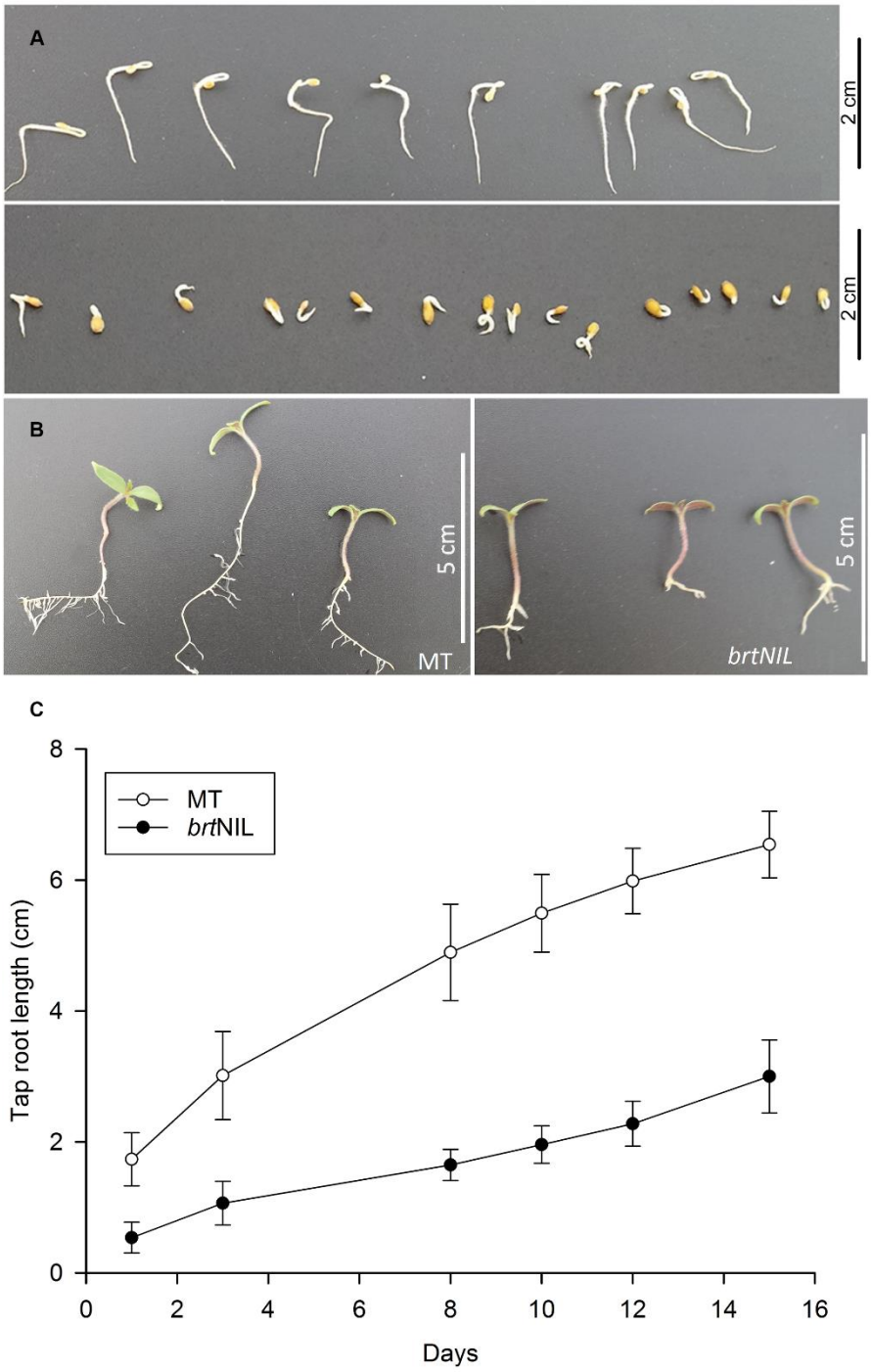


Figure8: Root development and taproot profile in *brt*NIL and MT: (a) Seedlings of Micro-Tom (MT) and *bushy root*-NIL (*brt*NIL) after the radical emergence, at the top and bottom panel, respectively. (b) The root system of both parents MT and *brt*NIL 11 days after sowing. (c) Taproot growth during the development of *brt*NIL and MT. Several plants were used for each time point (n=24). Means and standard deviations are sown (seeds were sown in filter paper).

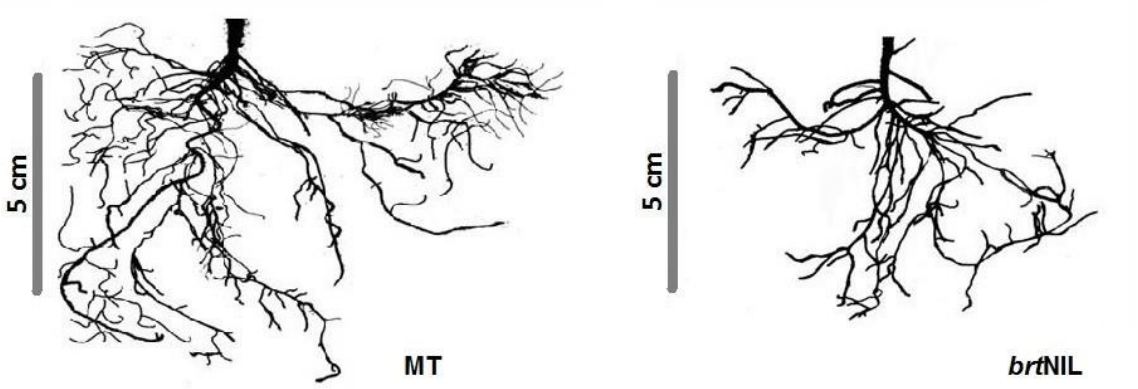


Figure 9: Root system architecture: Comparison of both phenotypes (plants at 25 days after germination), wild type (MT = Micro-Tom) and *bushy root*-NIL (*brt*NIL)

2.3.3 Seed area and maternal inheritance

As part of the phenotype characterisation, MT (WSS1414, n= 826), *brt*NIL (WSS1385, n= 2714) and F_1 (MT x *brt*NIL, n= 719 and *brt*NIL x MT, n= 788) seeds were analysed scanned and compared. The variation between both genotypes – *brt*NIL seeds are smaller (0.043±0.0002) than MT's (0.58±0.0003); is statistically significant ($p < 0.001$), but cannot be used as characterization feature, due to overlapping data (Table 3).

It was also determined that this trait is maternally inherited. It is known that *brt* vegetative phenotype shows a recessive Mendelian segregation (Zobel 1972; Zobel 1973; Voland 1987; Zobel 1991; Pino-Nunes & Peres 2005; Pino-Nunes & Peres 2009). Thus, if the seed size was controlled by the embryo genotype, a bimodal distribution might be expected, and 25% of the seed population will exhibit small seeds and 75% normal-sized seeds. However, this is not the case. The F_1 (MT♀ x *brt*NIL♂) seed area data shows a sharp peak similar to the MT seed population, differently from the outcome showed by F_1 (*brt*NIL♀ x MT♂) seed area (Figure 10).

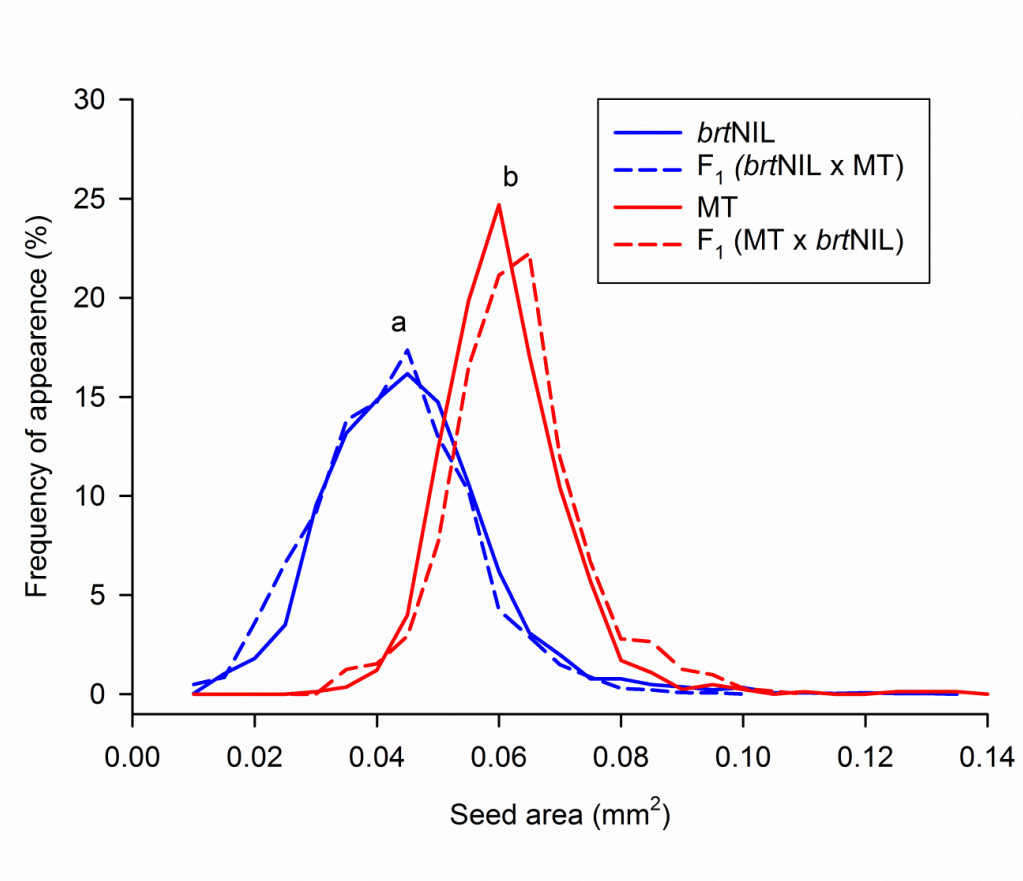


Figure 10: Seed area profile: Seeds from three different genotypes [*brt*NIL (WSS1385, n= 2714), MT (WSS1414, n= 826) and F_1 (MT x *brt*NIL, n= 719 and *brt*NIL x MT, n= 788)] were analysed compared. Statistical significance is represented by the letters ($p < 0.001$)

.

Table 3: Phenotype characterisation and comparison between Micro-Tom (MT) and *bushy root* (*brt*NIL) parental lines. Statistical differences are represented by letters. Errors are standard errors of the mean; populations used in each experiment per genotype are different and represented in the "N" column, except for the seed area, n = 2714 (*brt*NIL) and 826 (MT), 719 (F_1 MT x *brt*NIL) and 788 (F_1 *brt*NIL x MT) seeds. DAG = days after germination.

Trait	N	MT	*brt*NIL	*F1* *MT x* brt*NIL*	*F1* *brt*NIL *x MT*
Number of lateral roots at 35 DAG	30	95.5±3.3[a]	38.7±1.7[b]	n/d	n/d
Taproot length at 35 DAG (cm)	30	12.3±0.4[a]	6.8±0.2[b]	n/d	n/d
Total root length at 35 (cm) DAG	30	94.4±3.9[a]	51.9±2.6[b]	n/d	n/d
Root area at 34 DAG (cm^2), in perlite	10	127.9±6.7[a]	34.8±1.7[b]	n/d	n/d
Root area at 34 DAG (cm^2), in compost	10	108.4±5.3[a]	50.4±3.9[b]	n/d	n/d
Root depth at 34 DAG (cm), in perlite	10	43.6±1.9[a]	19.8±1.9[b]	n/d	n/d
Root depth at 34 DAG (cm), in compost	10	25.9±1.1[a]	10.9±0.6[b]	n/d	n/d
Root diameter at 34 DAG (cm), in perlite	10	7.6±0.5[a]	9.4±0.1[b]	n/d	n/d
Root diameter at 34 DAG (cm), in compost	10	6.1±0.1[a]	7.3±0.9[b]	n/d	n/d
Number of mature fruits (per plant)	10	21.1±0.6[a]	12.8±0.7[b]	n/d	n/d
Seed area (mm^2 $seed^{-1}$)	-	0.043±0.0002[a]	0.058±0.0003[b]	0.060±0.0003[b]	0.040±0.0004[a]

2.3.4 Resequencing of brtNIL and MT and delimiting the introgression

Initial genotyping experiments to find recombinants and delimit a target window were carried based on NGS data mapped to version SL2.30 of the reference genome. Results of those experiments placed the *brt* locus in a window of 54.4 Mbp, on chromosome 12. After the release of version SL2.50, *brt*NIL and MT were re-sequenced. As a result, from the Illumina sequencing of the genomic DNA, 132 million paired-end 126 bp reads for *brt*NIL.

To analyse both genomes individually, against the reference or against themselves, shared variants between *brt*NIL, MT and Heinz 1706 were excluded. Thus, the term "*unique"* is used to describe polymorphisms that are only present in one of the genotypes. After additional SNP filtration, by excluding polymorphisms with quality values under 210.8 and 112.8 (threshold based on the overall quality score and the validated polymorphism in this project) for SNPs and InDels, respectively, plots were created with 477,123 *brt*NIL and 147,244 MT unique SNPs (Figure 11) and with 81,437 *brt*NIL and 66,569 MT unique InDels (Figure 12).

The fact that *brt*NIL is a NIL of MT, has considerably reduced the amount of unique polymorphism. Very few SNPs were detected on chromosomes 3, 4, 6, 8, 9 and 10. In contrast to chromosomes 1, 2 and 12; where the number of SNPs in *brt*NIL is always higher than MT – although chromosomes 1 and 2 show peaks of SNPs, i.e. chr1: *brt*NIL 6,603 and MT 3,725 and chr2: *brt*NIL 4,871 and MT 1,010, the distribution pattern across the chromosome is the same. Differently than chromosome 12, location of the *BRT* locus, contains high peaks of SNPs across the whole chromosome. The *brt*NIL genome also shows more SNPs than MT across chromosome 5 and the top of chromosome 7. Interestingly, the only chromosome covered by more MT SNPs than *brt*NIL is chromosome 11, probably due to linkage of other loci with *mnt*.

The distribution of InDels follows the same pattern described above, with the exceptions of chromosomes 1 and 2, i.e. same distribution pattern, but the number of MT InDels overcomes the brtNIL variants – especially on

chromosome 1 where the peak of MT, 5,169 is much higher than *brt*NIL, 2098. However, Kevei et al. (2015) have also shown the same results, which would suggest that this region on chr1 is probably an introgression in Heinz. An inversion of the distribution pattern of SNPs was detected across chromosome 11, showing peaks of *brt*NIL 1,926 and MT 1,583.

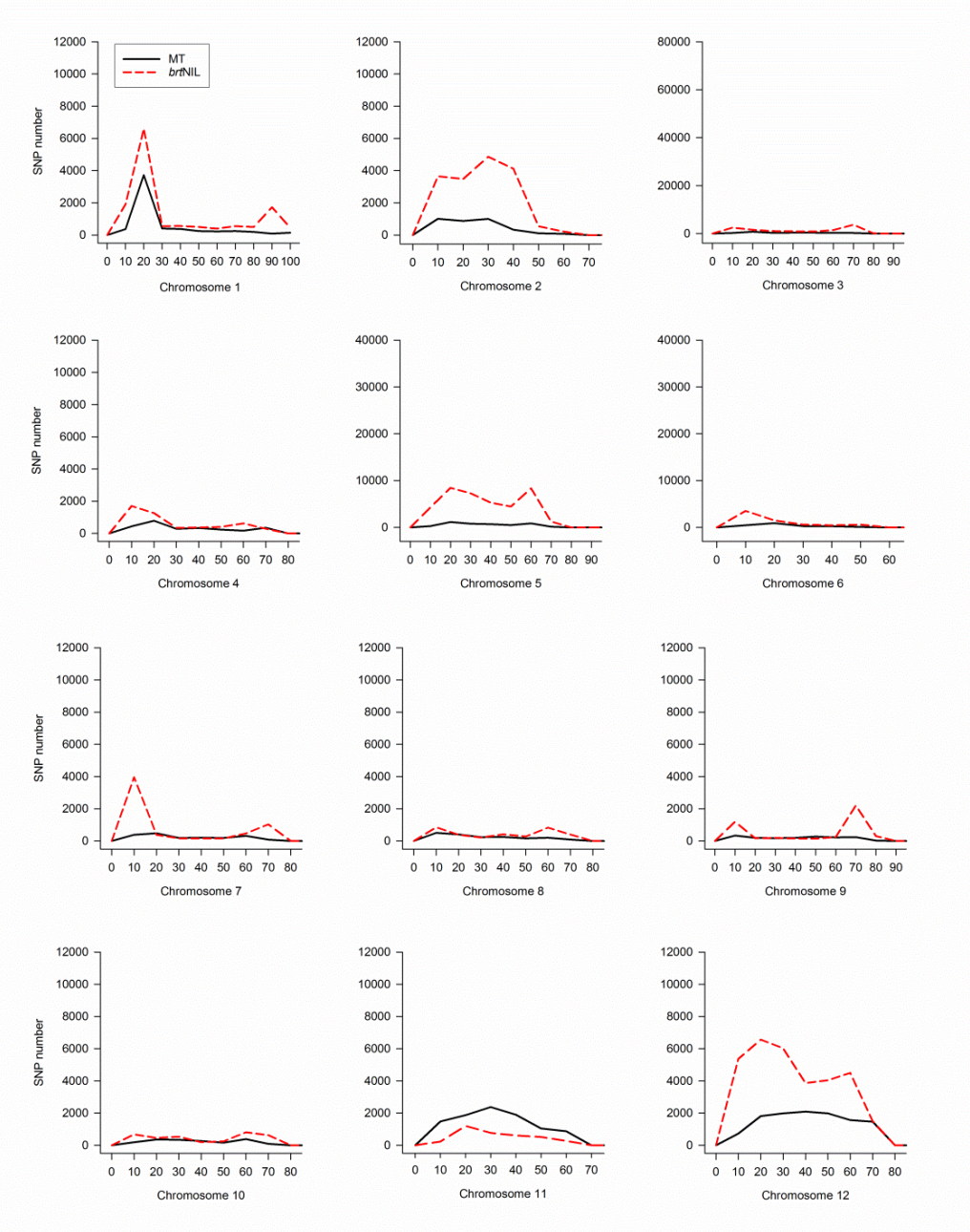


Figure 11: Genome-wide SNPs uniquely discovered in *bushy root* (*brt*NIL) and MT. Unique = only present in one of the genotypes. Both lines were compared to the genome reference (*S. lycopersicum* cv. Heinz 1706). Polymorphisms with quality lower

than 210.8 (quality index) or shared were excluded from the analysis presented above (unique - SNPs plotted on *brt*NIL and MT are 477,123 and 147,244, respectively); this threshold was based on validated SNPs on this project.

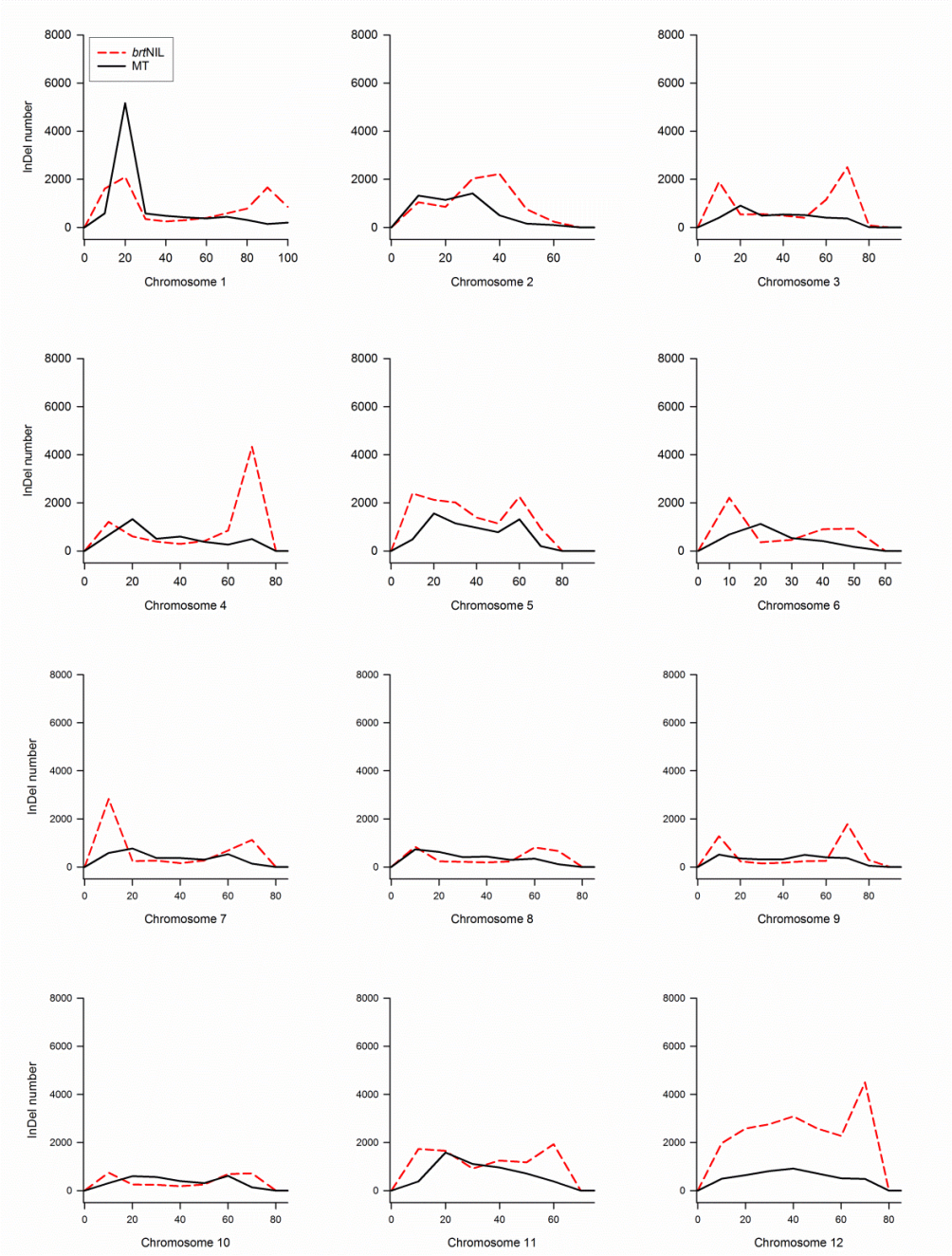


Figure 12: Genome-wide unique InDels discovered in *bushy root* (*brt*NIL) and MT. Unique = only present in one of the genotypes. Both lines were compared to the genome reference (*S. lycopersicum* cv. Heinz 1706). Polymorphisms with quality lower than 112.8 (quality index) or shared were excluded from the analysis presented above (InDel plotted on *brt*NIL and MT are 81,437 and 66,569, respectively).

2.3.5 Recombinants selection and mapping interval

After analysis of the NGS data, the introgression was again delimited (to 92.8% of the chromosome – figure 13) – which means that many SNPs were not discovered during the preliminary analysis (SL2.30). Molecular markers (*Alfa*, *Beta* and *Gama*; Table 4) were used to genotype a population of ~1,000 F_2 plants with the objective of breaking the introgression into smaller bins and define the target region. Genotyping results reduced the introgression to 4.86% (*Gamma 4;* 4,192,601 - *Gamma 11;* 7,025,456) of its original size.

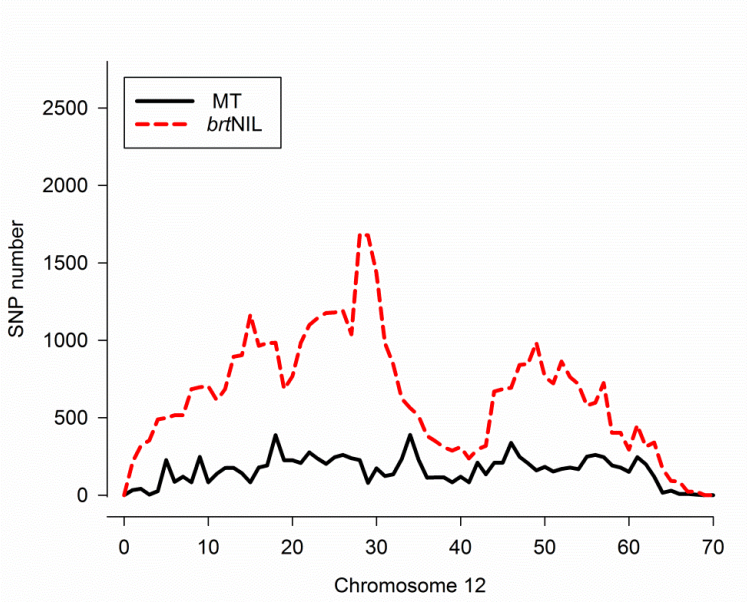
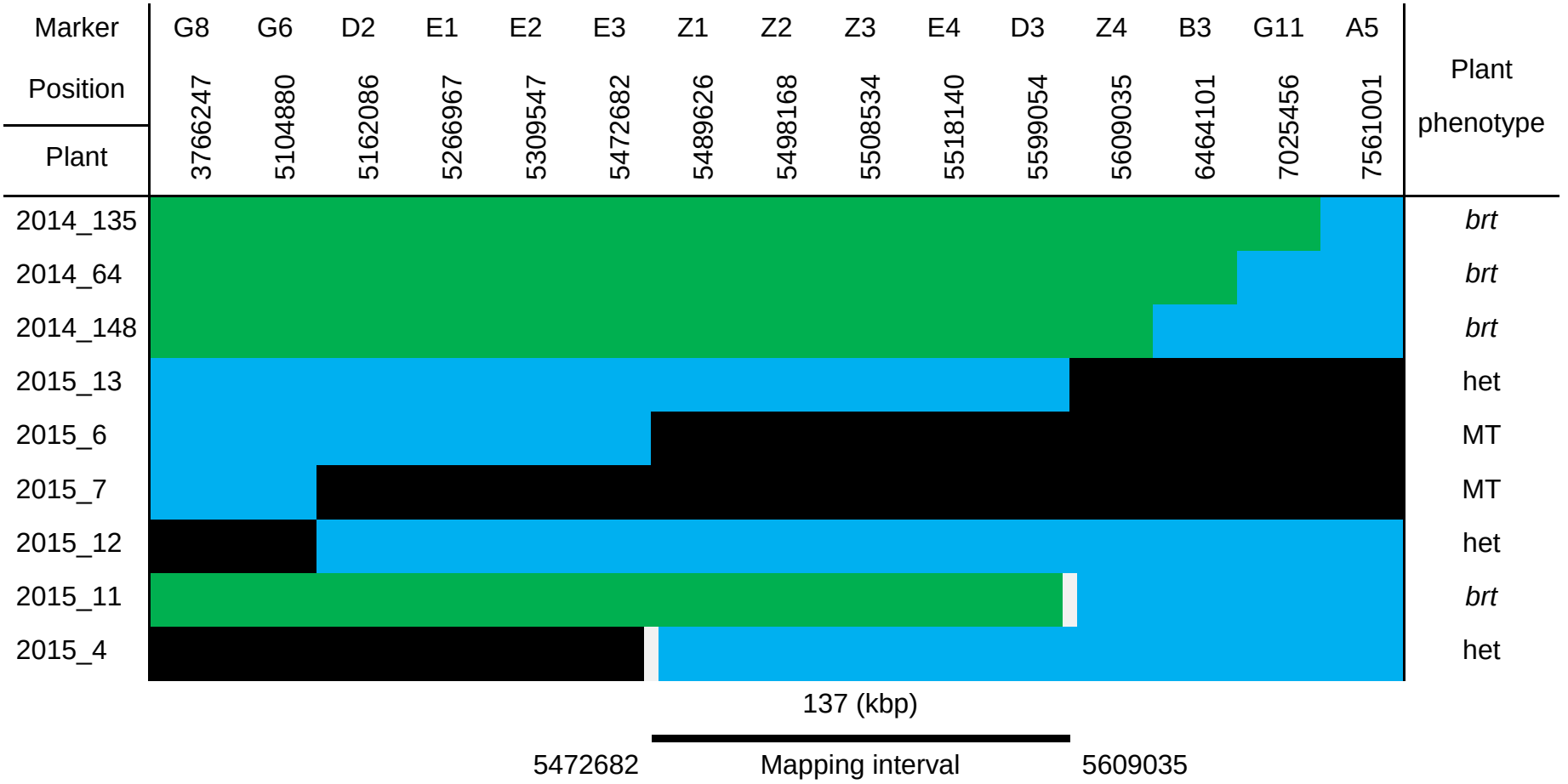


Figure 13: SNP discovery in *bushy root* (*brt*NIL) and Micro-Tom (MT) across chromosome 12. Both lines were compared to the reference genome (i.e. *S. lycopersicum* cv. Heinz 1706); only unique SNPs are shown (*brt*NIL n =477,123 and MT n = 147,244).

A larger population (n = 4,000) of F_2 plants was screened using *Gamma 4* and *Gamma 11*, to detect new recombinants and increase the mapping resolution in the target region. Thirteen recombinants were found and later tested with the other marker sets (*Epsilon* and *Zeta)*. The current search window (137 kbp – *Epsilon* 3; 5,472,682 - *Delta* 3; 5,599,054) contains 9 gene candidates.

Table 4: Recombinants genotyping and representation. Results were shown using representative recombinants and colour code (*brt*. *brt* = green; *brt*.+ = blue and +.+ = black) – Molecular marker are symbolized by the first letter and its positions are represented in bp. The phenotype was confirmed on the next generation (het plants).



2.3.6 SNP data from resequenced lines in fine-mapping

The VCF files published by other research groups (Aflitos et al. 2014; Lin et al. 2014) containing polymorphisms were used to exclude common variants, i.e. unlikely to be responsible for the *brt*NIL phenotype. The small mapping interval has 69 (64 SNPs and 5 InDels) polymorphisms – 30 (SNPs) and 4 (InDels) which are unique to the *brt*NIL genome, compared the other genomes (includes MT) – Table 5.

Table 5: Polymorphisms at the mapping interval. Amino-acid polarity changes are in bold. Base changes are separated by a slash. Annotation = upstream gene variant (UGV), downstream gene variant (DGV), missense variant (MV) and synonymous variant (SV).

Polymorphism	Position	Change	Annotation	Gene
SNP	5481593	T/C	UGV	Solyc12g014510
SNP	5485750	A/G	UGV	Solyc12g014510
SNP	**5485798**	**Q11G**	**MV**	**Solyc12g014510**
SNP	5485815	L17V	MV	Solyc12g014510
SNP	**5488162**	**T567R**	**MV**	**Solyc12g014510**
SNP	5488258	L599L	SV	Solyc12g014510
SNP	5488265	C601T	MV	Solyc12g014510
SNP	5488270	N603T	MV	Solyc12g014510
SNP	5489045	T/A	DGV	Solyc12g014510
SNP	5489309	C/T	DGV	Solyc12g014510
SNP	5489338	G/A	DGV	Solyc12g014510
SNP	5489357	T/G	DGV	Solyc12g014510
SNP	5489382	C/A	DGV	Solyc12g014510
SNP	5489388	G/A	DGV	Solyc12g014510
SNP	5489390	AC	DGV	Solyc12g014510
SNP	5486562	C/A	DGV	Solyc12g014520

SNP	**5489626**	**P1060Q**	**MV**	**Solyc12g014520**
SNP	**5489640**	**S1055R**	**MV**	**Solyc12g014520**
SNP	**5492482**	**S403P**	**MV**	**Solyc12g014520**
SNP	**5492587**	**S368P**	**MV**	**Solyc12g014520**
SNP	5494475	L.L	SV	Solyc12g014520
SNP	5494606	T/A	UGV	Solyc12g014520
SNP	5494614	G/A	UGV	Solyc12g014520
SNP	5494699	G/T	UGV	Solyc12g014520
SNP	5494861	G/C	UGV	Solyc12g014520
SNP	5494143	A/C	UGV	Solyc12g014530
SNP	5494162	C/T	UGV	Solyc12g014530
SNP	5508030	T/A	UGV	Solyc12g014530
SNP	5556051	T/C	DGV	Solyc12g014550
SNP	5556309	C/T	DGV	Solyc12g014550
SNP	5581778	T/A	UGV	Solyc12g014580
InDel	5582699	5582699^5582700insT	UGV	Solyc12g014590

2.3.7 Transcriptomics

Differentially expressed genes were identified using a threshold of 0.25 logFC and a FDR of 0.05 against possible contrasts (Table 6 and Appendix B). In the mapping interval, three genes (*Solyc12g014530*, *Solyc12g014540* and *Solyc12g014590*) showed a significantly different expression pattern. Further analysis showed that the only gene which showed differential expression between WT and *brt*NIL in both leaves and roots was the pirin gene (*Solyc12g014590*) – data not shown.

2.3.8 Candidate gene

After the exclusion of common polymorphisms presenting in several genomes, 6 genes – *Solyc12g014510, Solyc12g014520, Solyc12g014530,*

Solyc12g014550 and *Solyc12g014580, Solyc12g014590*; were left as possible candidates (Table 6). The excessive number of polymorphisms in a possible Mutator-like element (MULE) – two of the three candidates, and the transcriptomics results suggest that only *Solyc12g014590* may be the gene responsible for the *brt* phenotype (Table 7).

Table 6: Gene expression profile in *bushy root* (*brt*NIL) and Micro-Tom (MT)

Contrast	Overexpressed genes	Under expressed genes	Total
*brt*NIL (root) vs. MT (root)	3786	4440	8226
*brt*NIL (leaf) vs. MT (leaf)	117	123	240

2.4 DISCUSSION

As described, the mutant phenotype is a combination of pleiotropic effects caused by the *brt* gene. The phenotype includes smaller seeds, which is maternally inherited, a phenomenon that has been reviewed by many (Dolan 1984; Koelewijn & Van Damme 2005; Chaudhury & Berger 2001; Wolf & Wade 2009). Maternal effects "*are the influences of the maternal genotype or phenotype on the offspring phenotype*" (Wolf & Wade 2009); all the offspring, which are exposed to the same conditions on the mother plant, develop a phenotype that is dependent on the maternal genotype. It was suggested that maternal effects can be due to different factors, e.g. fertilization, the number of competing ovules per fruit and cytoplasmic and nuclear effects (Dolan 1984; Roach & Wulff 1987). The phytohormone cytokinin has been reported as a principle regulator in seed development (Bartrina et al. 2011; Swartzberg et al. 2011; Li et al. 2013; Li & Li 2016), so it could be hypothesized that the difference in seed size in *brt*NIL and MT might be due to the lower cytokinin sensitivity previously described by Pino-Nunes & Peres (2005).

Table 7: Candidate genes for the *brt*NIL phenotype. Unique polymorphisms = only present in *brt*NIL genome compared against the other resequenced lines.

Gene	Position	Size (bp)	Description	Unique Polymorphisms
Solyc12g014510	5485767 - 5488524	2757	Mutator-like transposase	Exon; Introns
Solyc12g014520	5489565 - 5494483	4918	Ulp1 protease family C-terminal catalytic domain containing protein	Promoter; Exons and Introns
Solyc12g014530	5498306 - 5503348	5042	L-aspartate oxidase	Intergenic
Solyc12g014540	5523526 - 5533996	10470	Lysine-specific demethylase 5C	None
Solyc12g014550	5553309 - 5554724	1415	Syntaxin-52	Intergenic
Solyc12g014560	5559684 - 5566991	7307	Cell division protein ftsZ	None
Solyc12g014570	5569788 - 5576277	6489	Glycerophosphoryl diester phosphodiesterase family protein	None
Solyc12g014580	5577417 - 5577847	430	Pollen allergen Ole e 6	None
Solyc12g014590	5578789 - 5581766	2977	Pirin	Promoter

The mutant smaller seeds can be associated with the results obtained from taproot experiment, *brt*NIL plants show reduced taproot length. The amount of endosperm available for the embryo is related with the seed size (Li & Li 2016). Therefore, the small seed size could slow down the initial development of *brt*NIL in comparison to MT due to limitation of seed reserve.

The available description of the *brt* root phenotype, highly branched root system, (Zobel 1972; Zobel 1973; Voland 1987; Zobel 1991) is very different than what was observed on *brt*NIL. The complex genetic background of the original cross might be responsible for such differences. The wording used was also challenging, once there was no formal botanical definition of what is considered "*bushy*" – Zobel (1991) suggests that the increased number of lateral roots could be responsible for the bushy phenotype in *brt*. However, the total number of lateral roots in the brtNIL was measured (Table 3), and found to be significantly fewer than in MT ($p < 0.005$), which would make the mutant less bushy if the definition was simply an increase in lateral roots.

However, many aspects of the root architecture were analysed and our results based on different models revelled that the most reliable and consistent way to describe bushiness must consider different parameters of the root system. Several models were applied to our data and bushiness should be defined as the "total root length" x "number of lateral roots" / "taproot length"; as these traits would represent a special and geometrical exploitation of the soil by the root system. In addition, the data available from the *brt* phenotype was collected by visual inspection by Zobel. Therefore, the data presented here are the only quantified and statistically analysed results for this mutant. The *brt* mutation does not make the root system bushier in the MT genetic background, but is does make the root system smaller and more compact.

Genotyping experiments confirmed that the *BRT* locus is indeed on the 12^{th} chromosome, as previously reported (Zobel 1972). The introgression in *brt*NIL represents 92.8% of chr 12, however, through fine-mapping a 137 kbp mapping interval was defined. From the nine genes located in this interval, six

of them (Solyc12g014510, Solyc12g014520, Solyc12g014530, Solyc12g014550 and Solyc12g014580, Solyc12g014590) contain unique polymorphisms compared to the 452 re-sequenced lines available (Aflitos et al. 2014; Lin et al. 2014).

The *Solyc12g014510* gene annotated as a transposon, which would make it a weak candidate for *brt*NIL. In addition, the excess of unique and shared variants in the exons, introns and promoters of the *Solyc12g014510* suggesting that it has a neutral function (Robbins et al. 2011). The lack of polymorphisms in other genes, *Solyc12g014540, Solyc12g014560* and *Solyc12g014570*, excluded them as likely candidates; in addition to the most probably non-causative, intergenic SNPs in other three candidates, i.e. *Solyc12g014530, Solyc12g014550* and *Solyc12g014580.*

Thus, the strongest candidates in the mapping interval are *Solyc12g014520* and *Solyc12g014590*; both genes are differently expressed in MT and *brt*NIL roots. The *Ulp1* gene (*Solyc12g014520*) mutant of *Arabidopsis* is reported to have severe changes in phenotype (i.e. reduced plant size and seed production, and slow development, (Murtas et al. 2003; García-Lorenzo et al. 2006; van Leeuwen et al. 2007; Hermkes et al. 2011). In addition to the similar phenotype found in the *Arabidopsis Ulp1* mutant, unique non-conservative amino-acid changes were also detected in the *brt*NIL line. However, the high number of polymorphisms in introns, exons and promoter region of this gene suggests it is not under any selection pressure, and may not encode a functional gene product. Furthermore, Ulp1 homologues are reported to be located next to transposons, which can suggest that Ulp1 is be part of a MULE, a super family of transposons (van Leeuwen et al. 2007).

EMS is usually responsible for very few alterations across the genome (Mohd-Yusoff et al. 2015), so is very unlikely to have been responsible for the large number of polymorphisms observed. These combinations of factors weaken the position of the *Ulp1* gene as viable candidate for the BRT locus.

Pirins are conserved bicupins proteins reported to be involved with a series of biological processes – e.g. apoptosis and abscission (Orzaez et al.

2001; Bar-Dror et al. 2011), germination (Lapik & Kaufman 2003; Orozco-Nunnelly et al. 2014), catabolism (Soo et al. 2007) and quercetinase activity (Tranchimand et al. 2010; Widiatningrum et al. 2015) or act as co-transcriptional factors. A single base insertion was detected in the promoter region of the pirin gene (Solyc12g014590) – Table 7. New RNA-seq data available at the new release of version SL3.0 shows expression in the promoter of the pirin in gene (SGN 2017); which suggests that an insertion in the promoter of *Solyc12g014590* can, indeed, be responsible for the *brt*NIL phenotype.

In conclusion, the combination of results shown here suggests that the mutation in the promoter region of the pirin gene is the best candidate in the nine-gene mapping interval to be responsible for the *brt*NIL compact root system and small seed phenotype. These findings provide a different perspective on the role of the pirin gene in plant development, as it that a possible interaction with cytokinin, due to *brt*NIL lower CK sensitivity. To confirm *Solyc12g014590* as responsible for the *brt* phenotype, transgenic experiments are required to increase or decrease the expression of the Solyc12g014590 gene, or to knock it out using CRISPR/Cas-9 genome editing.

2.5 ACKNOWLEDGEMENTS

DSF was supported by a Science without Borders PhD scholarship (246783/2012-5), funded by the Brazilian Government and the Brazilian National Council of Research and Development (CNPq). TK was supported by BBSRC grant BB/L011611/1; ZK was supported by BBSRC grant BB/L01954X/1. We are grateful to the Earlham Institute, Norwich for sequencing services. We thank the Tomato Genetics Resource Centre for providing the *brt* seed accessions to AJT and LEPP. We are also thankful for the previous work carried by Prof Richard Zobel (USA), Dr Lilian Pino-Nunes (Brazil) and Pauline Amiel (France), which helped us to reach the conclusions presented in this paper.

3 IDENTIFICATION AND CHARACTERISATION OF *BIFURICATE*, A NOVEL LOCUS ON CHROMOSOME 12 CONTROLLING TRUSS BRANCHING AND FLOWER NUMBER IN TOMATO

Silva Ferreira, D. K.[1]; Kevei, Z. L.[1]; Kurowski, T.[1]; De Noronha Fonseca, M.E.[2], Boiteux, L.S.[2], Mohareb, F.[1] and Thompson, A. J.[1]

[1]Cranfield Soil and AgriFood Institute, College road, Cranfield University, MK43 0AL, UK

[2]National Center for Vegetable Crops Research, CNPH – EMBRAPA Hortaliças, Brasília-DF, Brazil

Correspondence: Andrew Thompson. E-mail: a.j.thompson@cranfield.ac.uk

Highlight

At a newly discovered locus in tomato, *bifuricate*, a genomic segment introgressed from the wild species *Solanum galapagense*, increased inflorescence branching and flower number and contained two candidate genes, a MAP kinase and a BTB/POZ transcription factor.

Abstract

A naturally occurring recessive mutant named *bifuricate* (*bif*) was recovered from a single seed descent program intended to produce high-yield mini-plum tomato breeding lines. Plants with the *bif* phenotype showed an 24-fold increase in inflorescence (truss) branching in comparison to the wild type (WT) control line LAM183, which presented a mean of only 0.16 branches per truss. In addition, the number of flowers per truss was 235% higher in *bif* plants than WT, and a four-day low temperature treatment stimulated branching significantly more in *bif* than in LAM183. The genomes of the LAM183 and *bif* lines were resequenced and single nucleotide polymorphism markers were designed for genetic mapping in an F_2 population. The *BIF* gene was mapped to a 2.01 Mbp interval of chromosome 12 containing 53 genes. All coding region polymorphisms in the interval were surveyed and five candidate genes displayed altered protein sequences, of which two genes contained stop codons predicted to disrupt gene function. Both of these genes, *Solyc12g019420* (a BTB/TAZ transcription factor) and *Solyc12g019460* (a MAP kinase), are excellent candidates for inflorescence branching control. An introgression browser was used to demonstrate that the origin of this mutant haplotype was *Solanum galapagense*.

Key words

Branching; BTB-POZ BTB-TAZ; genome resequencing; inflorescence architecture; low temperature; MAP kinase; *Solanum galapagense*; *Solanum lycopersicum*.

3.1 INTRODUCTION

The flowering process and the number of flowers produced by a plant is controlled by a network of approximately 300 genes (Lemmon et al. 2016). An increase in the number of flowers will lead to a greater final fruit yield if supply of assimilate is abundant and reproductive growth is limited by sink strength (Périlleux et al. 2014). Conversely, flower development represents a considerable energy investment (Lemoine et al., 2013), and thus production of more flowers than can be sustained by the assimilate supply is a waste of resources and it may negatively affect the final fruit yield. In situations where assimilate supply is limiting, fruit number is inversely proportional to fruit size, and can be regulated by flower and fruit abscission in response to endogenous and environmental signals (Saglam & Yazgan 1999; Hidayatullah et al. 2008). Growers also manage fruit crops by thinning and pruning trusses to limit fruit number and to achieve larger, more uniform fruits (Cockshull & Ho 1995; Max et al. 2016).

Peduncle length, branching and the number of flowers per unit length of peduncle determines inflorescence architecture and the subsequent number of fruits that can be produced. It also has the potential to influence the uniformity of ripening that begins with proximal fruits and progresses to more distal fruits (Giovannoni 2001). Fruits borne on a highly branched structure will tend to exhibit a higher degree of synchronicity in their ripening, which may be advantageous for commercial fruit production (Ecker 1995; Moxon et al. 2008).

Flower initiation and development in tomato have been well characterized (Lippman et al. 2008), but the genetic mechanisms involved in controlling truss architecture are poorly understood. In the case of *Solanum lycopersicum* L. (tomato), the first inflorescence (truss) is initiated typically after production of 8 to 12 leaves (Lifschitz & Eshed 2006). After the appropriated flowering induction stimulus, the shoot apical meristem (SAM) transitions to the inflorescence meristem (IM), which develops floral meristems (FMs) and ultimately flowers (Lippman et al. 2008; Lozano et al. 2009). Truss architecture is extremely plastic and responsive to environmental factors. Several studies (reviewed by

Gratani 2014) have reported variations in truss architecture in response to external signals (e.g. lower temperatures increase branching and number of flowers – Calvert 1957; Calvert 1959).

In monopodial plants (e.g. *Arabidopsis*) the SAM is indeterminate and the reproductive structures are located on the sides (flanks). Sympodial plants (e.g. tomato), on the other hand, have a determinate SAM and the primary shoot is terminated with an inflorescence; subsequently a new vegetative cycle is initiated, progressing as the new primary shoot – resulting in three new leaves (vegetative nodes) and then again termination of the shoot in a new inflorescence, by the sympodial meristem (Schmitz & Theres 1999; Carmel-Goren et al. 2003; Quinet et al. 2006; Castel et al. 2010; Thouet et al. 2012). This pattern is repeated, forming consecutive sympodial segments which together constitute a sympodial shoot (Samach & Lotan 2007; Kirchoff & Claßen-Bockhoff 2013; Park et al. 2014; Astola et al. 2014).

Although the tomato and *Arabidopsis* model plants have distinct growth and flowering patterns, they share a number of orthologous genes controlling inflorescence architecture. In *Arabidopsis*, four key genes related to meristem identity have been studied in relation to the genetic control of inflorescence architecture: *TERMINAL FLOWER1* (*TFL1*), *APETALA1* (*AP1*), *UNUSUAL FLOWER ORGAN* (*UFO*) and *LEAFY* (*LFY*) (reviewed by: Weigel et al. 1992; Bradley et al. 1997; Kardailsky et al. 1999; Kobayashi et al. 1999; Hanano & Goto 2011; Chandler 2014). *TFL1* is responsible for early flowering after the development of rosette leaves; it delays the transition of IM to FM producing a terminal flower. *AP1* is upregulated in FM before floral stage 1 and it negatively regulates *TFL1* and controls FM initiation. The *UFO* gene is reported to regulate class B functions in the formation of petals and stamens (Samach et al. 1999), growth of floral primordium, and can regulate meristem identity by transforming FM back to IM (Levin & Meyerowitz 1995). UFO is also responsible for co-activating the *LFY* gene (Souer et al. 2008), which promotes floral fate by establishing and regulating floral identity (Kobayashi et al. 1999; Yang et al. 1999).

In tomato seven genes are known to be involved in the regulation of inflorescence architecture (Astola et al. 2014): in *FALSIFLORA* (*FA*), the tomato orthologue of *LFY*, the IM fails to assume floral identify, remaining in an intermediate stage between vegetative and reproductive states (Allen & Sussex 1996; Molinero-Rosales et al. 1999; Lozano et al. 2009); in *ANANTHA* (*AN*), the tomato orthologue of *UFO* (Souer et al. 2008) the IM propagates indefinitely, producing large inflorescences with immature flowers resembling cauliflower arrested inflorescences (Allen & Sussex 1996; Molinero-Rosales et al. 1999; Lozano et al. 2009); in *JOINTLESS* (*j*) the FM is produced, but, after 3-4 flowers are formed, the IM is converted to a vegetative meristem (VM) (Szymkowiak & Irish 1999; Mao et al. 2000); *BLIND* (*bl*) regulates the inflorescence architecture by controlling the formation of lateral meristems (Schmitz et al. 2002); the *UNIFLORA* (*UF*) gene controls the floral transition (Mero & Honma 1982; Dielen et al. 2004; Quinet et al. 2011); the *TERMINAL FLOWER* (*TMF*) gene upregulates vegetative growth by supressing *FA* expression (Périlleux et al. 2014) and early inflorescences have only a single flower in *tmf* mutants (Xu et al. 2016). Of the seven known genes that regulate inflorescence architecture, only the C*OMPOUND INFLORESCENCE* (*s*) mutant gene increases the number of flowers by increasing the number of peduncle branching points. The *s* gene encodes a transcription factor related to a gene called WESCHEL HOMEOBOX and it was mapped to chromosome 2L (*Solyc02g077390* – 42,337,50 bp, SL2.50 – Lippman et al. 2008; SGN 2017). Lippman et al. (2008) evaluated 6,000 tomato lines, from which 23 showed the highly-branched truss phenotype. Of those, 22 were allelic to *s* and had the same causative mutation (*s-classic* with amino acid substitution G82D), whereas a single line CC5721 showed the highly-branched truss phenotype but contained the wildtype s^+ allele. The line CC5721 therefore apparently had a different and unknown genetic cause for the phenotype (Lippman et al. 2008; Krieger et al. 2010). A different mutant allele of the *s* is also reported as *s-MULTIFLORA* (amino acid substitution I94F); it shows an intermediate phenotype between *an* and *s-classic,* having a high number of flowers/branching points and arrested flower development resembling the cauliflower head tissue. In the present paper, we

describe a new tomato locus, *BIFURICATE* (*BIF*), which causes a phenotype similar to *s* as well as the effect of a cold treatment on its phenotypic expression.

3.2 MATERIAL AND METHODS

3.2.1 Plant material

Seeds of LAM183 and *bifuricate* (*bif*) inbred lines originated from the tomato breeding program at CNPH, Brasília-DF, Brazil. After one more cycle of self-pollination carried out under glasshouse conditions in the UK (Cranfield University), plants of *bif* and LAM183 that clearly showed high or low branching, respectively, were selected and self-pollinated one more time. It was then confirmed that the contrasting branching phenotypes were inherited in all their progenies. All subsequent phenotyping experiments were conducted using the progeny from these representative stable lines, and they were also used to produce *bif* × LAM183 F_1 and F_2 seeds for genetic mapping (Figure 14).

3.2.2 Plant growth

To avoid potential pathogen transmission, all seeds were sterilized by immersion for 10 minutes in a solution of 50% v/v Domestos (Unilever, UK) commercial bleach, equivalent to 2.5% w/v hypochlorite, and then thoroughly washed with tap water to remove residual bleach. After sterilization, seeds were placed in 9 cm Petri dishes containing two layers of Whatman No. 1 filter paper (85 mm diameter) soaked with 3 ml of tap water. Petri dishes were placed in the dark at 25°C for 3 days, and the chitted seeds were sown into 3 L round pots, 15 cm diameter x 18 cm height, in a glasshouse in peat-based compost (Sinclair multi-purpose compost, LBS Horticulture Ltd, UK). Glasshouse temperature set points were 23/18ºC, 14/10 hours, day/night, with light provided by high pressure sodium lamps. Pots were irrigated according to demand, and were fed twice a week with Hoagland solution (5 mM K_2SO_4; 1 mM H_3PO_4; 5

mM $Ca(NO_3)_2$; 2 mM $MgSO_4$; 100 µM EDTA Fe-Na; 42.2 µM H_3BO_3; 9.1 µM $MnCl_2$; 0.76 µM $ZnSO_4$ and 0.32 µM $CuSO_4$. pH:5.8, adjusted using H_3PO_4) at half concentration before flowering and full concentration after flowering.

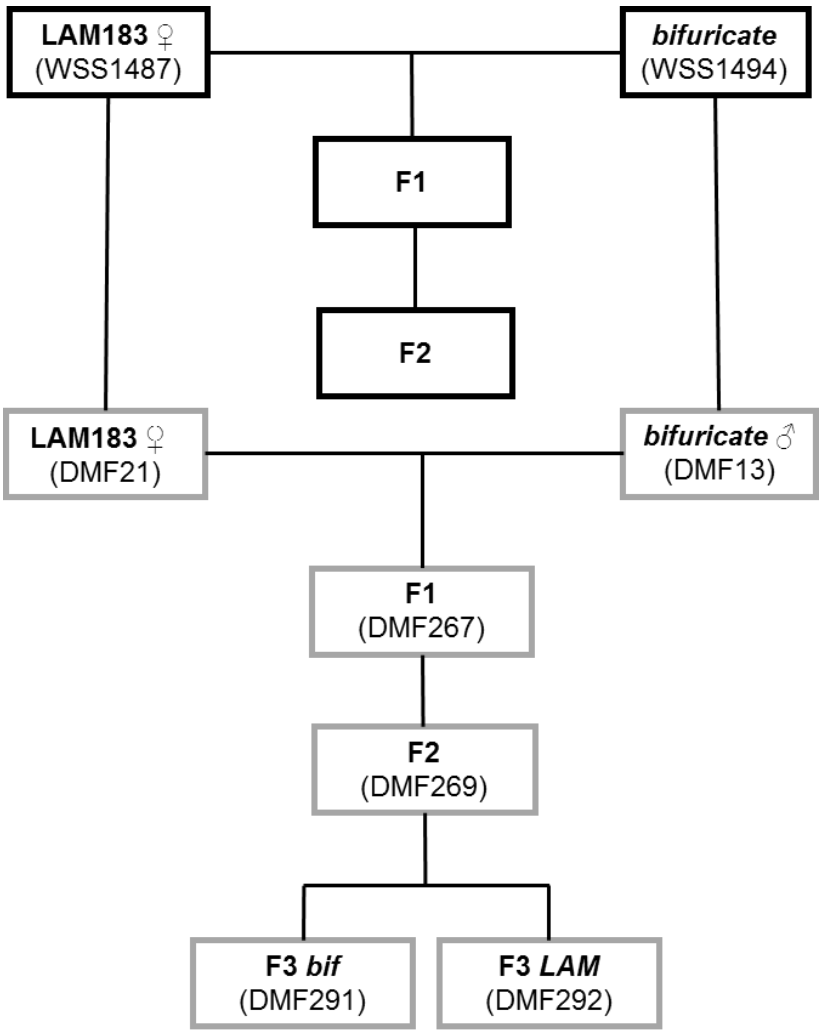


Figure 14: Seeds pedigree. Seeds produced in Brasilia (Brazil) are represented by the black outline, while seeds produced in Cranfield (UK) are in grey. F_2 seeds used on fine mapping were from seeds package DMF269.

3.2.3 Cold Experiment

Sixty chitted seeds from each inbred line (*bif:* DMF21 and LAM183: DMF13) were potted in the glasshouse in three randomized blocks (n = 20 plants per genotype per block). After the transfer, the seedlings were left one week in the glasshouse (set point 23ºC) for initial establishment. Groups of 24 plants (6

replicates x 2 genotypes x 2 temperatures) were transferred at five different stages of development to two illuminated growth cabinets (208 µmol m^{-2} s^{-1} photosynthetic photon flux) generated by cool white fluorescent bulbs (Panasonic, Bracknell, UK) set to provide a cold treatment (15ºC) and a control treatment (23ºC) (see Figure 15 for treatment regime). After 4 days at the differential temperature, the plants were transferred back to the glasshouse, maintaining the randomised block design. After the last transfer (i.e. week 6), all plants were grown in the glasshouse at a minimum temperature of 23ºC until the third truss was formed and the two first were mature enough to be scored.

3.2.4 DNA extraction, NGS genomic data generation and variant calling

Genomic DNA from LAM183 and *bif* plants was extracted using the DNeasy plant mini kit (Qiagen; Manchester, UK), according to the manufacturer's instructions. One lane was used to sequence both genomes, using 126 bp paired-end reads by Illumina HiSeq 2500. After the quality control performed by FastQC (Schmieder & Edwards 2011) , to ensure that the reads used had high quality (on average above 33); the raw data were mapped to the reference genome SL2.50 and genome annotation ITAG2.4, and analysed on Cranfield University's GRID, i.e. 32 CPUs 8GB/core, running a Linux environment on the Cranfield University GRID. Firstly, the reads were aligned by the Burrows-Wheeler aligner (BWA, version 0.7.4), using default specifications; secondly, the aligned reads were compressed into a binary (bam) format (Picard tools) and then sorted and indexed by Samtools (version 0.1.19); thirdly, the GATK package (Genome Analysis Tool Kit, Broad Institute, Cambridge, USA, version 3.3.0) was used to realign the Insertions and Deletions (InDels) and variant calling (HaplotypeCaller, using default settings). This pipeline produced variant call format (VCF) files. The annotation of the VCF files was completed by

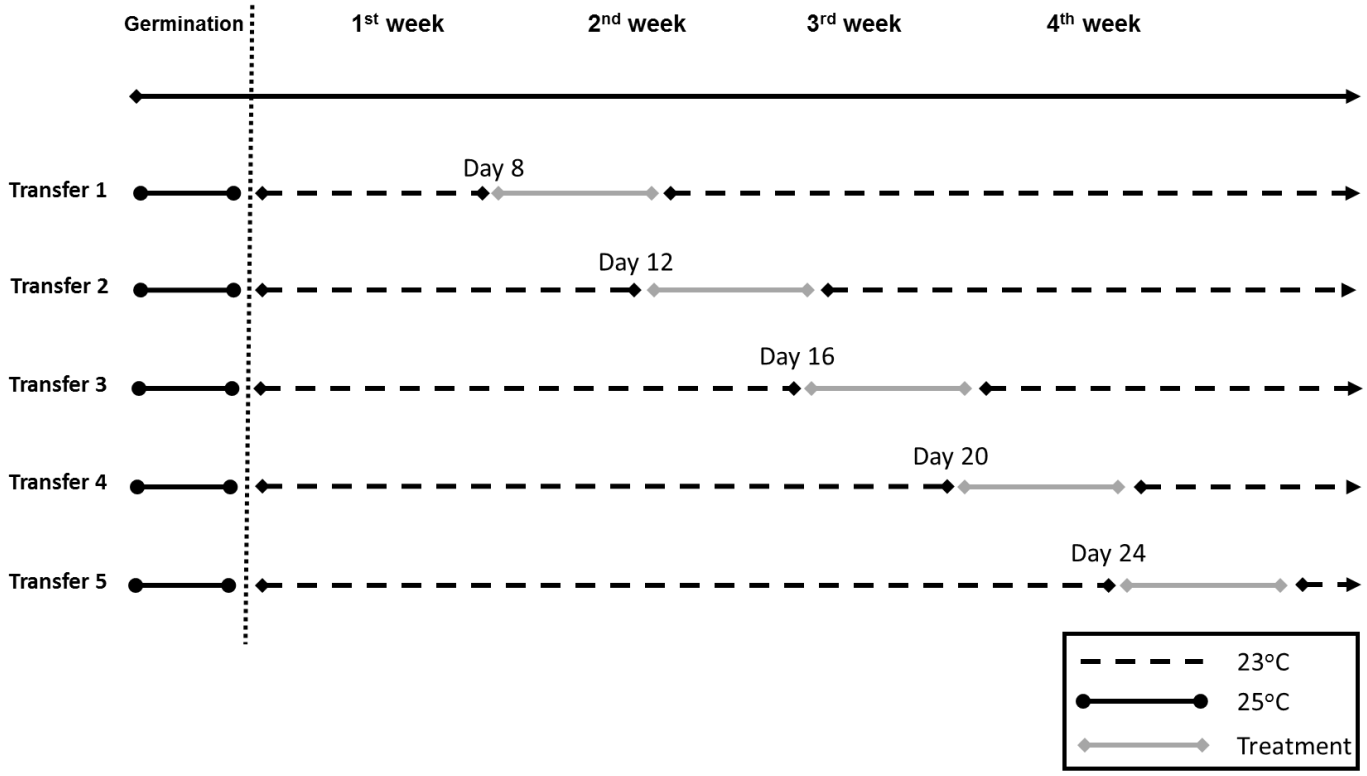


Figure 15: Transfer points. Plants were treated at different temperatures for four days – see Materials and Methods.

SnpEff (version 4.0) using ITAG2.40 (associated with tomato genome reference version SL2.50) available on the Sol website (http://solgenomics.net/). Finally, the variants were filtered using GATK's *variant filtration* tool (Quality Depth < 2, Fisher Strand > 60, Mapping Quality < 40, Haplotype Score > 13 and Mapping Quality RankSum < 12.5) (Kevei et al. 2015). Unique variants were filtered using a custom bash unpublished script which excluded polymorphisms shared between the data sets, similar to the mechanism of bedtools (Quinlan & Hall 2010). The filtered VCF file was uploaded to the GenoVerse Genome Browser (Bragin et al. 2012) which can be accessed on: http://elvis.misc.cranfield.ac.uk/GenoverseBifuricate.

3.2.5 SolCaps genotyping

Using the Genoverse genome browser, 48 SolCap markers (Scott 2010; Sacco et al. 2015; Sim et al. 2015; Kim et al. 2016) observed to be polymorphic between LAM183 and *bifuricate* (2 at the top; 2 at the bottom of each chromosome – avoiding heterochromatin) were manually selected. DNA extraction from leaf tissue of individual plants of the LAM183 × *bifuricate* F_2 population (n = 96) and Kompetitive Allele Specific PCR (KASP) genotyping of the 48 SolCap markers (Table 8) was performed by LGC (Teddington, UK).

Table 8: SolCap markers used for genotyping

Marker ID	SolCap Reference	Chromosome	Position	SNP
DSF1	solcap_snp_sl_59771	1	6229659	T/C
DSF2	solcap_snp_sl_16925	1	18519338	T/G
DSF3	solcap_snp_sl_19068	1	43255975	T/C
DSF4	solcap_snp_sl_22423	1	59436302	C/G
DSF5	solcap_snp_sl_17075	1	82722469	A/G
DSF6	solcap_snp_sl_17448	2	15153844	A/T

DSF7	solcap_snp_sl_6255	2	20334143	A/G
DSF8	solcap_snp_sl_13842	2	29361519	T/C
DSF9	solcap_snp_sl_67542	2	35723186	A/G
DSF10	solcap_snp_sl_100592	3	7100090	T/C
DSF11	solcap_snp_sl_1779	3	50094991	T/G
DSF12	solcap_snp_sl_21694	3	60059066	A/G
DSF13	solcap_snp_sl_20714	3	67127672	T/C
DSF14	solcap_snp_sl_2071114	4	5020084	G/T
DSF15	solcap_snp_sl_101013	4	54543361	A/T
DSF16	solcap_snp_sl_69262	4	60067355	C/G
DSF17	solcap_snp_sl_47590	4	64717474	T/C
DSF18	solcap_snp_sl_51106	5	8917895	T/G
DSF19	solcap_snp_sl_51600	5	20576788	T/C
DSF20	solcap_snp_sl_69404	5	52116995	T/C
DSF21	solcap_snp_sl_12268	5	63293808	C/G
DSF22	solcap_snp_sl_34975	6	3502385	A/G
DSF23	solcap_snp_sl_36705	6	32122683	T/G
DSF24	solcap_snp_sl_55874	6	37730097	A/G
DSF25	solcap_snp_sl_54417	6	48367262	A/G
DSF26	solcap_snp_sl_11180	7	1815826	A/G
DSF27	solcap_snp_sl_38939	7	57109419	A/G
DSF28	solcap_snp_sl_55505	7	63665011	A/C
DSF29	solcap_snp_sl_7305	8	711380	A/C
DSF30	solcap_snp_sl_56732	8	1916037	T/C
DSF31	solcap_snp_sl_4374	8	55937442	T/G
DSF32	solcap_snp_sl_34862	8	63766185	A/G
DSF33	solcap_snp_sl_28404	9	651775	A/T
DSF34	solcap_snp_sl_45095	9	6096766	A/G
DSF35	solcap_snp_sl_29222	9	67465994	T/C
DSF36	solcap_snp_sl_69743	9	70776666	T/C
DSF37	solcap_snp_sl_45992	10	49125	T/C
DSF38	solcap_snp_sl_46021	10	162566	T/C

DSF39	solcap_snp_sl_16511	10	58189616	T/C
DSF40	solcap_snp_sl_8834	10	64340314	T/C
DSF41	solcap_snp_sl_62736	11	5151004	T/G
DSF42	solcap_snp_sl_732	11	10015478	T/C
DSF43	solcap_snp_sl_2996	11	30715391	A/G
DSF44	solcap_snp_sl_53061	11	50649946	A/G
DSF45	solcap_snp_sl_41168	12	3036369	T/C
DSF46	solcap_snp_sl_16795	12	10579861	T/C
DSF47	solcap_snp_sl_59087	12	44105019	T/G
DSF48	solcap_snp_sl_53957	12	62088020	T/C

3.2.6 Linkage test

Linkage between the *BIF* locus and each of the 48 SolCap markers was detected by testing for a statistical difference from a 3:1 segregation ratio for each of the 48 SolCap markers in the sub-population of F_2 plants exhibiting the *bif* phenotype. The Chi squared test was applied in Microsoft Excel.

3.2.7 DNA extraction for genotyping

Genomic DNA was extracted from young leaves using a protocol based on the use of Chelex-100 (Bio-Rad, Hemel Hempstead, UK) with modifications (Wang et al. 2012; Casquet et al. 2012; Phillips et al. 2012; Walsh 1991; Turan et al. 2015; Pirllea et al. 2016). In brief, ~20-30 mg of leaf tissue from expanding leaves of < 2 cm length was disrupted in an extraction buffer containing 5% w/v Chelex 100 and 3 mm soda glass balls (Smith Scientific Ltd, Kent, UK) by a mixer mill (VWR, Lutterworth, UK); the samples were subsequently incubated at 100^{o}C for 5 min in a PCR machine with a hot-lid (MJ Research, Canada). Lastly, the samples were centrifuged for 4 minutes at 5,509 × g and stored at 4^{o}C for short-term or at -20^{o}C for long-term (96-well format protocol detailed in Appendix A).

3.2.8 Genotyping by polymerase chain reaction (PCR)-based markers

Fine mapping was performed using additional polymorphic KASP markers designed in the mapping interval. The reactions were performed using 1 µl of purified genomic DNA (~ 0.16 µg) in a 10 µl reaction volume containing 1x KASP master mix buffer (LGC, Teddington, UK). Using a CFX96 (Bio-Rad, Hemel Hempstead, UK) real-time PCR machine, thermal cycling was initiated at 94°C for 15 minutes, followed by 9 cycles of 94°C for 20 secs, 61 to 55°C for 1 minute (0.6°C drop per cycle), and then 25 cycles of 94°C for 20 sec, 55°C for 1 minute, then the temperature was decreased to 37°C for 1 minute for the final step of fluorescent plate-reading. KASP assays used fluorophores FAM and HEX for distinguishing genotypes; results were analysed in the Allelic Discrimination feature of CFX manager (BioRad). All KASP assays (primers) were developed by LGC (UK) based on polymorphisms and flanking sequence data provided (Table 9).

Table 9: SNPs used to develop the KASP markers on chromosome 12

Marker ID	SolCap Reference	Position	SNP
DSF50	solcap_snp_sl_41220	3240286	T/C
DSF51	solcap_snp_sl_32654	4777800	T/C
DSF52	solcap_snp_sl_63506	5833035	T/C
DSF53	solcap_snp_sl_20409	7479839	C/G
DSF54	solcap_snp_sl_40598	8948057	A/G
DSF55	solcap_snp_sl_52407	45654100	T/C
DSF56	-	51569050	G/A
DSF57	-	55725286	G/A
DSF58	solcap_snp_sl_42961	59225471	A/T
DSF59	solcap_snp_sl_53990	61861142	T/C
DSF60	solcap_snp_sl_53957	62088020	T/C

DSF61	solcap_snp_sl_16796	11159684	A/C
DSF62	solcap_snp_sl_53090	25928732	T/C
DSF63	solcap_snp_sl_38520	38590425	A/G
DSF64	solcap_snp_sl_52402	45339818	A/C
DSF65	solcap_snp_sl_18995	47510753	A/T
DSF66	solcap_snp_sl_52417	48203620	A/T
DSF67	-	50481042	C/G
DSF68	-	8566567	A/T
DSF69	solcap_snp_sl_40598	8948057	A/G
DSF70	solcap_snp_sl_16794	9973851	A/G
DSF71	solcap_snp_sl_16795	10579861	T/C
DSF72	-	10389589	T/A

3.2.9 Statistical analysis

Sample standard deviation and standard error were calculated using SigmaPlot. ANOVA was performed to evaluate the effects of the genotypes and treatments on the phenotype using SigmaPlot with significant difference claimed if $p < 0.05$ in a Tukey and Duun's post hoc test. Data was transformed prior to ANOVA to ensure the validity of the normality assumption: for flower number a $\log(x)$ transformation was used and for branching point number data, containing zero values, a $\log(x+1)$ transformation was used. Data was back-transformed prior to plotting.

3.2.10 Similarity map analysis

Firstly, an Introgression Browser (iBroswer) script was used to extract homozygous SNPs from the VCF files of the accessions available (Aflitos et al. 2014; Lin et al. 2014). Secondly, a custom Bash script used outcome file and FastTreeMP to generate the distance matrices and Newick trees in General Feature Format (GFF). Thirdly, the GFF files were organized into evenly-sized segments using another custom Bash. Lastly, the final result was plotted using

the iBrowser webserver script. This description was provided by Tom Kurowski, Cranfield University – March, 2017.

3.3 RESULTS

3.3.1 The bifuricate phenotype: initial characterization and inheritance

Multiple tomato germplasm lines and hybrids were crossed in a single seed descent (SSD) breeding programme at CNPH aiming to obtain small elongated fruits (“miniplum” or “grape” type) combining high-yield and high brix value. One inbred line was selected within the segregating populations due to its high number of branching points and flowers and it was named *bifuricate* due to the increased truss branching. LAM183 was an alternative inbred line developed from the same breeding programme with similar fruit morphology, high brix value, and general growth habit, but lacking the increase in truss branching. LAM183 and *bifuricate* lines are therefore phenotypically related, but the precise pedigree and genetic differences between them were unknown at the initiation of this study.

3.3.2 Truss development and characterisation in the contrasting inbred lines

The most obvious phenotypic difference between *bif* and LAM183 is the significantly higher number of flowers produced on *bif* trusses (Figure 16). Considering the mean of the first two trusses, *bif* produced 39.8 ± 1.6 flowers per truss, which was 3.3-fold higher than LAM183, which produced a mean of 12.0 ± 0.3 flowers per truss (Figure 17). The number of branching points within the truss was also affected – *bifuricate* trusses show a mean of 4.1 ± 1.8 branching points per truss in contrast with LAM183 which shows 0.16 ± 0.37, representing a 25.6-fold difference. Both traits are statistically significantly different between the two inbred lines (phenotype summary, Table 10).

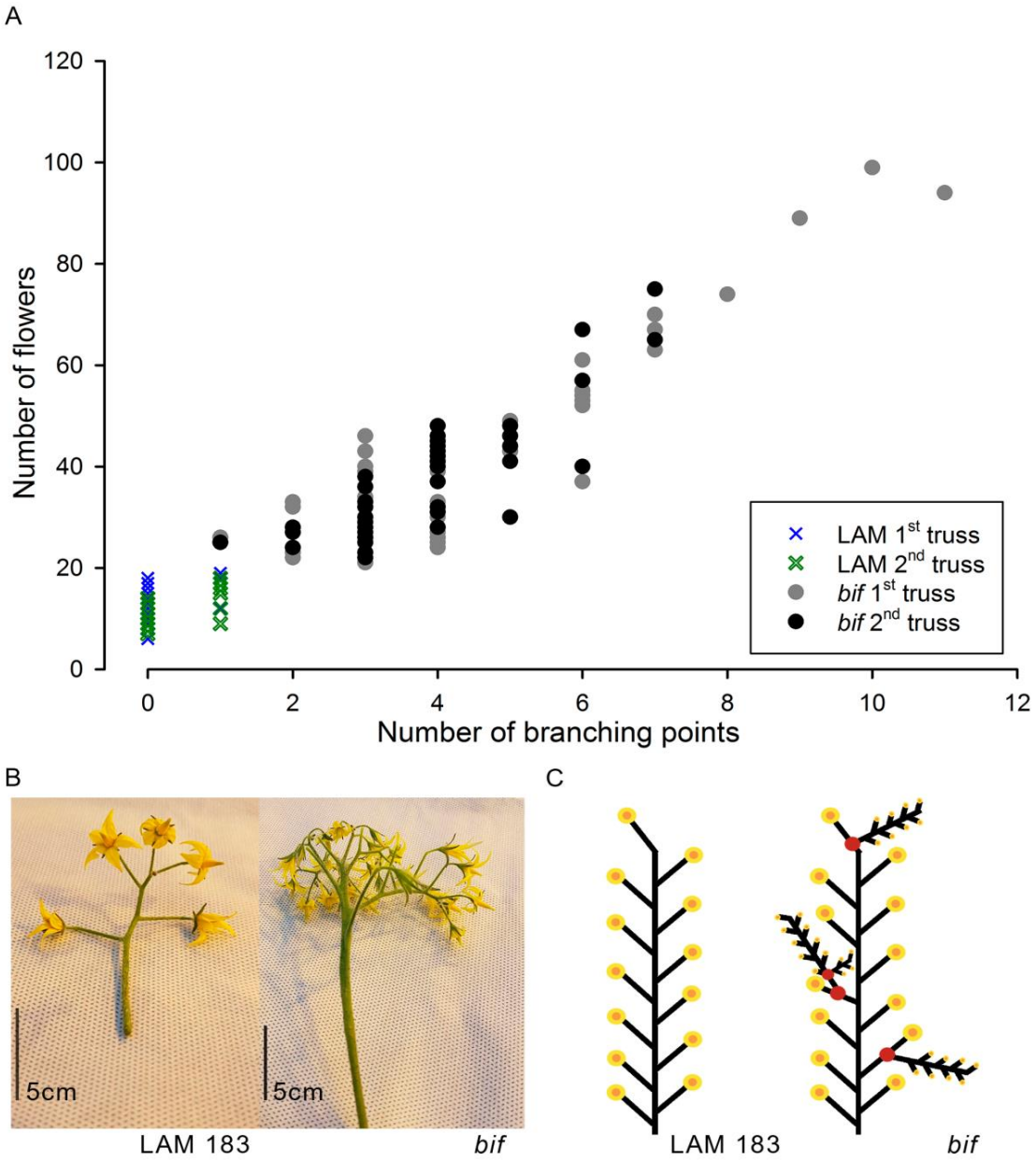


Figure 16: Characterization of the first and second truss of LAM183 (*bif*$^{+}$) and bifuricate (*bif*) inbred line plants. Phenotype was scored and numbers of flowers and branch points were recorded at 52 days after germination (DAG) in an F_2 population of 96 plants (A). Images of representative trusses 30 DAG are shown (B) and a schematic diagram illustrates the mean number of flowers (yellow circle) and branching points (red dots) on the first truss (C).

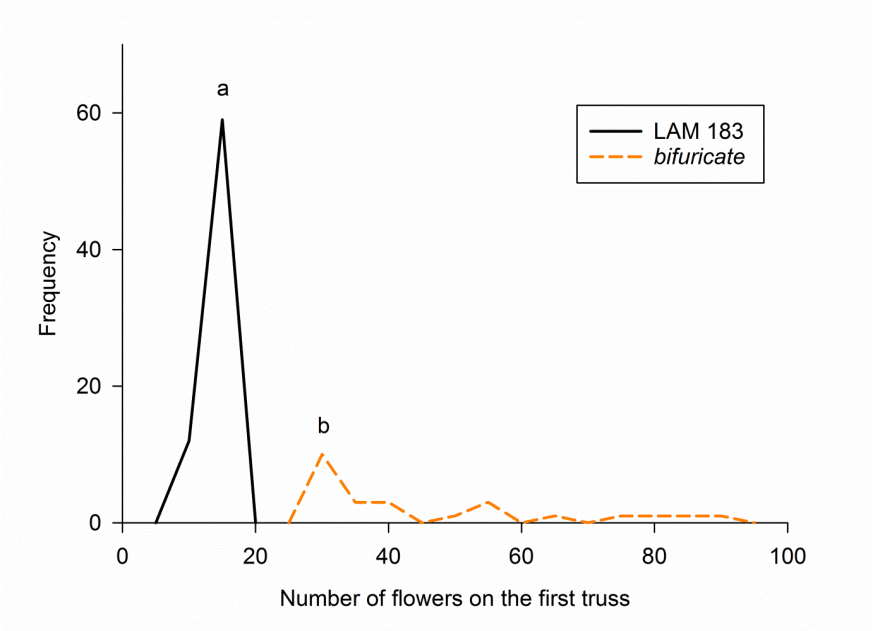


Figure 17: Floral production in the F_2 population. Phenotype was scored and numbers of flowers were recorded at 52 days after germination in an F_2 population of 96 plants in which 71 were scored as the wild type inbred line LAM183 and 25 as *bifuricate* (*bif*). The letters represent the statistical difference between both groups.

Table 10: Phenotype characterisation and comparison between LAM183 and *bifuricate* (*bif*) parental lines. Significant differences are represented by letters. Errors are standard errors of the mean; different population sizes were used: n = 15 (plant height), n = 20 (seeds per fruits); n = 30 (taproot length) and n = 48 (number of flowers and branching points/truss). DAG, days after germination.

Trait	LAM183		*bif*	
	1st truss	2nd truss	1st truss	2nd truss
Flowers per truss	12.85 ± 0.42[a]	11.08 ± 0.38[b]	41.81 ± 2.77[c]	37.75 ± 1.66[c]
Branch points per truss	0.12 ± 0.04[a]	0.20 ± 0.05[a]	4.37±0.31[b]	3.81 ± 0.16[b]
Leaves before first truss	6.66 ± 0.21[a]		6.86 ± 0.21[a]	
Plant height at 61 days (cm)	87.7 ± 3.7[a]		119.1 ± 2.7[b]	
Taproot length at 13 DAG (cm)	5.4±0.1[a]		10.4±0.2[b]	
Seeds per fruit	46.1 ± 1.1[a]		45.9 ± 1.4[a]	
Seed area (mm^2 $seed^{-1}$)	5.8 ± 0.04[a]		7.6 ± 0.05[b]	

LAM183 plants were observed to exhibit branching in some trusses, and occasionally unbranched trusses were found in the first truss of *bif* plants, although scoring of plants was unambiguous when looking at multiple trusses in older plants. Thus, incomplete penetrance and a genotype × environment interaction was apparently involved with the phenotypic expression of this trait. Environmental factors such as nutrient supply, irradiance and temperature are likely to be important regulators of the programmed development of sink strength.

3.3.3 Environmental interactions – Cold effect

The effects of different temperature treatments in tomato flowering is a well-established phenomenon (Calvert 1957; Calvert 1959). When LAM183 was grown in Brazil, rare truss branching was observed, but this was not consistent with lower temperature in the UK leading to occasional branching in LAM183. Therefore, an experiment was conducted to test if low temperature could induce truss branching, and if there was an interaction between genotype and temperature. LAM183 and *bif* plants were transferred from a glasshouse at 23^{o}C to growth cabinets either at 15^{o}C (cold) or at 23^{o}C (control) for four day periods at weekly intervals over five weeks (Figure 15) during initiation and development of inflorescences, and then the subsequent truss development was recorded (Figure 18). There were more flowers and branching points in *bif* than in LAM183 ($p < 0.001$; Appendix C) at both the first and second trusses, as expected.

The cold treatment significantly increased the number of flowers produced on the first truss of *bif* (cold = 38.86 ± 2.83; control = 28.80 ± 1.78) regardless of when the plants were exposed to the lower temperature (Appendix C). The same effect was not seen in the first truss of LAM183 (cold = 12.36 ± 0.52; control = 11.85 ± 0.18), and the response of the two genotypes to cold was significantly different (p = 0.025 for the genotype × treatment interaction; Appendix C).

On the second truss there was a significant interaction between treatment and transfer point ($p < 0.001$), because the cold treatment had opposite effects depending on whether the transfer was early or late in truss development: the cold treatment significantly reduced the number of flowers in *bif* up to and including the 3rd transfer (means of the first three transfers: cold = 31.26 ± 1.89; control = 39.61 ± 0.82); after this threshold, the exposure to lower temperatures increased the number of flowers produced (means of the last two transfers: cold = 43.13 ± 5.43; control = 28.27 ± 2.75). This effect of the later transfers was similar to that observed for the first truss at all five transfer points (Figure 18).

The impact of cold treatment on flower number in the second truss was much smaller in LAM183 compared to the *bif* line, but there was no statistical evidence of a different pattern of behaviour between genotypes based on an absence of a significant genotype × treatment interaction ($p = 0.838$; Appendix C) or a genotype × treatment × transfer point interaction ($p = 0.130$; Appendix C).

As expected, the number of branching points followed a similar pattern exhibited by the number of flowers. On the first truss, the *bif* plants showed a significant increase in branching points in the cold treatment (cold = 3.55 ± 0.22; control = 1.93 ± 0.27), compared to a non-significant difference in LAM183 (cold = 0.02 ± 0.02; control = 0), and there was a significant genotype × treatment interaction ($p < 0.004$; Appendix C), but no interaction with transfer point, similar to the observation for flower number.

On the second truss, the *bif* inflorescences were less branched when exposed to lower temperatures up to and including the 3rd transfer (means of the first three transfers: cold = 2.96 ± 0.22; control = 4.07 ± 0.21), and the effect was inverted by the 4th and 5th transfers (cold = 4.14 ± 0.55; control = 2.01 ± 0.23), resulting in a highly significant treatment × transfer point interaction ($p < 0.001$; Appendix C). However, as observed for flower numbers, there was no statistical evidence that the response of the two genotypes to cold was different, indeed LAM183 did have higher number of branching points after the 4th and 5th

transfers to cold treatment (cold = 0.53 ± 0.11; control = 0.26 ± 0.01), but the magnitude of the increase was much less than observed in *bif* line (Figure 18).

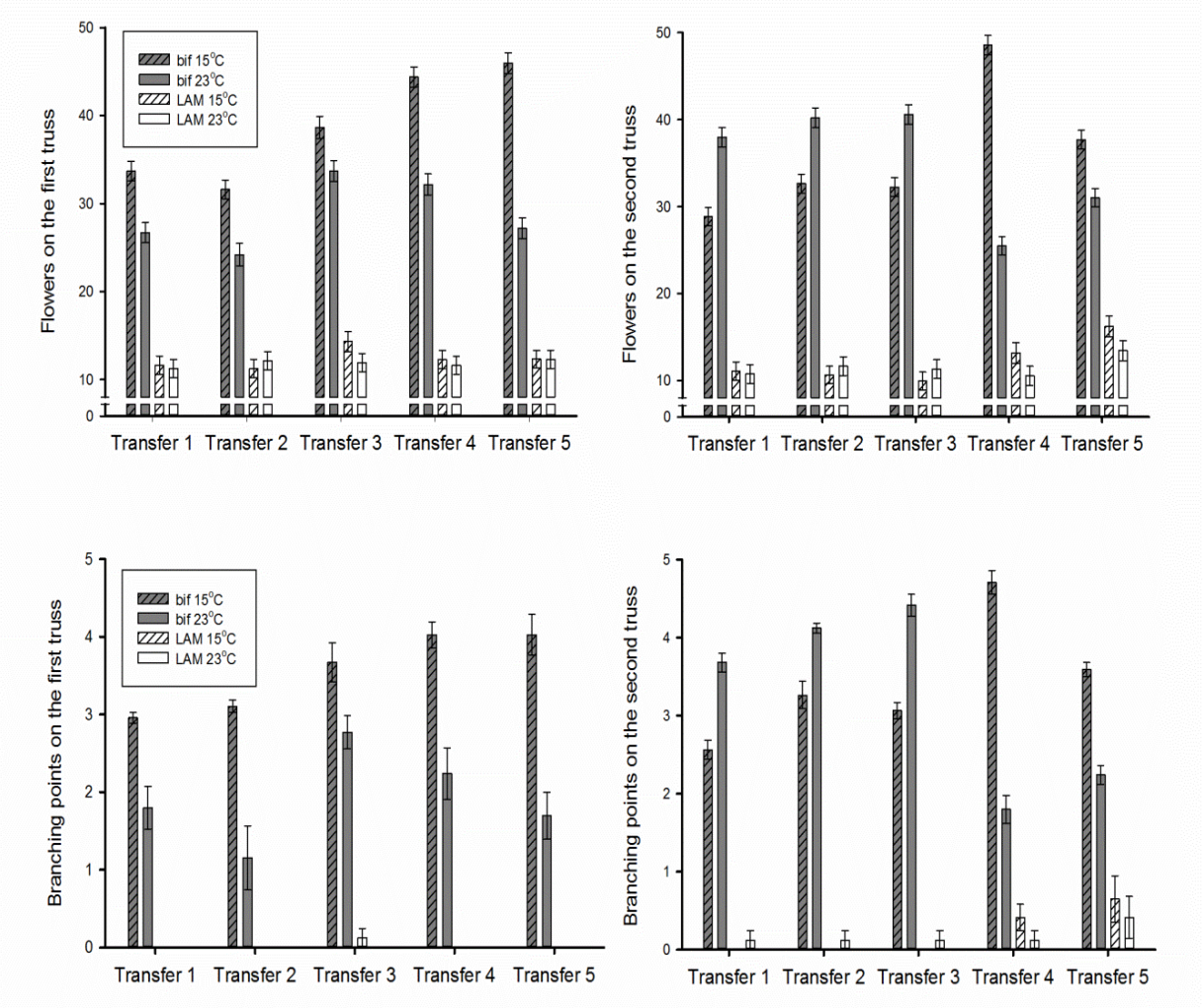


Figure 18: Cold effects on both parental lines. The numbers of flowers and branch points were recorded at 47 days after germination (DAG) in a population of 60 plants per genotype. Standard error among the replicate is represented by the bars (n = 6) – ANOVA results are shown on Appendix C: In summary, the number of flowers produced on both trusses is statistically different when *bif* plants are exposed to lower temperatures, the same effect is not seen on LAM183; the genotype effect is still present. The number of branching points is also significantly altered. Interestingly, on the second truss, an inverse effect is seen between the first 3 transfers and the last 2, probably caused by the stage of development, i.e. the second truss is formed later, thus, later transfers affect it differently.

3.3.4 The *BIFURICATE* gene: Map-based cloning

A LAM183 × *bif* F_2 population was produced and 96 plants were visually scored for truss branching and flower numbers to establish the inheritance pattern. The number of branching points for LAM183 is not always zero, but the flower number is always significantly lower than the mutant. Therefore, the phenotypic scoring was done using a combination of both traits whereby LAM183 plants have 0 or 1 branching points and max 18 flowers on the first truss. Also, if necessary, each plant could be unambiguously scored by observing multiple trusses in later development.

Twenty-three plants were scored as *bif*; seventy-three plants were scored as LAM183. A Chi-squared test indicated no statistically significant deviation from a 3:1 segregation ratio. Therefore, *bif* is determined by a single recessive gene. In order to genetically map *BIF*, the LAM183 and *bif* lines were resequenced to obtain polymorphic markers, and to provide useful information to assist CNPH-EMBRAPA's continuing breeding work with this material.

3.3.5 Resequencing of bif and LAM183 inbred lines

Illumina sequencing of genomic DNA resulted in 148 million paired-end 126 bp reads for LAM183 and 138 million reads for *bif*. The raw reads were mapped to the tomato genome reference (*S. lycopersicum* cv. Heinz 1706, version SL.2.50), with a 33- and 34-fold coverage for LAM and *bif*, respectively.

Both inbred lines used as parents came from SSD populations with a relatively large genetic basis, so a high degree of polymorphisms was expected at multiple loci. Without any filtering, 96 million and 126 million changes were found between LAM and *bif* (when compared against the reference), respectively. A large number of those polymorphisms could be due to sequencing errors, mis-mapping, or regions with insufficient coverage. Nevertheless, the density of Single Nucleotide Polymorphisms (SNPs) and Insertions and Deletions (InDels) were presented to indicate the regions where LAM183 and *bif* differ genetically.

To analyse both genomes individually, against the reference or against themselves; shared variants between *bif*, LAM183 and Heinz 1706 were excluded. Thus, the term "*unique"* is used to describe polymorphisms that are only present in one of the genotypes. After additional SNP filtration, by excluding polymorphisms with quality values under 230.8 (threshold based on the overall quality score and the SolCap validated polymorphism), plots were created with the 625,887 *bif* and 479,247 LAM183 unique SNPs (Figure 19) and with 77,049 *bif* and 81,894 LAM183 unique InDels (Figure 20).

The *bif* genome presents a single relevant peak of SNPs (33,085) at the top of chromosome 1, in contrast with LAM183 that presents more SNPs (11,343) at the bottom of the same chromosome; at the same time, on chromosome 6 LAM183 shows a peak with 45,544 SNPs and *bif* does not. No relevant information was found on chromosomes 2 (*s* locus), 3, 9 and 10. Even with repeated SNPs deleted, the same distribution pattern was displayed on chromosomes 4 and 5. On chromosomes 7 and 8 a single small peak of SNPs was detected in each genotype (chromosome 7: *bif,* 6,651; LAM183, 12,720; chromosome 8: *bif,* 10,727; LAM183, 6,665).

The most interesting results were found on chromosomes 11 and 12, which are covered by unique SNPs. On chromosome 11, *bif* is very similar to Heinz 1706 (reference genome) and LAM183 is different. Whereas, on chromosome 12 LAM183 is the closest to Heinz 1706 and *bif* presents a large number of SNPs across the whole chromosome; suggesting that linkage blocks were preserved during the SSD breeding program. The InDels distribution pattern was consistent with the SNP plot.

Results from the variant calling were used to identify 48 SNP markers for the initial genetic mapping (supplementary table S3). Validated SNPs were used to choose existing SolCap markers; ~4 per chromosome (i.e. 2 at the top, 2 at the bottom). Linkage analyses showed that the *bif* phenotype was closely linked with two markers on chromosome 12 (DSF46 and DSF47), at the border of the heterochromatin (Figure 21).

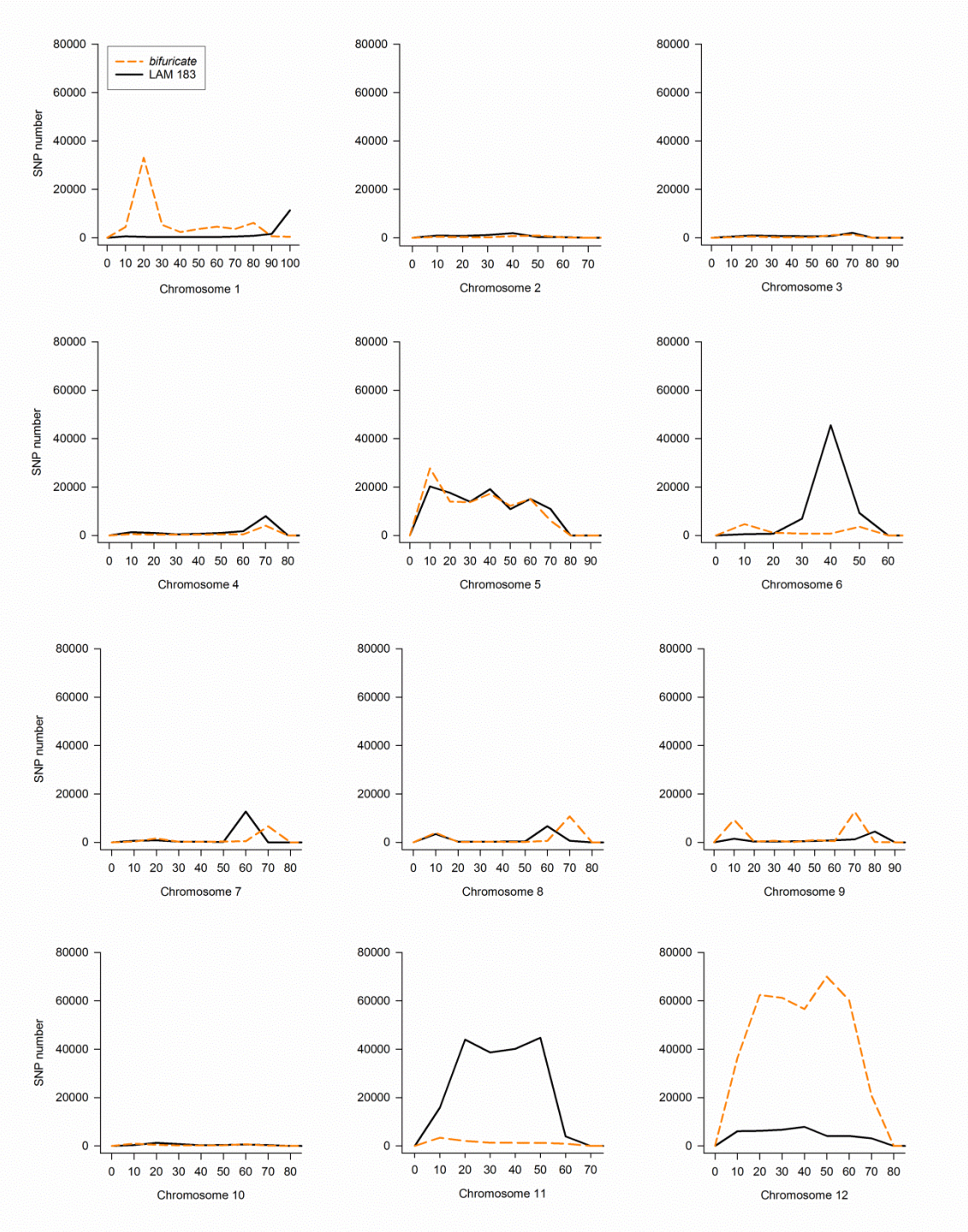


Figure 19: Genome-wide SNPs uniquely discovered in *bifuricate* and LAM183. Unique = only present in one of the genotypes. Both lines were compared against the genome reference (S. lycopersicum cv. Heinz 1706); Polymorphisms with quality lower than 230.8 or repeated were excluded from the analysis presented above (unique - SNPs plotted on *bif* and LAM are 625,887 and 479,247, respectively); this threshold was defined based on validated SNPs on this project.

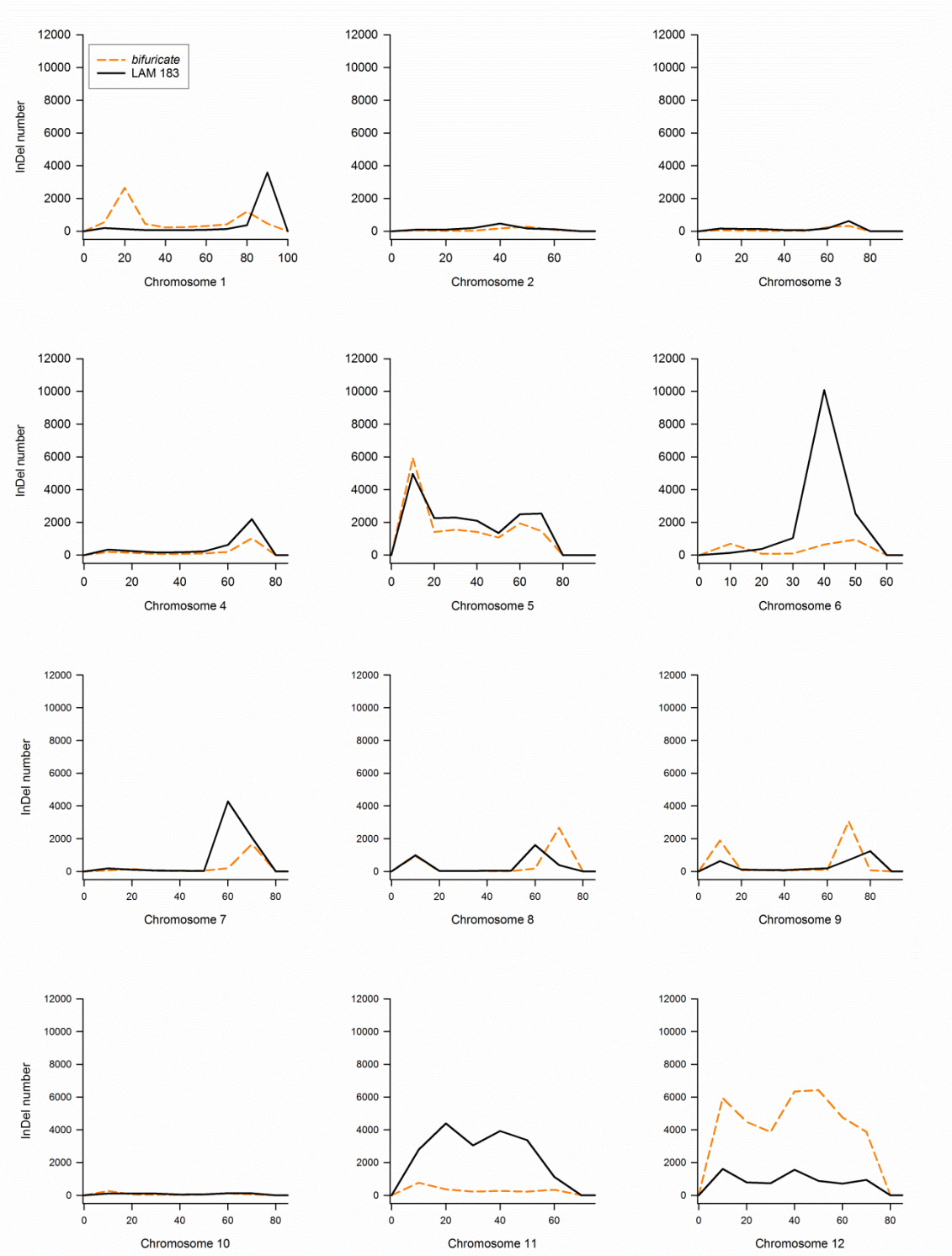


Figure 20: Genome-wide unique InDels discovered in *bifuricate* and LAM183. Unique = only present in one of the genotypes. Both lines were compared against the genome reference (*S. lycopersicum* cv. Heinz 1706). Polymorphisms with quality lower than 230.8 or repeated were excluded from the analysis presented above (InDel plotted on LAM and *bif* are 81,894 and 77,049, respectively).

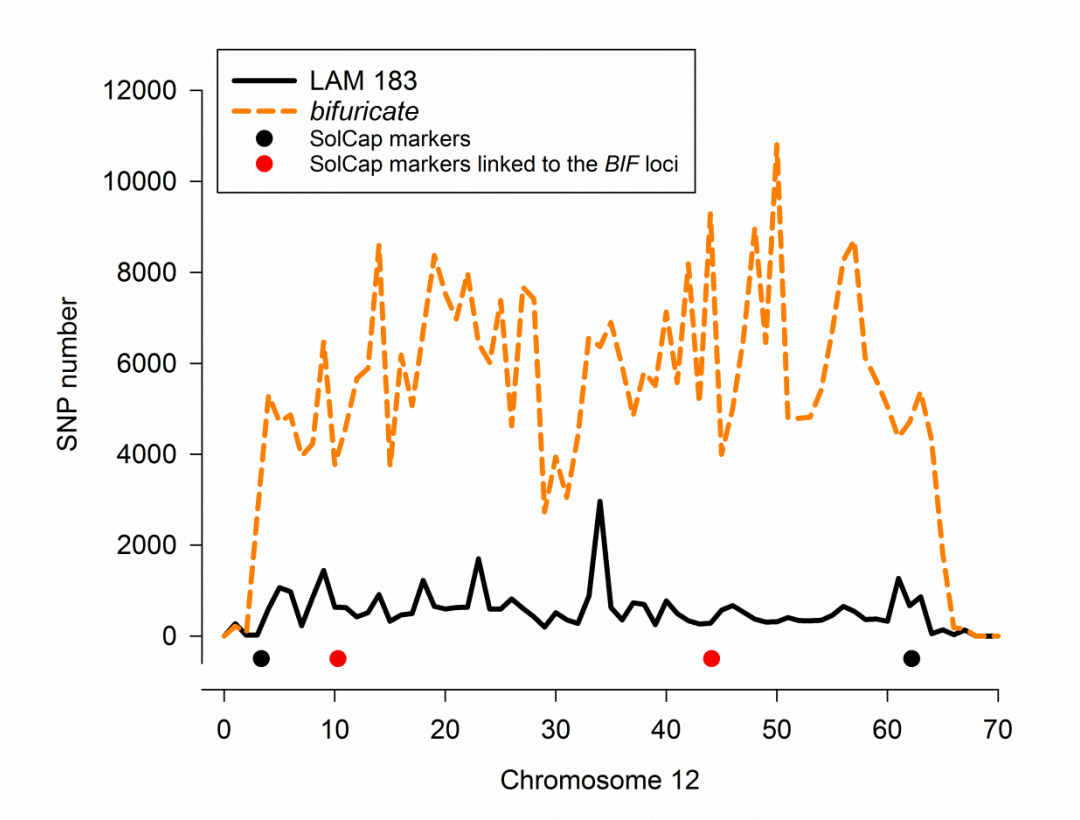


Figure 21: SNP discovery in *bifuricate* (*bif*) and LAM183 across chromosome 12. Both lines were compared against the reference genome (i.e. *Solanum lycopersicum* cv. Heinz 1706); only unique SNPs are shown (*bif* n = 367,568 and LAM183 n = 38,460). The four dots are representing the SolCap markers using for primarily mapping (markers characterized by the red dots are linked with the *BIF* locus).

3.3.6 High resolution gene mapping

The population of 96 F_2 plants was genotyped with additional markers. Recombinant plants identified with the first batch of markers (i.e. DSF50-DSF60; see supplementary table S3) narrowed the target region from 59.05 Mbps (DSF45: 3,036,369 bp – DSF48: 62,088,020 bp) to 44.08 Mbps (DSF53: 7,479,839 bp and DSF56: 51,569,050 bp). DSF53 and DSF56 were used to screen an F_2 population of 6,000 plants: 600 recombinants were recovered and phenotyped. A new batch of seven markers (i.e. DSF61-DSF67) was used to increase the resolution in the target area on the 600 recombinants. The genotyping of the recombinants reduced the target window to 3.68 Mbp

(DSF53: 7,479,839 bp – DSF61: 11,159,684 bp). A small subset of the original recombinants (n = 19), were genotyped with another series markers (i.e. DSF68-DSF72); the result narrowed the region containing the *BIF* gene to 2.01 Mbp (DSF68: 8,566,567 – DSF71: 10,579,861) containing 53 candidates (Table 12).

3.3.7 Candidate gene analyses

From this list of 53 genes (appendix D), many were unlikely to be functional: four transposons and 17 apparently pseudogenes (which had no expression recorded in transcriptomics databases and/or presented a single short exon). Another 26 genes presented only synonymous (15) or conservative missense changes in exons (11).

These analyses narrowed down to six remaining candidate genes. Four of which, have SNPs changing the amino-acid polarity (*Solyc12g019130*, *Solyc12g019140*, *Solyc12g019200* and *Solyc12g019320*); therefore, considered as moderate effects on the coded proteins. The other two genes in the interval displayed polymorphisms considered as being of potentially high phenotypic impact – a frameshift mutation on the first exon of *Solyc12g019420* and a stop codon gained on fifth exon of *Solyc12g019460* (Table 13).

3.3.8 Germplasm origin of the genomic region encompassing the BIF locus

The SNPs in the mapping interval of *bif* were compared to LAM183 and other 63 tomato accessions (Figure 22); this region in *bif* was found to be very similar to *S. galapagense* accessions LA1044, LA1401, LA0483, and LA0528. At a chromosomal level, the differences observed between the other wild species (*S. habrochaites, S. arcanum, S. pennellii, S. chilense, S. peruvianum, S. huaylasense, S. corneliomulleri, S. chmielewskii*, *and S. neorickii*) were suggesting that there was an introgression from *S. galapagense* on chromosome 12 of the *bif* line (Figure 23 and Figure 24).

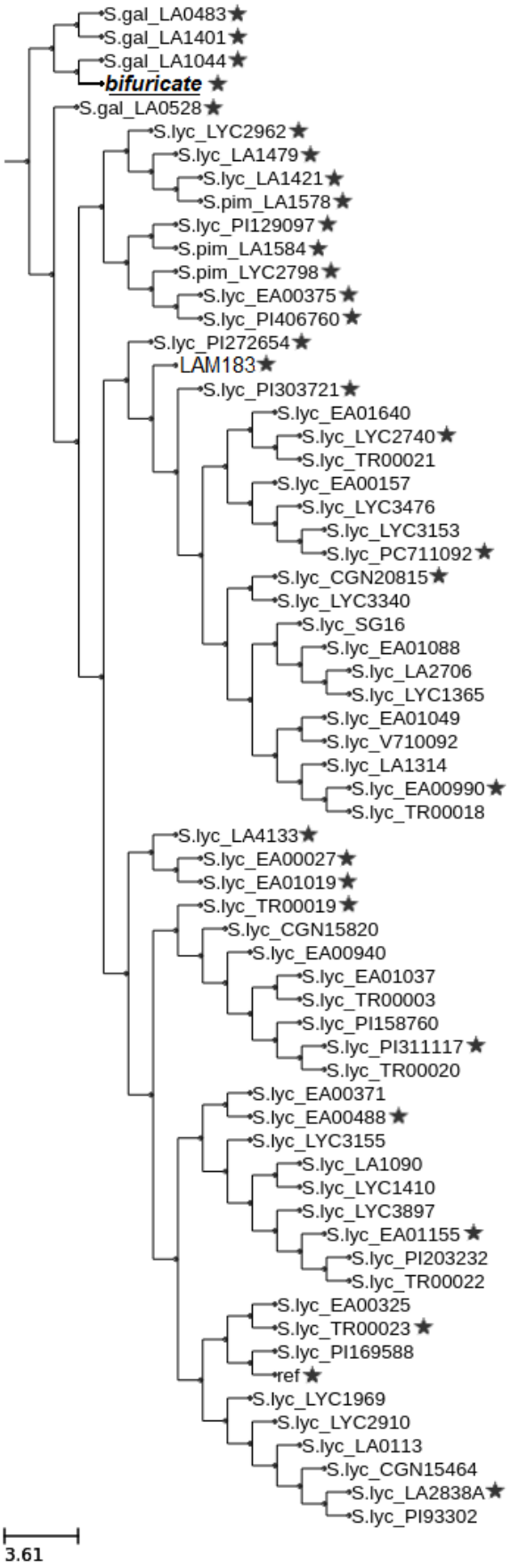
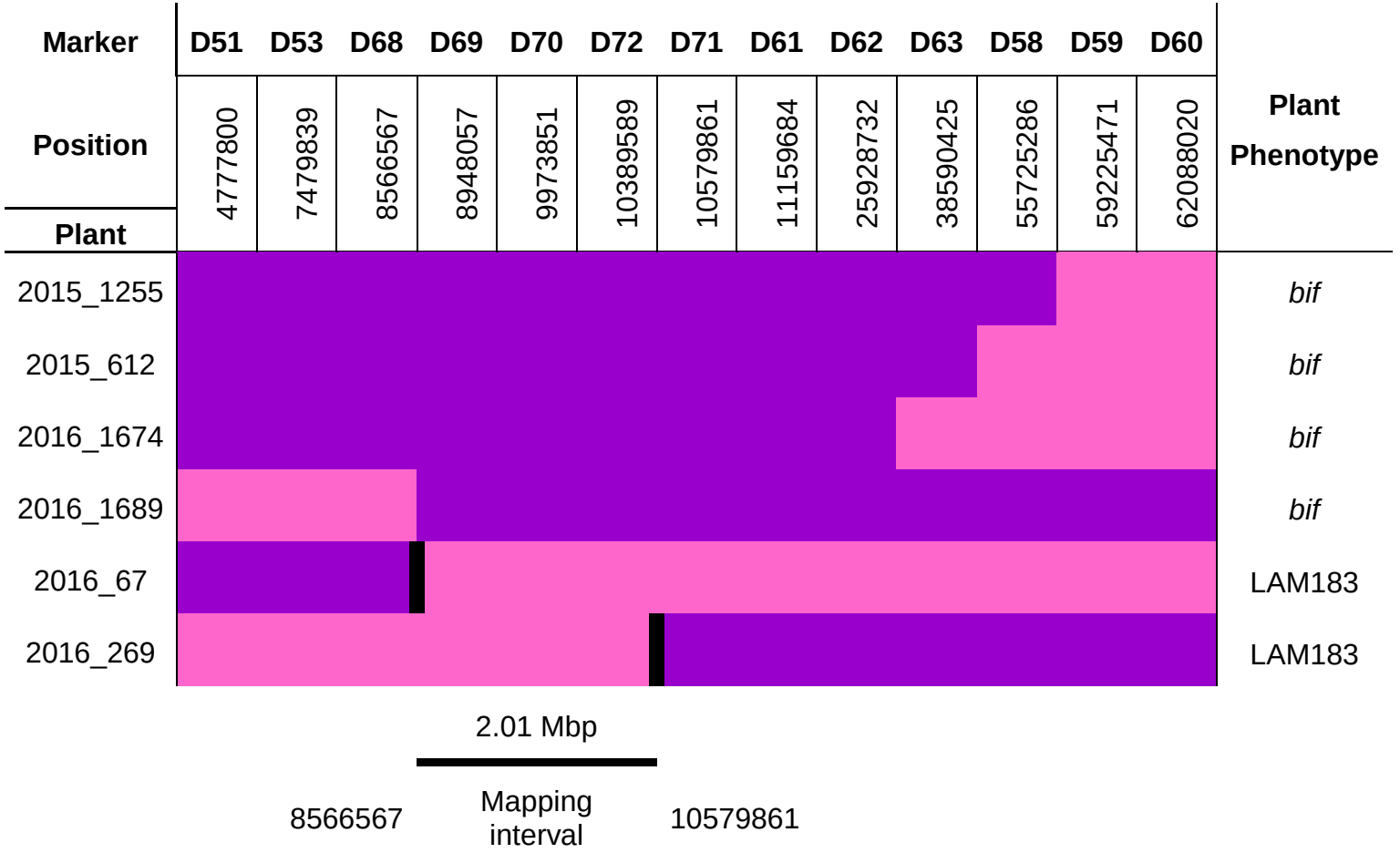


Figure 22: Similarity tree based on the SNPs in the defined mapping interval (2.01 Mbp) on chromosome 12. Accessions of *Solanum* wild species (64) were grouped according to how similar they were in comparison to the *bifuricate* genome. Black stars represent the lines selected to be part of Figure 25 and Figure 26.

Table 11 Genotyping of recombinants for fine mapping of *bif*. Results were summarised using representative recombinants and a colour code (*bif*.*bif* = purple; *bif*.+ = pink) – DSF marker are symbolized by the first letter and positions are represented in bp.



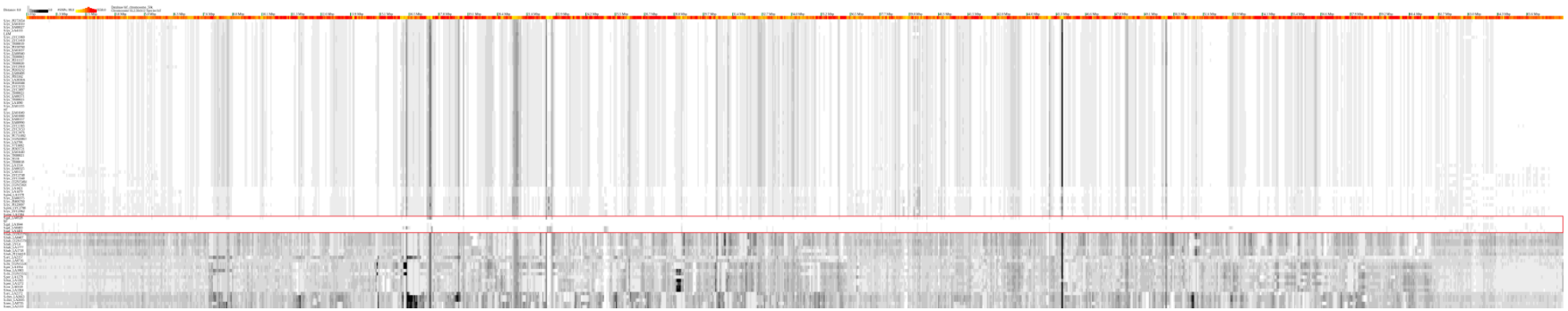


Figure 23: SNP similarities map in *bifuricate*, LAM183, Heinz 1706 and other genomes. Each signal covers a 50kb bin. The grey scale represents the differences in the bin, i.e. higher number of different SNPs result in darker tone. The red box highlights *bif* and the 4 accessions of *Solanum galapagense* (LA1044, LA1401, LA0483 and LA0528). This figure can be downloaded for visualization from: http://dx.doi.org/10.17862/cranfield.rd.4721560.

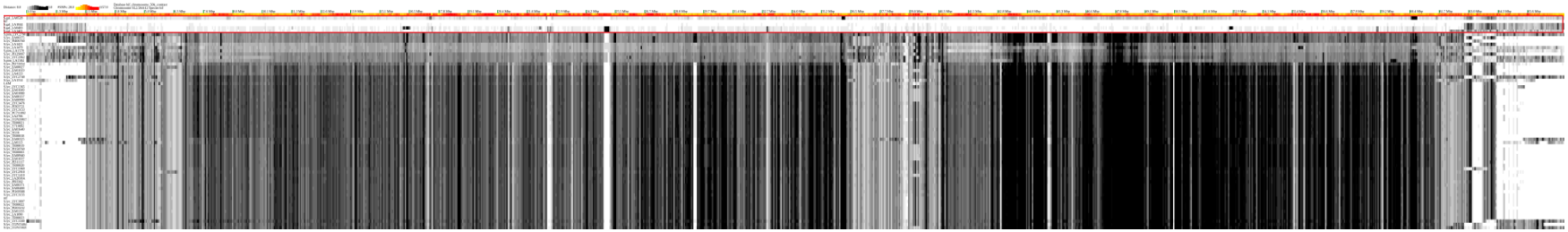


Figure 24: High contrast SNP similarities map in *bifuricate*, LAM183, Heinz 1706 and 28 other lines. Each signal is equivalent to a 50kb bin. The grey scale represents the differences in the bin, i.e. more different SNPs in the site, darker the tone. 23 accessions of wild species (i.e. *S. habrochaites, S. arcanum, S. pennellii, S. chilense, S. peruvianum, S. huaylasense, S. corneliomulleri, S. chmielewskii and S. neorickii*) were deleted from this analysis to increase the contrast of the lines that are similar to *bif*; the red box highlights *bif* and the other accessions of *S. galapagense*. Selected accessions *of S. lycopersicum* can be seen on Figure 22. This figure can be downloaded for visualization from: http://dx.doi.org/10.17862/cranfield.rd.4721560.

Table 12: Candidate genes for the *bif* phenotype

Gene	Position	Size (bp)	Type of polymorphism in coding region	Change
Solyc12g019130	9244294 - 9245985	1691	Non-conservative change	T381A and G194R
Solyc12g019140	9293132 - 9295123	1991	Non-conservative change	D206G
Solyc12g019200	9547294 - 9548500	1206	Non-conservative change	Q29P and G35C
Solyc12g019320	9971385 - 9976838	5453	Non-conservative change	I249T
Solyc12g019420	10234086 - 10236575	2489	Frameshift InDel	7^8insT
Solyc12g019460	10385358 - 10395971	10613	Stop gained	L291*****

After removing them from the analysis (Figure 24), it was clear that even though the *bif-containing* locus is similar to the *S. galapagense*, some regions (0.1 to 2.5 Mbp and ~63.5 to 66 Mbp) were closer related to *S. lycopersicum* than to *S*. *galapagense*. It was also possible to notice that the *S. galapagense* accession LA0528 (sequenced and published by Lin et al. 2014) is different from the other three *S. galapagense* accessions (sequenced and published by Aflitos et al. 2014). Similar results can be obtained, specifically for the mapping interval (Figure 26); using the same lines selected on Figure 22 (for all lines, see figure 25), *S. galapagense* LA0528 displays variants that are not present in LA0483, LA1401, LA1044 and *bif*. However, it was verified that all four accessions display the same causative SNP on the fifth exon of *Solyc12g019460*. The InDel causing a frameshift mutation in *Solyc12g019420* was detected by Aflitos et al. (2014) in all three sequenced accessions.

However, we were not able to verify if LA0528 displayed the same InDel as, the VCFs provided by Lin at al. 2014 only contain SNPs.

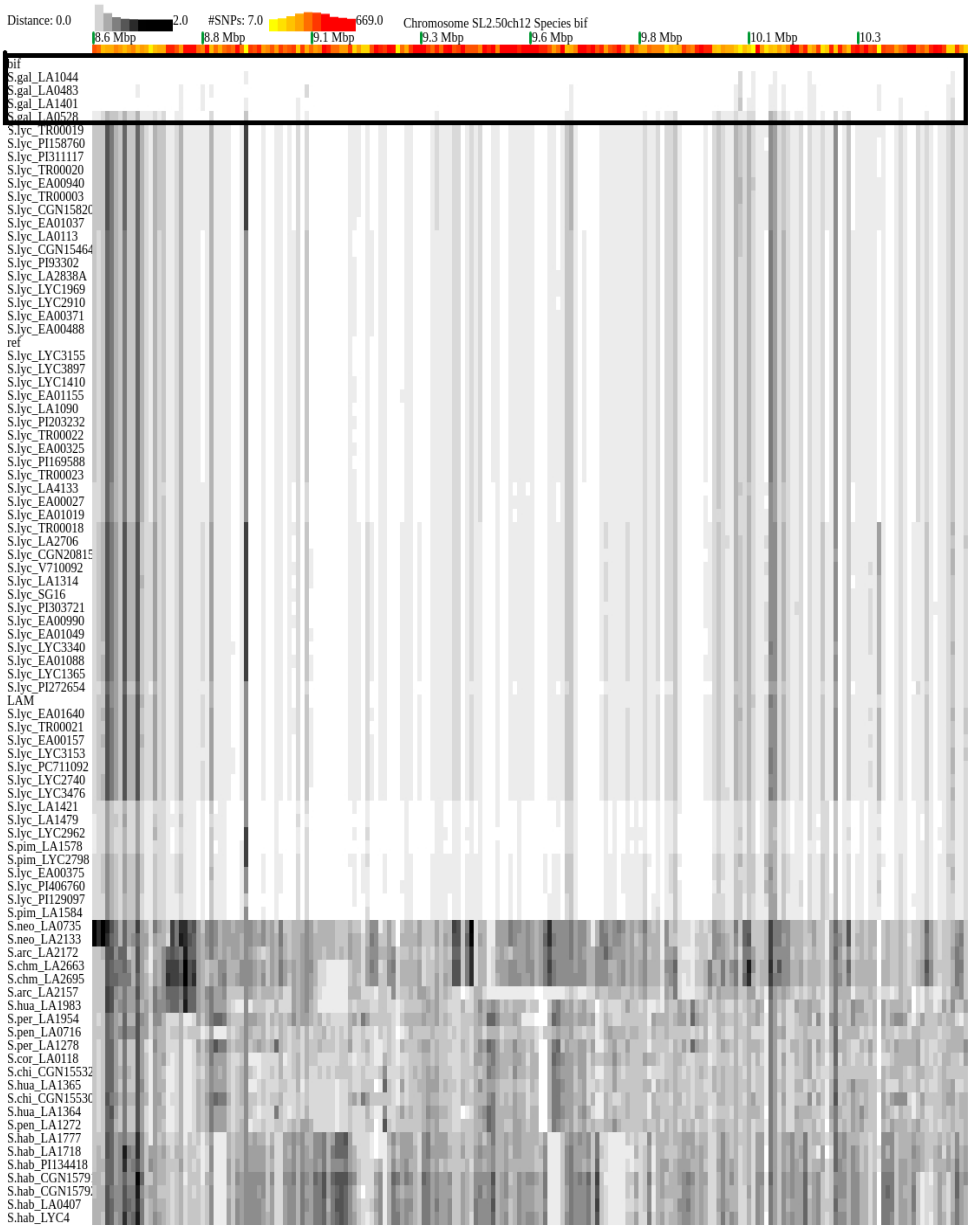


Figure 25: SNP heat map in the mapping interval. Each bin is equivalent to 10kb. The grey scale represents the differences in the bin, i.e. more different SNPs in the site, darker tone. The black box highlights the similarity between *bif* and the other *Solanum galapagense* (LA1044, LA1401, LA0483 and LA0528) accession

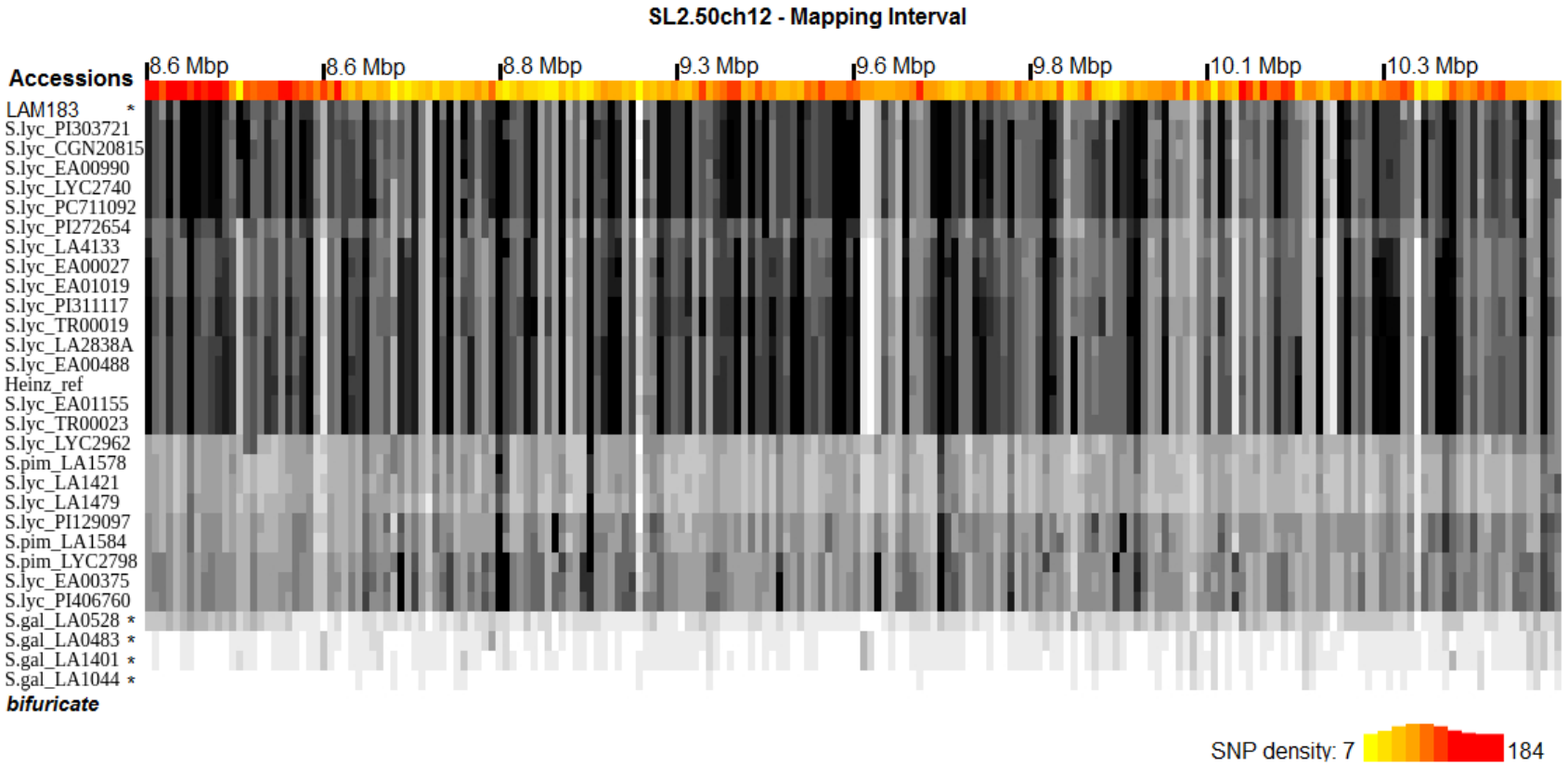


Figure 26: SNP heat map in the *bifuricate* (*bif*) gene mapping interval using the lines that are more similar to *bif* locus. (complete analysis with all 64 accessions Figure 25); The grey scale represents the differences in the bin when compared to the *bif* line, i.e. more different SNPs in the site, darker tone. Each bin is equivalent to 10 kb.

3.4 DISCUSSION

Although the transition between the vegetative and reproductive stages is well known (Yang et al. 1999; Kobayashi et al. 1999; Quinet et al. 2006; Samach & Lotan 2007; Kwiatkowska 2008), the mechanisms synchronizing the inflorescence development on the initial and the sympodial segments are still poorly understood (Hong et al. 2010; Quinet et al. 2011; Thouet et al. 2012; Poyatos-Pertíñez et al. 2016). Thus, at this moment we were not able to explain the factors responsible for the significant difference on flower production between the first and second truss on LAM183, or why this phenomenon is not shown on *bif* plants.

And yet, the plant and environment interaction results gave us a better perspective on the mutant characterization. Tomato plants, in general, are very sensitive to the cold (Shah et al. 2016). The cold stress initiates reactions at different levels (i.e. cellular, molecular and physiological) to adapt to the new environment (La Porta et al. 2015). As shown on Figure 18, the interaction between lower temperature, genotype and different exposures is very complex. But, consistently with what was demonstrated by Calvert (1957 and 1959), lower temperatures increase the number of flowers in tomato (Lewis 1953; Calvert 1957; Calvert 1959; Sawhney 1983; Adams et al. 2001). Interestingly, the interaction of the genotype and treatment was also significant.

The *BIF* initial mapping established linkage with two markers close to the heterochromatin border, on chromosome 12. The phylogenetic tree of the mapping interval suggests that the source of *BIF* locus is *S. galapagense*. In the mapping interval specifically, *bif* is clustered with *S. galapagense* LA1044 (Figure 22), due to the high level of similarities between both genomes. Interestingly, on the *Isla Santa Cruz* (origin of the less similar line, LA0528) the vegetation is significantly denser than on the other Galapagos islands, which might have contributed for the superior genetic variations among the *S. galapagense* accessions (Figure 27). These germplasms (LA0483, 1956; LA1401, 1971; LA1044, 1966 and LA0528, 1957) were collected decades ago

and no detailed descriptions of the inflorescence architecture were provided to the tomato germplasm collection of TGRC.

The fact that *S. galapagense bif* introgression does not cover the top or bottom chromosome is very interesting. Quantitative trait loci (QTLs) related to plant height, number of nodes, flower node number, leaf number, and plant fresh weight were described in LA0483 (Goldman et al. 1995; Paran et al. 1997), although they are all located close to the top (TG68, TG50A, TG111 and TG367) and bottom (TG296, TG350 and TG602) of chromosome 12 (Goldman et al. 1995; Paran et al. 1997; SGN 2017).

Figure 27: Origin of the *Solanum galapagense* accessions. The locations shown above are based on the description provided by the germplasm donor to the Tomato Genetics Resource Centre (TGRC, 2017).

Due to the proximity to the heterochromatin, approximately 6000 F2 plants were needed to be screened in order to define the map position of *bif* to a 2 Mbp segment, which is a region with an overall low gene density (~1 expressed gene per 62,000 bp). Results from the molecular markers on the current window, narrowed the number of possible candidates to 53 genes. Of those, 48 were discarded as candidates for four reasons: 4 were transposons, 17 artefacts, 15 had synonymous changes and 11 conservative missense changes; consolidating 6 genes as candidates for the *BIF* locus. Although, we cannot exclude *Solyc12g019130*, *Solyc12g019140*, *Solyc12g019200* and *Solyc12g019320* it is not likely that they are responsible for the phenotype, especially when considering the tissue where these genes were reported to be expressed (mainly on roots, seeds and leafs, respectively). In addition, the putative effect caused by those changes on *Solyc12g019130*, *Solyc12g019140*, *Solyc12g019200* and *Solyc12g019320* is moderate when compared to high effect polymorphisms present in the other two candidate genes (*Solyc12g019420* and *Solyc12g019460*).

The frameshift insertion in the first codon of *Solyc12g019420* (7^8insT) changes the next eleven amino-acids until gaining a stop codon, clearly abolishing the protein function and causing loss of BTB domain. The stop codon gained due to a SNP on *Solyc12g019460* (L291*****) disrupt the amino-acid sequence of the MAPK, in the PK_c and STKc domains. In addition to the high quality score of the mentioned polymorphisms (*Solyc12g019420*: 7^8insT; and *Solyc12g019460*: L291*****), both were also detected by the tomato genome sequencing consortium (Aflitos et al. 2014), which suggests that they are not sequencing/mapping errors, but "*real variants*".

Blast results showed that *Solyc12g019460* and MPK6 in *Arabidopsis* contain similar domains. Mitogen-activated protein kinases (MAPKs) are involved in the regulation of many different attributes in eukaryotes (Bush & Krysan 2007), including plant growth and development (Müller et al. 2010; Xu & Zhang 2015). In 2012, it was reported that due to its role on cell proliferation, MPK6 was involved in the inflorescence stem and pedicel elongation in *Arabidopsis*, which would affect its inflorescence architecture (Meng et al.

2013). In the homologous gene in tomato, a mutation in *bif* is creating a stop codon in the middle of the polypeptide sequence, most probably resulting in a dysfunctional protein, implicating MPK6 as a strong candidate for the *BIF* locus (Table 13).

At the same time, a single base insertion (thymine – 10,236,554 bp) caused a frameshift on the transcript of *Solyc12g019420*, leading to the change of Arg to a stop codon. *Solyc12g019420* is a BTB (Broad complex, Tramtrack, and Bric-a-brac)/TAZ (Transcriptional adapter zinc binder) protein domain – also known as BTB/POZ (Pox virus zinc-finger). This class of protein was reported to be involved in a different number of signalling pathways such as cytoskeleton regulation (Kang et al. 2004) , transcriptional repression (Ahmad et al. 2003), tetramerization and gating of ion channels (Du & Poovaiah 2004), protein ubiquitination/degradation (Mandadi et al. 2009). Recently published results from the protein-protein interactions of the tomato TERMINATING FLOWERING (TMF) gene indicated its association with three transcription cofactors with BTB domains – SlBOP1 (*Solyc04g04040220*), SlBOP2 (*Solyc10g0794460*) and SlBOP3 (*Solyc10g079750*), demonstrated that the interaction of this three SlBOP are involved with the control of inflorescence architecture in tomato due to its dominant role in oligomerization (Xu et al. 2016) – also making *Solyc12g019420* as a strong candidate for the *BIF* locus.

In conclusion, to verify which of the two strongest candidates are responsible for the *bif* phenotype (i.e. highly branched trusses with more flower and bigger seeds), more recombinants should be generated to separate both mutated genes. But, as they are closely located on chromosome 12, to find recombinants between the two candidate genes would be extremely laborious. Thus, transgenic experiments related to the individual genes seem to be the best alternative for the functional analyses of both genes. Polymorphisms scored within candidate genes with high phenotypical effects are uncommon. In our mapping interval we located two, of them, in genes with highly conserved domains. Thus, the *bif* phenotype might be caused by one of these more likely candidate genes or the combined effect of the truncated proteins of MPK6 and BTB/TAZ. In addition, extensive characterization experiments on diverse *S.*

galapagense germplasm indicated that trusses with 2-3 branches were frequently found in different accessions (Darwin et al. 2003). Unfortunately, the accessions were branching was found are not mentioned, nevertheless supports the data presented in the present study.

3.5 ACKNOWLEDGEMENTS

This project was initiated by a UK-Brazil Sustainable Research Networks Grant from the British Council awarded to AT and LSB. DSF was supported by a Science without Borders PhD scholarship (246783/2012-5), funded by the Brazilian Government and the Brazilian National Council of Research and Development (CNPq). TK was supported by BBSRC grant BB/L011611/1; ZK was supported by BBSRC grant BB/L01954X/1. We thank Ryan Ramsey of Syngenta for highlighting truss branching as a commercially interesting trait prior to the initiation of this project. We are grateful to the Earlham Institute, Norwich for sequencing services.

4 GENERAL DISCUSSION and FUTURE WORK

Breeders are constantly examining traits found in plants that perform better than others in different conditions, e.g. drought resistance, (Chapin et al. 1993; Barbosa et al. 2014). Depending on the studied trait, the research focuses on the above- or belowground part of the plant (Khamzina et al. 2006; Noulèkoun et al. 2017). Even though, the response in different parts of the plant is reported to be different due to several reasons, the whole-plant perspective is required to understand the metabolic control of the plant architecture.

Some research groups are now pursuing genes reported to act on particular traits to understand its global effects on plant architecture. For instance, the *CLAVATA* (*CLE*) gene family in *Arabidopsis*, which have a dominant role in plant growth, development and environmental responses (Endo et al. 2014; Wang et al. 2016). Recently, it was reported the interaction between the CLE peptides and cytokinin (Qiang et al. 2013; Endo et al. 2014; Wang et al. 2016), and thus, the effect caused on the whole plant architecture, e.g. overexpression of *CLE14*/*CLE20*, produced a plant with a reduced root phenotype (Meng & Feldman 2010). Another example of a gene controlling the general architecture would be the *C-TERMINALLY ENCODED PEPTIDE*, *CEP* (Ohyama et al. 2009). Recent work by Delay et al.(2013) reported that this family of regulatory peptides is involved in the regulation of different aspects of the root-, e.g. LR development, and shoot architecture, e.g. SAM maintenance (Christina Delay et al. 2013).

The work carried in this PhD project is especially important when considering *Solyc12g019460* (*SIMAPK1*, candidate gene for *bif*). MAP kinase signaling cascades regulate many stress and defense responses in plants, but they also control many aspects of plant growth and developmentotropic effects are clear in the mutant, experiments to link the *brt* root architecture to the shoot architecture are still to be done. I would suggest experiments with treatments of the mutant with different concentrations of P and N to verify the impact in the root architecture and how this would reflect the shoot development and architecture.

In addition, grafting experiments would be interesting to connect the shoot and root phenotype. Transcriptomics results showed that the candidate gene (*Solyc12g014590*) is differently expressed in *brt*NIL tissues; the gene is downregulated in *brt*NIL roots, when compared to *brt*NIL leaves, and MT roots and leaves. It would be also interesting to see if the *brt* shoot phenotype, i.e. dwarf with reduced yield, remains the same with WT root system. Lastly, a study of the genetic interaction of *brt* with other root mutants (e.g. *dwarf* and *rosette*) in double and triple mutants would extend our understanding about the root architecture regulation and its molecular basis.

Even though the genetic work on *brt* is slightly more advanced due to the small mapping window, the phenotypic characterisation of the *bifuricate* (*bif*) accomplished; including its reaction to environmental effects (i.e. low temperature exposure). Although both candidate genes (*Solyc12g019420* and *Solyc12g019420*) are expressed in the root and shoot (INRA 2017), grafting experiments would be useful once again to characterize the mutant shoot behaviour in a different rootstock – i.e. whether a vigorous root system increase or reduce the yield, once the sink strength is augmented. I would suggest an experiment to test different levels of N and P again to see how the alterations caused in the root architecture affect the phenotype. As mentioned before, there are similarities between the *bif* and *s* phenotype; to verify if the phenotype (more branching points and flowers) is enhanced by the other locus (epistasis), a cross (F1) between them, followed by the selection and characterization of double mutants on the F_2 population would be required.

Genome editing, e.g. CRISPR/Cas-9, and transgenic experiments would reveal in both mutants if *Solyc12g014590* as responsible for the *brt*NIL and *Solyc12g019420* and/or *Solyc12g019420* for *bif* phenotype.

REFERENCES

Abe, S. et al., 2014. Carlactone is converted to carlactonoic acid by *MAX1* in *Arabidopsis* and its methyl ester can directly interact with AtD14 in vitro. *Proceedings of the National Academy of Sciences* , 111(50), pp.18084–18089.

Adams, S.R., Cockshull, K.E. & Cave, C.R.J., 2001. Effect of temperature on the growth and development of tomato fruits. *Annals of Botany*, 88(5), pp.869–877.

Aflitos, S. et al., 2014. Exploring genetic variation in the tomato (*Solanum* section *Lycopersicon*) clade by whole-genome sequencing. *The Plant Journal*, 80(1), pp.136–148.

Ahmad, K.F. et al., 2003. Mechanism of smrt corepressor recruitment by the BCL6 BTB domain. *Molecular Cell*, 12(6), pp.1551–1564..

Allen, K.D. & Sussex, I.M., 1996. FALSIFLORA and ANANTHA control early stages of floral meristem development in tomato (L*ycopersicon esculentum Mill*.). *Planta*, 200(2), pp.254–264.

Anderson, J. V et al., 2012. Auxin and ABA act as central regulators of developmental networks associated with paradormancy in Canada thistle (C*irsium arvense*). *Functional & Integrative Genomics*, 12(3), pp.515–531.

Andrews, K.R. et al., 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, 17(2), pp.81–92.

Ansari, M.J. et al., 2013. Characterization and gene mapping of a chlorophyll-deficient mutant *clm1* of *Triticum monococcum* L. *Biologia Plantarum*, 57(3), pp.442–448.

Arms, E.M., Bloom, A.J. & St Clair, D.A., 2015. High-resolution mapping of a major effect QTL from wild tomato *Solanum habrochaites* that influences water relations under root chilling. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*.

Asamizu, E. et al., 2012. Mapping of Micro-Tom BAC-end sequences to the reference tomato genome reveals possible genome rearrangements and polymorphisms. *International Journal of Plant Genomics*, 2012.

Asins, M.J. et al., 2015. Genetic dissection of tomato rootstock effects on scion traits under moderate salinity. *Theoretical and Applied Genetics*, 128(4), pp.667–679.

Astola, L. et al., 2014. Inferring the gene network underlying the branching of tomato inflorescence. *PLoS ONE*, 9(4).

Auldridge, M.E., McCarty, D.R. & Klee, H.J., 2006. Plant carotenoid cleavage oxygenases and their apocarotenoid products. *Current Opinion in Plant Biology*, 9(3), pp.315–321.

Bahmani, M. & Maali-Amiri, R., 2015. Genetic regulation of cross tolerance in plants under biotic and abiotic stresses. *Genetics in the Third Millennium*, 13(3).

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(5), pp.1085–94.

Bar-Dror, T. et al., 2011. Programmed cell death occurs asymmetrically during abscission in tomato. *Plant Cell*, 23(11).

Barbosa, E.R.M. et al., 2014. Tree species from different functional groups respond differently to environmental changes during establishment. *Oecologia*, 174(4), pp.1345–1357..

Bari, R. et al., 2006. *PHO2*, microRNA399, and *PHR1* define a phosphate-signaling pathway in plants. *Plant Physiology*, 141(3), pp.988–999

Barthélémy, D. & Caraglio, Y., 2007. Plant architecture: A dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. *Annals of Botany*, 99(3), pp.375–407.

Bartrina, I. et al., 2011. Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in *Arabidopsis thaliana*. *Plant Cell*, 23(1), pp.69–80.

Batley, J. & Edwards, D., 2009. Genome sequence data: management, storage, and visualization. *BioTechniques*, 46(5 SPEC. ISSUE), pp.333–335

Beaudoin, N. et al., 2000. Interactions between abscisic acid and ethylene signaling cascades. *Plant Cell*, 12(7), pp.1103–1115.

Bell, E.M. et al., 2012. *Arabidopsis* lateral organ boundaries negatively regulates brassinosteroid accumulation to limit growth in organ boundaries. *Proceedings of the National Academy of Sciences of the United States of America*, 109(51), pp.21146–21151.

Bell, P.A. et al., 2002. SNPstream® UHT: Ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery. *BioTechniques*, 32(6 SUPPL.), pp.S70–S77.

Bell, P.R., 1992. Green plants: their origin and diversity. *Green plants: their origin and diversity*.

Belović, M.M. et al., 2016. Tomato (*Solanum Lycopersicum* L.) Processing Main Product (Juice) and By-Product (Pomace) Bioactivity Potential Measured as Antioxidant Activity and Angiotensin-Converting Enzyme Inhibition. *Journal of Food Processing and Preservation*, 40(6).

Bénard, C. et al., 2009. Effects of low nitrogen supply on tomato (*Solanum lycopersicum*) fruit yield and quality with special emphasis on sugars, acids, ascorbate, carotenoids, and phenolic compounds. *Journal of Agricultural and Food Chemistry*, 57(10), pp.4112–4123.

Bennett, T. et al., 2006. The *Arabidopsis MAX* pathway controls shoot branching by regulating auxin transport. *Curr Biol*, 16.

Bernstad, A.K., Cánovas, A. & Valle, R., 2017. Consideration of food wastage along the supply chain in lifecycle assessments: A mini-review based on the case of tomatoes. *Waste Management and Research*, 35(1).

Beveridge, C.A., 2006. Axillary bud outgrowth: sending a message. *Current Opinion in Plant Biology*, 9(1), pp.35–40.

Beveridge, C.A., Ross, J.J. & Murfet, I.C., 1996. Branching in pea: action of genes *Rms3* and *Rms4*. *Plant Physiology*, 110(3), pp.859–865.

Beveridge, C.A., Ross, J.J. & Murfet, I.C., 1994. Branching Mutant r*ms-2* in *Pisum sativum* (Grafting Studies and Endogenous Indole-3-Acetic Acid Levels). *Plant Physiology*, 104(3), pp.953–959.

Bianco, M.S., Cecílio Filho, A.B. & de Carvalho, L.B., 2015. Nutritional status of the cauliflower cultivar “verona” grown with omission of out added macronutrients. *Plos One*, 10(4), p.e0123500.

Bilavcik, A., Zamecnik, J. & Faltus, M., 2015. Cryotolerance of apple tree bud is independent of endodormancy. *Frontiers in Plant Science* , 6, p.695.

Boerjan, W. et al., 1995. *Superroot*, a recessive mutation in *Arabidopsis*, confers auxin overproduction. *Plant Cell*, 7(9), pp.1405–1419.

Bolger, A.M., Lohse, M. & Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), pp.2114–2120.

Bortolin, S. et al., 2004. Analytical validation of the tag-it high-throughput microsphere-based universal array genotyping platform: application to the multiplex detection of a panel of thrombophilia-associated single-nucleotide polymorphisms. *Clinical Chemistry*, 50(11), p.2028 LP-2036.

Bradley, D. et al., 1997. Inflorescence commitment and architecture in *Arabidopsis*. *Science*, 275(5296).

Bragin, E., 2012. Interactive genome browser based on html5 and related web technologies. *Cranfield University, UK.*

Braun, A., Little, D.P. & Köster, H., 1997. Detecting *CFTR* gene mutations by using primer oligo base extension and mass spectrometry. *Clinical Chemistry*, 43(7), pp.1151–1158.

Busch, B.L. et al., 2011. Shoot branching and leaf dissection in tomato are regulated by homologous gene modules. *The Plant Cell*, 23(10), pp.3595–3609.

Bush, S.M. & Krysan, P.J., 2007. Mutational evidence that the *Arabidopsis* MAP kinase *MPK6* is involved in anther, inflorescence, and embryo development. *Journal of Experimental Botany*, 58(8).

Busov, V.B., Brunner, A.M. & Strauss, S.H., 2008. Genes for control of plant stature and form. *New Phytologist*, 177(3), pp.589–607.

Calegari, F., 2012. *CyclinD2* at the edge: Splitting up cell fate. *EMBO Journal*, 31(8), pp.1850–1852.

Calvert, A., 1957. Effect of the early environment on development of flowering in the tomato: I. Temperature. *Journal of Horticultural Science*, 32(1), pp.9–17..

Calvert, A., 1959. Effect of the early environment on the development of flowering in tomato II. Light and temperature interactions. *Journal of Horticultural Science*, 34(3), pp.154–162.

Cannon, W.A., 1949. A tentative classification of root systems. *Ecology*, 30(4), pp.542–548.

Cardoso, C. et al., 2014. Natural variation of rice strigolactone biosynthesis is associated with the deletion of two *MAX1* orthologs. *Proceedings of the National Academy of Sciences,* 111(6), pp.2379–2384.

Carmel-Goren, L. et al., 2003. The *SELF-PRUNING* gene family in tomato. *Plant Molecular Biology*, 52(6), pp.1215–1222.

Casal, J.J., Sanchez, R.A. & Deregibus, V.A., 1986. The effect of plant density on tillering: The involvement of R/FR ratio and the proportion of radiation intercepted per plant. *Environmental and Experimental Botany*, 26(4), pp.365–371.

Cashman, J.R. et al., 2001. Population distribution of human flavin-containing monooxygenase form 3: Gene polymorphisms. *Drug Metabolism and Disposition*, 29(12), pp.1629–1637.

Casimiro, I. et al., 2001. Auxin transport promotes *Arabidopsis* lateral root initiation. *Plant Cell*, 13(4), pp.843–852.

Casimiro, I. et al., 2003. Dissecting *Arabidopsis* lateral root development. *Trends in Plant Science*, 8(4), pp.165–171.

Casquet, J., Thebaud, C. & Gillespie, R.G., 2012. Chelex without boiling, a rapid and easy technique to obtain stable amplifiable DNA from small amounts of ethanol-stored spiders. *Molecular Ecology Resources*, 12(1).

Castel, R., Kusters, E. & Koes, R., 2010. Inflorescence development in petunia: through the maze of botanical terminology. *Journal of Experimental Botany*, 61(9), pp.2235–2246.

Catchen, J.M. et al., 2011. Stacks: building and genotyping loci d*e novo* from short-read sequences. *G3: Genes, Genomes, Genetics*, 1(3), pp.171–182.

Chandler, J.W., 2014. Patterns and polarity in floral meristem and floral organ initiation. *Critical Reviews in Plant Sciences*, 33(6), pp.457–469.

Chapin, F.S., Autumn, K. & Pugnaire, F., 1993. Evolution of suites of traits in response to environmental stress. *The American Naturalist*, 142, pp.78–92.

Char, S.N. et al., 2017. An agrobacterium-delivered CRISPR-Cas9 system for high-frequency targeted mutagenesis in maize. *Plant Biotechnology Journal*, 15(2), pp.257–268.

Chaudhury, A.M. & Berger, F., 2001. Maternal control of seed development. *Seminars in Cell and Developmental Biology*, 12(5), pp.381–386.

Chaves, M.M., Maroco, J.P. & Pereira, J.S., 2003. Understanding plant responses to drought - from genes to the whole plant. *Functional Plant Biology*, 30(3), pp.239–264.

Chayut, N. et al., 2015. A bulk segregant transcriptome analysis reveals metabolic and cellular processes associated with orange allelic variation and fruit β-carotene accumulation in melon fruit. *BMC Plant Biology*, 15(1).

Chuck, G. et al., 2002. The control of spikelet meristem identity by the branched *silkless1* gene in maize. *Science*, 298(5596), pp.1238–1241.

Chuck, G. et al., 2007. The maize *TASSELSEED4* microRNA controls sex determination and meristem cell fate by targeting *TASSELSEED6*/*INDETERMINATE SPIKELET1*. *Nature Genetics*, 39(12), pp.1517–1521.

Chuck, G., Meeley, R. & Hake, S., 2008. Floral meristem initiation and meristem cell fate are regulated by the maize *AP2* genes *ids1* and *sid1*. *Development*, 135(18), pp.3013–3019.

Chuck, G., Meeley, R.B. & Hake, S., 1998. The control of maize spikelet meristem fate by the *APETALA2*-like gene indeterminate *spikelet1*. *Genes and Development*, 12(8), pp.1145–1154.

Chun, L. et al., 2005. Genetic analysis of maize root characteristics in response to low nitrogen stress. *Plant and Soil*, 276(1–2), pp.369–382.

Cingolani, P. et al., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w 1118; iso-2; iso-3. *Fly*, 6(2), pp.80–92

Clevenger, J. et al., 2015. Single nucleotide polymorphism identification in polyploids: A review, example, and recommendations. *Molecular Plant*, 8(6), pp.831–846.

Cline, M.G., 1997. Concepts and terminology of apical dominance. *American Journal of Botany*, 84(8), pp.1064–1069.

Cline, M.G. & Deppong, D.O., 1999. The role of apical dominance in paradormancy of temperate woody plants: a reappraisal. *Journal of Plant Physiology*, 155(3), pp.350–356.

Cockshull, K.E. & Ho, L.C., 1995. Regulation of tomato fruit size by plant density and truss thinning. *Journal of Horticultural Science*, 70(3).

Cong, B. & Tanksley, S.D., 2006. *FW2.2* and cell cycle control in developing tomato fruit: A possible example of gene co-option in the evolution of a novel organ. *Plant Molecular Biology*, 62(6), pp.867–880.

Cordell, D., Drangert, J.-O. & White, S., 2009. The story of phosphorus: Global food security and food for thought. *Global Environmental Change*, 19(2), pp.292–305.

Cornelis, S. et al., 2017. Forensic SNP genotyping using nanopore MinION Sequencing. *Scientific Reports*, 7.

Coutts, M.P., 1983. Root architecture and tree stability. *Plant and Soil*, 71(1–3), pp.171–188.

Cuartero, J. & Fernández-Muñoz, R., 1999. Tomato and salinity. *Scientia Horticulturae*, 78(1–4), pp.83–125.

Danecek, P. et al., 2011. The variant call format and VCFtools. *Bioinformatics*, 27(15), pp.2156–2158.

Darwin, S.C., Knapp, S. & Peralta, I.E., 2003. Taxonomy of tomatoes in the galapagos islands: native and introduced species of *Solanum* section *Lycopersicon* (*Solanaceae*). *Systematics and Biodiversity*, 1(1), pp.29–53.

Dechaine, J.M. et al., 2007. Constraints on the evolution of adaptive plasticity: Costs of plasticity to density are expressed in segregating progenies. *New Phytologist*, 176(4), pp.874–882.

Delay, C., Imin, N. & Djordjevic, M.A., 2013. *CEP* genes regulate root and shoot development in response to environmental cues and are specific to seed plants. *Journal of Experimental Botany*, 64(17), pp.5383–5394.

Delay, C., Imin, N. & Djordjevic, M.A., 2013. Regulation of *Arabidopsis* root development by small signaling peptides. *Frontiers in Plant Science*, 4(SEP).

DESA, 2014. Department of economic and social affairs, (June/16).

Dielen, V. et al., 2004. *UNIFLORA*, a pivotal gene that regulates floral transition and meristem identity in tomato (*Lycopersicon esculentum*). *New Phytologist*, 161(2), pp.393–400.

Dillon, S. et al., 2006. Development of a codominant CAPS marker linked to *PRSV-P* resistance in highland papaya. *Theoretical and Applied Genetics*, 113(6), pp.1159–1169.

Dobin, A. et al., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), pp.15–21.

Doebley, J., Stec, A. & Hubbard, L., 1997. The evolution of apical dominance in maize. *Nature*, 386(6624), pp.485–488.

Dolan, R.W., 1984. The effect of seed size and maternal source on individual size in a population of *Ludwigia leptocarpa* (*Onagraceae*). *American Journal of Botany*, 71(9), pp.1302–1307.

Domagalska, M.A. & Leyser, O., 2011. Signal integration in the control of shoot branching. *Nature Reviews Molecular Cell Biology*, 12(4), pp.211–221.

Du, L. & Poovaiah, B.W., 2004. A novel family of Ca 2+/calmodulin-binding proteins involved in transcriptional regulation: Interaction with *fsh/Ring3* class transcription activators. *Plant Molecular Biology*, 54(4), pp.549–569.

Dunham, M.J., 2012. Two flavors of bulk segregant analysis in yeast. *Methods in Molecular Biology*, 871, pp.41–54.

Ecker, J.R., 1995. The ethylene signal transduction pathway in plants. *Science*, 268(5211), pp.667–675.

Edwards, D., Batley, J. & Snowdon, R.J., 2013. Accessing complex crop genomes with next-generation sequencing. *Theoretical and Applied Genetics*, 126(1), pp.1–11.

Eid, J. et al., 2009. Real-time DNA sequencing from single polymerase molecules. *Science*, 323(5910), pp.133–138.

Elsayed, A.Y. et al., 2012. Inheritance of resistance to late blight in tomato and identification of SSR markers linked to resistance through bulked segregant analysis. *Acta Horticulturae*, 929, pp.245–252.

Elshire, R.J. et al., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLOS ONE*, 6(5), pp.19379.

Emmanuel, E. & Levy, A.A., 2002. Tomato mutants as tools for functional genomics. *Current Opinion in Plant Biology*, 5(2), pp.112–117.

Endo, S., Betsuyaku, S. & Fukuda, H., 2014. Endogenous peptide ligand–receptor systems for diverse signaling networks in plants. *Current Opinion in Plant Biology*, 21, pp.140–146.

Ewels, P. et al., 2016. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32(19), pp.3047–3048.

Faiss, M. et al., 1997. Conditional transgenic expression of the *IPT* gene indicates a function for cytokinins in paracrine signalling in whole tobacco plants. *The Plant Journal*, 12(2), pp.401–415.

Fan, J.-B. et al., 2000. Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays. *Genome Research*, 10(6), pp.853–860.

FAOSTAT, 2014. Food and Agriculture Organization of the United Nations (June/30). Available at: http://faostat.fao.org/.

Farooqi, M.Q.U. et al., 2016. Bulk segregant analysis (BSA) for improving cold stress resistance in maize using SSR markers. *Genetics and Molecular Research*, 15(4).

Farrar, R.R. & Kennedy, G.G., 1991. Relationship of leaf lamellar-based resistance to *Leptinotarsa decemlineata* and *Heliothis zea* in a wild tomato, *Lycopersicon hirsutum* f. glabratum. *Entomologia Experimentalis et Applicata*, 58(1), pp.61–67.

Fauser, F., Schiml, S. & Puchta, H., 2014. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis* thaliana. *Plant Journal*, 79(2), pp.348–359.

Fernandez, A.I. et al., 2009. Flexible tools for gene expression and silencing in tomato. *Plant Physiology*, 151(4).

Fischer, I. et al., 2011. Adaptation to drought in two wild tomato species: The evolution of the *Asr* gene family. *New Phytologist*, 190(4), pp.1032–1044.

Fitter, A.H., 1987. An architectural approach to the comparative ecology of plant root systems. *New Phytologist*, 106, pp.61–77.

Fitter, A.H. et al., 1991. Architectural analysis of plant root systems 1. Architectural correlates of exploitation efficiency. *New Phytologist*, 118(3), pp.375–382.

Foolad, M.R., 2007. Genome mapping and molecular breeding of tomato. *International Journal of Plant Genomics*, 2007.

Foolad, M.R. et al., 2014. Response of accessions within tomato wild species, *Solanum pimpinellifolium* to late blight. *Plant Breeding*, 133(3), pp.401–411.

Fridman, E. et al., 2001. Two tightly linked QTLs modify tomato sugar content via different physiological pathways. *Molecular Genetics and Genomics*, 266(5), pp.821–826.

Furutani, M. et al., 2004. *PIN-FORMED1* and PINOID regulate boundary formation and cotyledon development in *Arabidopsis* embryogenesis. *Development*, 131(20), pp.5021–5030.

Galli, M. & Gallavotti, A., 2016. Expanding the Regulatory Network for Meristem Size in Plants. *Trends in Genetics*, 32(6).

Gandour, M., Chedly, A. & Taamalli, W., 2013. Phenotypic plasticity of plants in response to environmental change. *Phenotypic Plasticity: Molecular Mechanisms, Evolutionary Significance and Impact on Speciation*, pp.87–100.

Gao, Z. et al., 2009. *Dwarf 88*, a novel putative esterase gene affecting architecture of rice plant. *Plant Molecular Biology*, 71(3), pp.265–276.

García-Lorenzo, M. et al., 2006. Protease gene families in Populus and *Arabidopsis*. *BMC Plant Biology*, 6(1), p.30.

Garofalo, P. et al., 2017. Environmental sustainability of agri-food supply chains in Italy: The case of the whole-peeled tomato production under life cycle assessment methodology. *Journal of Food Engineering*, 200.

Garrison, E. & Marth, M., 2012. Haplotype-based variant detection from short-read sequencing.

Gebhardt, C., 2016. The historical role of species from the *Solanaceae* plant family in genetic research. *Theoretical and Applied Genetics*, 129(12).

Gendron, J.M. et al., 2012. Brassinosteroids regulate organ boundary formation in the shoot apical meristem of *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 109(51), pp.21152–21157.

Gerald, J.N.F. et al., 2006. Identification of quantitative trait loci that regulate *Arabidopsis* root system size and plasticity. *Genetics*, 172(1), pp.485–498.

Gibbs, R.A., Nguyen, P.-N. & Caskey, C.T., 1989. Detection of single DNA base differences by competitive oligonucleotide priming. *Nucleic Acids Research*, 17(7), pp.2437–2448.

Giovannoni, J., 2001. Molecular biology of fruit maturation and ripening. *Annual Review of Plant Biology*, 52, pp.725–749.

Givnish, T.J., 2002. Ecological constraints on the evolution of plasticity in plants. *Evolutionary Ecology*, 16(3), pp.213–242.

Glass, A.D.M. et al., 2002. The regulation of nitrate and ammonium transport systems in plants. *Journal of Experimental Botany*, 53(370), pp.855–864.

Glaubitz, J.C. et al., 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE*, 9(2).

Goda, H. et al., 2002. Microarray analysis of brassinosteroid-regulated genes in *Arabidopsis*. *Plant Physiology*, 130(3), pp.1319–1334.

Goldman, I.L., Paran, I. & Zamir, D., 1995. Quantitative trait locus analysis of a recombinant inbred line population derived from a *Lycopersicon esculentum* x *Lycopersicon cheesmanii* cross. *Theoretical and Applied Genetics*, 90(7–8), pp.925–932.

Grandillo, S. et al., 2011. *Solanum* sect. *Lycopersicon*. In C. Kole, ed. *Wild Crop Relatives: Genomic and Breeding Resources SE - 9*. Springer Berlin Heidelberg, pp. 129–215.

Gratani, L., 2014. Plant phenotypic plasticity in response to environmental factors. *Advances in Botany*, 2014, pp.1–17.

Haff, L.A. & Smirnov, I.P., 1997. Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Research*, 7(4), pp.378–388.

Hagen, G. & Guilfoyle, T., 2002. Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Molecular Biology*, 49(3), pp.373–385.

Hammer, G.L. et al., 2009. Can changes in canopy and/or root system architecture explain historical maize yield trends in the u.s. corn belt? *Crop Science*, 49, pp.299–312.

Hammond, J.P. & White, P.J., 2011. Sugar signaling in root responses to low phosphorus availability. *Plant physiology*, 156(3), pp.1033–40.

Hanano, S. & Goto, K., 2011. *Arabidopsis TERMINAL FLOWER1* is involved in the regulation of flowering time and inflorescence development through transcriptional repression. *Plant Cell*, 23(9).

Hardenbol, P. et al., 2005. Highly multiplexed molecular inversion probe genotyping: Over 10,000 targeted SNPs genotyped in a single tube assay. *Genome Research*, 15(2), pp.269–275.

Hardenbol, P. et al., 2003. Multiplexed genotyping with sequence-tagged molecular inversion probes. *Nature Biotechnology*, 21(6), pp.673–678.

Harrop, T.W.R. et al., 2016. Gene expression profiling of reproductive meristem types in early rice inflorescences by laser microdissection. *The Plant Journal*, 86(1), pp.75–88.

Hartweck, L.M., 2008. Gibberellin signalling. *Planta*, 229(1), p.1.

Hatirli, S.A., Ozkan, B. & Fert, C., 2006. Energy inputs and crop yield relationship in greenhouse tomato production. *Renewable Energy*, 31(4), pp.427–438.

Le Hellard, S. et al., 2002. SNP genotyping on pooled DNAs: comparison of genotyping technologies and a semi-automated method for data storage and analysis. *Nucleic acids research*, 30(15).

Hermkes, R. et al., 2011. Distinct roles for *Arabidopsis SUMO* protease *ESD4* and its closest homolog *ELS1*. *Planta*, 233(1), pp.63–73.

Hidayatullah et al., 2008. Path coefficient analysis of yield component in tomato (*Lycopersicon esculentum*). *Pakistan Journal of Botany*, 40(2), pp.627–635.

Higashide, T. et al., 2014. Suppression of elongation and growth of tomato seedlings by auxin biosynthesis inhibitors and modelling of the growth and environmental response. *Scientific Reports*, 4.

Hirschhorn, J.N. et al., 2000. SBE-TAGS: An array-based method for efficient single-nucleotide polymorphism genotyping. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), pp.12164–12169.

Hobbie, L. & Estelle, M., 1995. The *axr4* auxin-resistant mutants of *Arabidopsis* thaliana define a gene important for root gravitropism and lateral root initiation. *The Plant Journal*, 7(2), pp.211–220.

Hodge, A. et al., 2009. Plant root growth, architecture and function. *Plant and Soil*, 321(1–2), pp.153–187.

Holland, P.M. et al., 1991. Detection of specific polymerase chain reaction product by utilizing the 5′ → 3′ exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 88(16), pp.7276–7280.

Hong, P.I., Chen, J.T. & Chang, W.C., 2010. Shoot development and plant regeneration from protocorm-like bodies of *Zygopetalum mackayi*. *In Vitro Cellular and Developmental Biology - Plant*, 46(3), pp.306–311.

Hoogendoorn, B. et al., 1999. Genotyping single nucleotide polymorphisms by primer extension and high performance liquid chromatography. *Human Genetics*, 104(1), pp.89–93.

Hu, M.-J. et al., 2016. Cloning and characterization of *TaTGW-7A* gene associated with grain weight in wheat via SLAF-seq-BSA. *Frontiers in Plant Science*, 7(DECEMBER2016).

Huang, X. et al., 2009. Natural variation at the *DEP1* locus enhances grain yield in rice. *Nature Genetics*, 41(4), pp.494–497.

Huther, C.M. et al., 2013. Physiological response to heat stress of tomato “Micro-Tom” plants expressing high and low levels of mitochondrial sHSP23.6 protein. *Plant Growth Regulation*, 70(2), pp.175–185.

Hwang, S. et al., 2015. Systematic comparison of variant calling pipelines using gold standard personal exome variants. *Scientific Reports*, 5.

Ichihashi, Y. & Sinha, N.R., 2014. From genome to phenome and back in tomato. *Current Opinion in Plant Biology*, 18(1), pp.9–15.

INRA, 2017. TOMEXPRESS. Available at: http://gbf.toulouse.inra.fr/tomexpress/www/welcomeTomExpress.php [Accessed February 13, 2017].

Ito, Y. et al., 2015. CRISPR/Cas9-mediated mutagenesis of the *RIN* locus that regulates tomato fruit ripening. *Biochemical and Biophysical Research Communications*, 467(1), pp.76–82.

Janssen, B.J., Drummond, R.S.M. & Snowden, K.C., 2014. Regulation of axillary shoot development. *Current Opinion in Plant Biology*, 17, pp.28–35.

Jędrszczyk, E. et al., 2012. The influence of weather conditions during vegetation period on yielding of twelve determinate tomato cultivars. 40(2), pp.203–209.

Jeon, J.-S. et al., 2000. *LEAFY HULL STERILE 1* is a homeotic mutation in a rice MADS box gene affecting rice flower development. *Plant Cell*, 12(6), pp.871–884.

Ji, Y., Scott, J., et al., 2007. Sources of resistance, inheritance, and location of genetic loci conferring resistance to members of the tomato-infecting begomoviruses. In H. Czosnek, ed. *Tomato Yellow Leaf Curl Virus Disease SE - 20*. Springer Netherlands, pp. 343–362.

Ji, Y., Scott, J.W., et al., 2007. Tomato Yellow Leaf Curl Virus Disease H. Czosnek, ed., Dordrecht: Springer Netherlands.

Jiang, L. et al., 2013. *DWARF 53* acts as a repressor of strigolactone signalling in rice. *Nature*, 504(7480), pp.401–405.

Just, D. et al., 2013. Micro-Tom mutants for functional analysis of target genes and discovery of new alleles in tomato. *Plant Biotechnology*, 30(3), pp.225–231.

Kaneko, M. et al., 2003. Where do gibberellin biosynthesis and gibberellin signaling occur in rice plants? *Plant Journal*, 35(1), pp.104–115.

Kang, M.-I. et al., 2004. Scaffolding of *Keap1* to the actin cytoskeleton controls the function of *Nrf2* as key regulator of cytoprotective phase 2 genes. *Proceedings of the National Academy of Sciences of the United States of America*, 101(7), pp.2046–2051.

Kaplinsky, N.J. & Freeling, M., 2003. Combinatorial control of meristem identity in maize inflorescences. *Development*, 130(6), pp.1149–1158.

Kaul, S. et al., 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408(6814), pp.796–815.

Kebrom, T.H., Brutnell, T.P. & Finlayson, S.A., 2010. Suppression of sorghum axillary bud outgrowth by shade, phyB and defoliation signalling pathways. *Plant, Cell and Environment*, 33(1), pp.48–58.

Kellogg, E.A., 2007. Floral displays: genetic control of grass inflorescences. *Current Opinion in Plant Biology*, 10(1), pp.26–31.

Kennedy, G.C. et al., 2003. Large-scale genotyping of complex DNA. *Nature Biotechnology*, 21(10), pp.1233–1237.

Kevei, Z. et al., 2015. Resequencing at ≥ 40-fold depth of the parental genomes of a *Solanum lycopersicum* x *S. pimpinellifolium* recombinant inbred line population and characterization of frame-shift indels that are highly likely to perturb protein function. *G3; Genes|Genomes|Genetics*, 5(5), pp.971–981.

Khamzina, A. et al., 2006. Assessing the potential of trees for afforestation of degraded landscapes in the aral sea basin of Uzbekistan. *Agroforestry Systems*, 66(2), pp.129–141.

Khlestkina, E.K. & Shumny, V.K., 2016. Prospects for application of breakthrough technologies in breeding: The CRISPR-Cas9 system for plant genome editing. *Russian Journal of Genetics*, 52(7), pp.676–687.

Kim, B. et al., 2016. Development of a high-resolution melting marker for selecting Fusarium crown and root rot resistance in tomato. *Genome*, 59(3), pp.171–183.

Kim, H.B. et al., 2006. The regulation of *DWARF4* expression is likely a critical mechanism in maintaining the homeostasis of bioactive brassinosteroids in *arabidopsis*. *Plant Physiology*, 140(2), pp.548–557

Kim, J.-E. et al., 2014. Genome-wide SNP calling using next generation sequencing data in tomato. *Molecules and Cells*, 37(1), pp.36–42.

Kim, S. et al., 2002. Solid phase capturable dideoxynucleotides for multiplex genotyping using mass spectrometry. *Nucleic acids research*, 30(16).

Kim, S. & Misra, A., 2007. SNP genotyping: technologies and biomedical applications. *Annual Review of Biomedical Engineering*, 9, pp.289–320.

Kirchoff, B.K. & Claßen-Bockhoff, R., 2013. Inflorescences: concepts, function, development and evolution. *Annals of Botany*, 112(8), pp.1471–1476.

Kirda, C. et al., 2004. Yield response of greenhouse grown tomato to partial root drying and conventional deficit irrigation. *Agricultural Water Management*, 69(3), pp.191–201.

Klepper, B., 1992. Development and growth of crop root systems. In J. L. Hatfield & B. A. Stewart, eds. *Limitations to Plant Root Growth*. New York, NY: Springer New York, pp. 1–25.

Van Kleunen, M. et al., 2005. Constraints on the evolution of adaptive phenotypic plasticity in plants. *New Phytologist*, 166(1), pp.49–60.

Knapp, S. et al., 2004. Solanaceae - A model for linking genomics with biodiversity. *Comparative and Functional Genomics*, 5(3), pp.285–291.

Kobayashi, M. et al., 2014. Genome-wide analysis of intraspecific dna polymorphism in “micro-tom”, a model cultivar of tomato (*Solanum lycopersicum*). *Plant and Cell Physiology*, 55(2), pp.445–454.

Kobayashi, Y. et al., 1999. A pair of related genes with antagonistic roles in mediating flowering signals. *Science*, 286(5446), pp.1960–1962.

Koelewijn, H.P. & Van Damme, J.M.M., 2005. Effects of seed size, inbreeding and maternal sex on offspring fitness in gynodioecious Plantago coronopus. *Journal of Ecology*, 93(2), pp.373–383.

Kohlen, W., Ruyter-Spira, C. & Bouwmeester, H.J., 2011. Strigolactones: A new musician in the orchestra of plant hormones. *Botany*, 89(12), pp.827–840.

Komatsu, M. et al., 2003. *FRIZZY PANICLE* is required to prevent the formation of axillary meristems and to establish floral meristem identity in rice spikelets. *Development*, 130(16), p.3841 LP-3850.

Krieger, U., Lippman, Z.B. & Zamir, D., 2010. The flowering gene *SINGLE FLOWER TRUSS* drives heterosis for yield in tomato. *Nature Genetics*, 42(5), pp.459–463.

Kroll, K.W. et al., 2014. Quality control for RNA-Seq (QuaCRS): an integrated quality control pipeline. *Cancer Informatics*, 13, pp.8–14.

Kwiatkowska, D., 2008. Flowering and apical meristem growth dynamics. *Journal of Experimental Botany*, 59(2), pp.187–201.

Labate, J. et al., 2007. Tomato. In C. Kole, ed. *Vegetables SE - 1*. Genome Mapping and Molecular Breeding in Plants. Springer Berlin Heidelberg, pp. 1–125. Available at: http://dx.doi.org/10.1007/978-3-540-34536-7_1.

Lambers, H., Martinoia, E. & Renton, M., 2015. Plant adaptations to severely phosphorus-impoverished soils. *Current opinion in plant biology*, 25, pp.23–31.

Landegren, U. et al., 1988. A ligase-mediated gene detection technique. *Science*, 241(4869), pp.1077–1080.

Langmead, B. & Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), pp.357–359.

Lapik, Y.R. & Kaufman, L.S., 2003. The *Arabidopsis* cupin domain protein *AtPirin1* interacts with the G protein alpha-subunit GPA1 and regulates seed germination and early seedling development. *The Plant cell*, 15(7), pp.1578–1590.

Larbat, R. et al., 2014. Phenolic characterization and variability in leaves, stems and roots of Micro-Tom and patio tomatoes, in response to nitrogen limitation. *Plant Science*, 224, pp.62–73.

Laskowski, M. et al., 2006. Expression profiling of auxin-treated *Arabidopsis* roots: Toward a molecular analysis of lateral root emergence. *Plant and Cell Physiology*, 47(6), pp.788–792

Laskowski, M.J. et al., 1995. Formation of lateral root meristems is a two-stage process. *Development*, 121(10), pp.3303–3310.

Latorra, D. et al., 2003. Enhanced allele-specific PCR discrimination in SNP genotyping using 3′ locked nucleic acid (LNA) primers. *Human Mutation*, 22(1), pp.79–85.

Laudencia-Chingcuanco, D. & Hake, S., 2002. The indeterminate floral *APEX1* gene regulates meristem determinacy and identity in the maize inflorescence. *Development*, 129(11), pp.2629–2638.

Lee, D.-Y. et al., 2007. The rice heterochronic gene *SUPERNUMERARY BRACT* regulates the transition from spikelet meristem to floral meristem. *Plant Journal*, 49(1), pp.64–78.

Lee, D. & An, G., 2012. Two AP2 family genes, *SUPERNUMERARY BRACT (SNB)* and *OsINDETERMINATE SPIKELET 1 (OsIDS1)*, synergistically control inflorescence architecture and floral meristem establishment in rice. *The Plant Journal*, 69(3), pp.445–461.

Lee, H.C. et al., 2012. Bioinformatics tools and databases for analysis of next-generation sequence data. *Briefings in Functional Genomics*, 11(1), pp.12–24.

van Leeuwen, H., Monfort, A. & Puigdomenech, P., 2007. Mutator-like elements identified in melon, *Arabidopsis* and rice contain *ULP1* protease domains. *Molecular Genetics and Genomics*, 277(4), pp.357–364.

Lemmon, Z.H. et al., 2016. The evolution of inflorescence diversity in the nightshades and heterochrony during meristem maturation. *Genome Research*, November (7).

Leung, C. et al., 2017. Real-time visualization of perforin nanopore assembly. *Nature Nanotechnology*.

Levene, H.J. et al., 2003. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science*, 299(5607), pp.682–686.

Levin, J.Z. & Meyerowitz, E.M., 1995. *UFO*: an *Arabidopsis* gene involved in both floral meristem and floral organ development. *The Plant cell*, 7(5), pp.529–548.

Lewis, D., 1953. Some factors affecting flower production in the tomato. *Journal of Horticultural Science*, 28(3), pp.207–220.

Leyser, O., 2005. The fall and rise of apical dominance. *Current Opinion in Genetics & Development*, 15(4), pp.468–471.

Li-xiang, Z. & Dan, L., 2014. Effects of different phosphorus applications on growth and platycodin content of balloon flower (*Platycodon grandiflorum*). *Scientia Horticulturae*, 178, pp.8–13.

Li, B. & Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, 12(1), p.323.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.

Li, H., 2011. Tabix: fast retrieval of sequence features from generic TAB-delimited files. *Bioinformatics*, 27(5), pp.718–719.

Li, H. et al., 2009. The Sequence Alignment/Map format and SAMtools. , 25(16), pp.2078–2079.

Li, J. et al., 2013. Integration of epigenetic and genetic controls of seed size by cytokinin in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), pp.15479–15484.

Li, N. & Li, Y., 2016. Signaling pathways of seed size control in plants. *Current Opinion in Plant Biology*, 33.

Li, R. et al., 2009. SNP detection for massively parallel whole-genome resequencing. *Genome Research*, 19(6), pp.1124–1132.

Liang, L. et al., 2017. DNA sequencing by two-dimensional materials: As theoretical modeling meets experiments. *Biosensors and Bioelectronics*, 89, pp.280–292.

Lifschitz, E. & Eshed, Y., 2006. Universal florigenic signals triggered by *FT* homologues regulate growth and flowering cycles in perennial day-neutral tomato. *Journal of experimental botany*, 57(13), pp.3405–3414.

Lin, K.-H. et al., 2006. RAPD markers for the identification of yield traits in tomatoes under heat stress via bulked segregant analysis. *Hereditas*, 143(2006), pp.142–154.

Lin, T. et al., 2014. Genomic analyses provide insights into the history of tomato breeding. *Nat Genet*, 46(11), pp.1220–1226.

Linkohr, B.I. et al., 2002. Nitrate and phosphate availability and distribution have different effects on root system architecture of *Arabidopsis*. *Plant Journal*, 29(6), pp.751–760.

Lippman, Z.B. et al., 2008. The making of a compound inflorescence in tomato and related nightshades. *PLoS Biology*, 6(11), pp.2424–2435.

Little, D.Y. et al., 2005. The putative high-affinity nitrate transporter *NRT2.1* represses lateral root initiation in response to nutritional cues. *Proceedings of the National Academy of Sciences of the United States of America*, 102(38), pp.13693–13698.

Liu, Q. et al., 2013. The involvement of auxin in root architecture plasticity in *Arabidopsis* induced by heterogeneous phosphorus availability. *Biologia Plantarum*, 57(4), pp.739–748.

Liu, W. et al., 1998. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. *Nucleic Acids Research*, 26(6), pp.1396–1400. A

Livak, K.J., 1999. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genetic Analysis - Biomolecular Engineering*, 14(5–6), pp.143–149.

Lizardi, P.M. et al., 1998. Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nature Genetics*, 19(3), pp.225–232.

Ljung, K., Bhalerao, R.P. & Sandberg, G., 2001. Sites and homeostatic control of auxin biosynthesis in *Arabidopsis* during vegetative growth. *The Plant Journal*, 28(4), pp.465–474.

Lombardi-Crestana, S. et al., 2012. The tomato (*Solanum Lycopersicum* cv. Micro-Tom) natural genetic variation *rg1* and the *della* mutant procera control the competence necessary to form adventitious roots and shoots. *Journal of Experimental Botany*, 63(15), pp.5689–5703.

López-Bucio, J. et al., 2002. Phosphate availability alters architecture and causes changes in hormone sensitivity in the *Arabidopsis* root system. *Plant Physiology*, 129(1), pp.244–256.

López-Bucio, J., Cruz-Ramírez, A. & Herrera-Estrella, L., 2003. The role of nutrient availability in regulating root architecture. *Current Opinion in Plant Biology*, 6(3), pp.280–287.

Lozano, R. et al., 2009. Genetic analysis of reproductive development in tomato. *International Journal of Developmental Biology*, 53(8–10).

Lu, F. et al., 2013. Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genetics*, 9(1).

Lunde, C.F. & Hake, S., 2005. Florets; rosettes: meristem genes in maize and *Arabidopsis*. *Maydica*, 50(3–4), pp.451–458.

Luo, H.H. et al., 2015. Response of cotton root growth and yield to root restriction under various water and nitrogen regimes. pp.384–392.

Lyamichev, V. et al., 1999. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nature Biotechnology*, 17(3), pp.292–296.

Mandadi, K.K. et al., 2009. BT2, a BTB protein, mediates multiple responses to nutrients, stresses, and hormones in *Arabidopsis*. *Plant Physiology*, 150(4), pp.1930–1939.

Mao, L. et al., 2000. *JOINTLESS* is a MADS-box gene controlling tomato flower abscission zone development. *Nature*, 406(6798), pp.910–913.

Marincs, F. et al., 2017. Large-scale amplicon sequencing of the *SP3D* gene responsible for fruit-yield heterosis in tomato. *Plant Gene*, 9.

Martí, E. et al., 2006. Genetic and physiological characterization of tomato cv. Micro-Tom. *Journal of Experimental Botany*, 57(9), pp.2037–2047.

Martín-Trillo, M. et al., 2011. Role of tomato *BRANCHED1*-like genes in the control of shoot branching. *Plant Journal*, 67(4), pp.701–714.

MASHIGUCHI, K. et al., 2009. Feedback-Regulation of Strigolactone Biosynthetic Genes and Strigolactone-Regulated Genes in *Arabidopsis*. *Bioscience, Biotechnology, and Biochemistry*, 73(11), pp.2460–2465.

Mason, M.G. et al., 2014. Sugar demand, not auxin, is the initial regulator of apical dominance. *Proceedings of the National Academy of Sciences*, 111(16), pp.6092–6097.

Matsuzaki, H. et al., 2004. Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. *Genome Research*, 14(3), pp.414–425.

Max, J.F.J. et al., 2016. Effects of shoot pruning and inflorescence thinning on plant growth, yield and fruit quality of greenhouse tomatoes in a tropical climate. *Journal of Agriculture and Rural Development in the Tropics and Subtropics*, 117(1).

McCormack, J.E. et al., 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, 66(2), pp.526–538.

McKenna, A. et al., 2010. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), pp.1297–1303.

McLaren, W. et al., 2010. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, 26(16), pp.2069–2070.

McSteen, P. & Zhao, Y., 2008. Plant hormones and signaling: common themes and new developments. *Developmental Cell*, 14(4), pp.467–473.

Medintz, I. et al., 2001. High-performance multiplex SNP analysis of three hemochromatosis-related mutations with capillary array electrophoresis microplates. *Genome Research*, 11(3), pp.413–421.

Meissner, R. et al., 1997. A new model system for tomato genetics. *Plant Journal*, 12(6), pp.1465–1472.

Memelink, J., 2009. Regulation of gene expression by jasmonate hormones. *Phytochemistry*, 70(13–14), pp.1560–1570.

Meng, L. & Feldman, L.J., 2010. *CLE14*/*CLE20* peptides may interact with *CLAVATA2*/*CORYNE* receptor-like kinases to irreversibly inhibit cell division in the root meristem of *Arabidopsis*. *Planta*, 232(5), pp.1061–1074.

Meng, X. et al., 2013. A MAPK cascade downstream of *ERECTA* receptor-like protein kinase regulates *Arabidopsis* inflorescence architecture by promoting localized cell proliferation. *Plant Cell*, 24(12).

Mero, C.E. & Honma, S., 1982. Inheritance of a *UNIFLORA* mutant in the tomato. *Journal of Heredity*, 73(1).

Michelmore, R.W., Paran, I. & Kesseli, R. V, 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the United States of America*, 88(21), pp.9828–9832.

Mistrik, I. & Mistrikova, B., 1995. Uptake, transport and metabolism of phosphate by individual roots of *Zea mays* L.

Miyawaki, K., Matsumoto-Kitano, M. & Kakimoto, T., 2004. Expression of cytokinin biosynthetic isopentenyltransferase genes in *Arabidopsis*: tissue specificity and regulation by auxin, cytokinin, and nitrate. *The Plant Journal*, 37(1), pp.128–138.

Mohammed, K.A.S. et al., 2016. Phenotypic plasticity of shoot and root traits of tomato in response to different rates of K, Mg and N supply. *International Journal of ChemTech Research*, 9(5).

Mohd-Yusoff, N.F. et al., 2015. Scanning the effects of ethyl methanesulfonate on the whole genome of *Lotus japonicus* using second-generation sequencing analysis. *G3: Genes|Genomes|Genetics*, 5(4), pp.559–567.

Molinero-Rosales, N. et al., 1999. *FALSIFLORA*, the tomato orthologue of *FLORICAULA* and *LEAFY*, controls flowering time and floral meristem identity. *Plant Journal*, 20(6), pp.685–693.

Mori, M. et al., 2002. Isolation and characterization of a rice dwarf mutant with a defect in brassinosteroid biosynthesis. *Plant Physiology*, 130(3), pp.1152–1161.

Morris, S.E. et al., 2005. Auxin dynamics after decapitation are not correlated with the initial growth of axillary buds. *Plant Physiology*, 138(3), pp.1665–1672.

Mouchel, C.F., Osmont, K.S. & Hardtke, C.S., 2006. *BRX* mediates feedback between brassinosteroid levels and auxin signalling in root growth. *Nature*, 443(7110), pp.458–461.

Moxon, S. et al., 2008. Deep sequencing of tomato short RNAs identifies microRNAs targeting genes involved in fruit ripening. *Genome Research*, 18(10), pp.1602–1609.

Muigai, S.G. et al., 2003. Greenhouse and field screening of wild *Lycopersicon* germplasm for resistance to the whitefly *Bemisia argentifolii*. *Phytoparasitica*, 31(1), pp.27–38.

Müller, J. et al., 2010. *Arabidopsis MPK6* is involved in cell division plane control during early root development, and localizes to the pre-prophase band,

phragmoplast, trans-Golgi network and plasma membrane. *The Plant Journal*, 61(2), pp.234–248.

Murtas, G. et al., 2003. A Nuclear Protease Required for Flowering-Time Regulation in *Arabidopsis* Reduces the Abundance of *SMALL UBIQUITIN-RELATED MODIFIER* Conjugates. *Plant Cell*, 15(10), pp.2308–2319.

Müssig, C., Fischer, S. & Altmann, T., 2002. Brassinosteroid-regulated gene expression. *Plant Physiology*, 129(3), pp.1241–1251.

Nadira, U.A. et al., 2014. The changes in physiological and biochemical traits of Tibetan wild and cultivated barley in response to low phosphorus stress. *Soil Science and Plant Nutrition*, 60(6), pp.832–842.

Nagel, O.W., Konings, H. & Lambers, H., 2001. Growth rate and biomass partitioning of wildtype and low-gibberellin tomato (*Solanum lycopersicum*) plants growing at a high and low nitrogen supply. *Physiologia Plantarum*, 111(1), pp.33–39.

Nakamura, A. et al., 2003. Brassinolide induces *IAA5*, *IAA19*, and *DR5*, a synthetic auxin response element in *Arabidopsis*, implying a cross talk point of brassinosteroid and auxin signaling. *Plant Physiology*, 133(4), pp.1843–1853.

Nakamura, H. et al., 2013. Molecular mechanism of strigolactone perception by *DWARF14*. *Nature Communications*, 4.

Nakano, A. et al., 2012. Automated analysis of fine-root dynamics using a series of digital images. *Journal of Plant Nutrition and Soil Science*, 175(5), pp.775–783.

Neff, M.M. et al., 1999. *BAS1*: A gene regulating brassinosteroid levels and light responsiveness in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 96(26), pp.15316–15323.

Nemhauser, J.L., Mockler, T.C. & Chory, J., 2004. Interdependency of brassinosteroid and auxin signaling in *Arabidopsis*. *PLoS Biology*, 2(9).

Nibau, C., Gibbs, D.J. & Coates, J.C., 2008. Branching out in new directions: the control of root architecture by lateral root formation. *New Phytologist*, 179(3), pp.595–614.

Nielsen, R. et al., 2011. Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, 12(6), pp.443–451.

Nikiforov, T.T. et al., 1994. Genetic bit analysis: A solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Research*, 22(20), pp.4167–4175.

Niklas, K.J., 2006. A phyletic perspective on the allometry of plant biomass-partitioning patterns and functionally equivalent organ-categories. *New Phytologist*, 171(1), pp.27–40.

Nilsson, M. et al., 1994. Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science*, 265(5181), pp.2085–2088.

Niu, Y.F. et al., 2013. Responses of root architecture development to low phosphorus availability: a review. *Annals of botany*, 112(2), pp.391–408.

Noguchi, T. et al., 1999. Brassinosteroid-insensitive dwarf mutants of *Arabidopsis* accumulate brassinosteroids. *Plant Physiology*, 121(3), pp.743–752.

Nordström, A. et al., 2004. Auxin regulation of cytokinin biosynthesis in *Arabidopsis thaliana*: A factor of potential importance for auxin–cytokinin-regulated development. *Proceedings of the National Academy of Sciences of the United States of America*, 101(21), pp.8039–8044.

Noulèkoun, F. et al., 2017. Shoot and root responses of woody species to silvicultural management for afforestation of degraded croplands in the Sudano-Sahelian zone of Benin. *Forest Ecology and Management*, 385, pp.254–263.

O'Driscoll, A., Daugelaite, J. & Sleator, R.D., 2013. "Big data", Hadoop and cloud computing in genomics. *Journal of Biomedical Informatics*, 46(5), pp.774–781.

O'Rawe, J. et al., 2013. Low concordance of multiple variant-calling pipelines: Practical implications for exome and genome sequencing. *Genome Medicine*, 5(3).

Obenchain, V. et al., 2014. Variant Annotation: A bioconductor package for exploration and annotation of genetic variants. *Bioinformatics*, 30(14), pp.2076–2078.

Ohyama, K. et al., 2009. A glycopeptide regulating stem cell fate in *Arabidopsis* thaliana. *Nature Chemical Biology*, 5(8), pp.578–580.

Okonechnikov, K., Conesa, A. & García-Alcalde, F., 2015. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics*, 32(2), pp.292–294.

Orozco-Nunnelly, D.A. et al., 2014. *Pirin1* (*PRN1*) is a multifunctional protein that regulates quercetin, and impacts specific light and uv responses in the seed-to-seedling transition of *Arabidopsis thaliana*. *PLoS ONE*, 9(4).

Orzaez, D., De Jong, A.J. & Woltering, E.J., 2001. A tomato homologue of the human protein *PIRIN* is induced during programmed cell death. *Plant Molecular Biology*, 46(4).

Osmont, K.S., Sibout, R. & Hardtke, C.S., 2007. Hidden branches: Developments in root system architecture. *Annual Review of Plant Biology*, 58, pp.93–113.

Osorio, S., Ruan, Y.-L. & Fernie, A.R., 2014. An update on source-to-sink carbon partitioning in tomato. *Frontiers in Plant Science*, 5(OCT).

Ozkan, B., Ceylan, R.F. & Kizilay, H., 2011. Energy inputs and crop yield relationships in greenhouse winter crop tomato production. *Renewable Energy*, 36(11), pp.3217–3221.

Ozsolak, F. et al., 2010. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell*, 143(6), pp.1018–1029.

Ozsolak, F. & Milos, P.M., 2011. RNA sequencing: Advances, challenges and opportunities. *Nature Reviews Genetics*, 12(2), pp.87–98.

Ozsolak, F. & Milos, P.M., 2011. Transcriptome profiling using single-molecule direct rna sequencing bt - high-throughput next generation sequencing: methods and applications. In Y. M. Kwon & S. C. Ricke, eds. Totowa, NJ: Humana Press, pp. 51–61.

Pabinger, S. et al., 2014. A survey of tools for variant analysis of next-generation genome sequencing data. *Briefings in Bioinformatics*, 15(2), pp.256–278.

Paran, I., Goldman, I. & Zamir, D., 1997. QTL analysis of morphological traits in a tomato recombinant inbred line population. *Genome*, 40(2), pp.242–248.

Pareek, C.S., Smoczynski, R. & Tretyn, A., 2011. Sequencing technologies and genome sequencing. *Journal of Applied Genetics*, 52(4), pp.413–435.

Park, S.J., Eshed, Y. & Lippman, Z.B., 2014. Meristem maturation and inflorescence architecture-lessons from the *Solanaceae*. *Current Opinion in Plant Biology*, 17(1), pp.70–71.

Pastinen, T. et al., 1997. Minisequencing: A specific tool for DNA analysis and diagnostics on oligonucleotide arrays. *Genome Research*, 7(6), pp.606–614.

Périlleux, C., Lobet, G. & Tocquin, P., 2014. Inflorescence development in tomato: gene functions within a zigzag model. *Frontiers in Plant Science*, 5(MAR).

Phillip, I.D.J., 1975. Apical Dominance. *Annual Review of Plant Physiology*, 26(1), pp.341–367.

Phillips, K., McCallum, N. & Welch, L., 2012. A comparison of methods for forensic DNA extraction: Chelex-10 ® and the QIAGEN DNA Investigator Kit (manual and automated). *Forensic Science International: Genetics*, 6(2).

Pino-Nunes, L.E. et al., 2009. Induced mutagenesis and natural genetic variation in tomato “micro-tom.” *Acta Horticulturae*, 821, pp.63–72.

Pino-Nunes, L.E. & Peres, L.E.P., 2009. Controle do desenvolvimento vegetal pela interação auxina-citocinina. Uma nova abordagem baseada no estudo de mutantes de tomateiro (*Solanum lycopersicum* cv Micro-Tom). Available at: http://www.teses.usp.br/teses/disponiveis/64/64133/tde-15102009-112355/pt-br.php.

Pino-Nunes, L.E. & Peres, L.E.P., 2005. Obtenção e uso de mutantes com alterações no balanço auxina/citocinina no estudo da competência organogênica em micro-tomateiro (*Lycopersicon esculentum* cv Micro-Tom). Available at: http://www.bdpi.usp.br/single.php?_id=001440484.

Pîrlea, S. et al., 2016. Permanganate-assisted removal of PCR inhibitors during the DNA Chelex extraction from stained denim samples. *International Journal of Legal Medicine*.

Pnueli, L. et al., 1998. The S*ELF-PRUNING* gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of *CEN* and *TFL1*. *Development*, 125(11), pp.1979–1989.

Pnueli, L. et al., 2001. Tomato *SP*-interacting proteins define a conserved signaling system that regulates shoot architecture and flowering. *The Plant Cell*, 13(12), pp.2687–2702.

Pons, C. et al., 2014. A bulk segregant gene expression analysis of a peach population reveals components of the underlying mechanism of the fruit cold response. *PLoS ONE*, 9(3).

Poorter, P., 2015. Commentary The limits to leaf and root plasticity : what is so special about specific root length? *Phytologist*, (2012), pp.1188–1190.

La Porta, N. et al., 2015. Identification of low temperature stress regulated transcript sequences and gene families in italian cypress. *Molecular Biotechnology*, 57(5), pp.407–418.

Poulsen, D.M.E. et al., 1995. The use of bulk segregant analysis to identify a RAPD marker linked to leaf rust resistance in barley. *Theoretical and Applied Genetics*, 91(2), pp.270–273.

Poyatos-Pertíñez, S. et al., 2016. Genetic interactions of the *unfinished flower development (ufd)* mutant support a significant role of the tomato *UFD* gene in regulating floral organogenesis. *Plant Reproduction*.

Premstaller, A. et al., 2002. Multiplex analysis of single-nucleotide extension products on a 16-capillary, denaturing, high-performance liquid chromatography array. *Genomics*, 79(6), pp.793–798.

Puchta, H., 2017. Applying CRISPR-Cas for genome engineering in plants: the best is yet to come. *Current Opinion in Plant Biology*, 36, pp.1–8.

Qi, X. et al., 2015. *De novo* transcriptome sequencing and analysis of *Coccinella septempunctata* L. in non-diapause, diapause and diapause-terminated states to identify diapause-associated genes. *BMC Genomics*, 16(1).

Qiang, Y. et al., 2013. *CLE* peptides in vascular development. *Journal of Integrative Plant Biology*, 55(4), pp.389–394.

Quinet, M. et al., 2006. Characterization of tomato (*Solanum lycopersicum* L.) mutants affected in their flowering time and in the morphogenesis of their reproductive structure. *Journal of Experimental Botany*, 57(6), pp.1381–1390.

Quinet, M., Kinet, J.-M. & Lutts, S., 2011. Flowering response of the *uniflora*:*Blind*:*self-pruning* and *jointless*:*uniflora*:*self-pruning* tomato (*Solanum lycopersicum*) triple mutants. *Physiologia Plantarum*, 141(2).

Quinlan, A.R. & Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), pp.841–842.

Ragoussis, J., 2009. Genotyping technologies for genetic research. *Annual Review of Genomics and Human Genetics*, 10, pp.117–133.

Rani, M.G. et al., 2013. Identification of new source for *Pi-1* blast gene for neck blast resistance through bulk segregant analysis in rice. *Research on Crops*, 14(3), pp.661–665.

Remans, T. et al., 2006. A central role for the nitrate transporter *NRT2.1* in the integrated morphological and physiological responses of the root system to nitrogen limitation in *Arabidopsis*. *Plant Physiology*, 140(3), pp.909–921.

Rick, C.M. & Butler, L., 1956. Cytogenetics of the tomato. *Advances in Genetics*, 8(C), pp.267–382.

Riefler, M. et al., 2006. *Arabidopsis* cytokinin receptors mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *Plant Cell*, 18(1), pp.40–54

Rigano, M.M. et al., 2013. Production of pharmaceutical proteins in *Solanaceae* food crops. *International Journal of Molecular Sciences*, 14(2).

Rimmer, A. et al., 2014. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nature genetics*, 46(8), pp.912–918.

Roach, D.A. & Wulff, R.D., 1987. Maternal effects in plants. *Annual review of ecology and systematics. Vol. 18*, pp.209–235.

Robbins, M.D. et al., 2011. Mapping and linkage disequilibrium analysis with a genome-wide collection of SNPs that detect polymorphism in cultivated tomato. *Journal of Experimental Botany*, 62(6).

Rogers, E.D. & Benfey, P.N., 2015. Regulation of plant root system architecture: implications for crop advancement. *Current Opinion in Biotechnology*, 32, pp.93–98.

Ronaghi, M. et al., 1996. Real-Time DNA Sequencing Using Detection of Pyrophosphate Release. *Analytical Biochemistry*, 242(1), pp.84–89.

Ross, P. et al., 1998. High level multiplex genotyping by MALDI-TOF mass spectrometry. *Nature Biotechnology*, 16(13), pp.1347–1351.

Sacco, A. et al., 2015. Exploring a tomato landraces collection for fruit-related traits by the aid of a high-throughput genomic platform. *PLoS ONE*, 10(9).

Sachs, T. & Thimann, K. V, 1967. The role of auxins and cytokinins in the release of buds from dominance. *Amer J Bot*, 54.

Saglam, N. & Yazgan, A., 1999. Effect of fruit number per truss on yield and quality in tomato. In *Acta Horticulturae*. International Society for Horticultural Science (ISHS), Leuven, Belgium, pp. 261–264.

Saito, T. et al., 2009. Mutant resources for the miniature tomato (*Solanum lycopersicum* L.) “Micro-Tom.” *Journal of the Japanese Society for Horticultural Science*, 78(1), pp.6–13.

Saleem, M.Y., Iqbal, Q. & Asghar, M., 2013. Genetic variability, heritability, character association and path analysis in F1 hybrids of tomato. *Pakistan Journal of Agricultural Sciences*, 50(4), pp.649–653.

Salunkhe, A.S. et al., 2011. Fine mapping QTL for drought resistance traits in rice (*Oryza sativa* L.) using bulk segregant analysis. *Molecular Biotechnology*, 49(1), pp.90–95.

Samach, A. et al., 1999. The *UNUSUAL FLORAL ORGANS* gene of *Arabidopsis thaliana* is an F-box protein required for normal patterning and growth in the floral meristem. *The Plant Journal*, 20(4), pp.433–445.

Samach, A. & Lotan, H., 2007. The transition to flowering in tomato. *Plant Biotechnology*, 24(1), pp.71–82.

Sanchez, J.J. et al., 2003. Multiplex PCR and minisequencing of SNPs - A model with 35 Y chromosome SNPs. *Forensic Science International*, 137(1), pp.74–84.

Sandhu, N. et al., 2014. Identification and mapping of stable QTL with main and epistasis effect on rice grain yield under upland drought stress. *BMC Genetics*, 15.

Santa-Cruz, A. et al., 2002. The rootstock effect on the tomato salinity response depends on the shoot genotype. *Plant Science*, 162(5), pp.825–831.

Santos, M., Szathmáry, E. & Fontanari, J.F., 2015. Phenotypic plasticity, the baldwin effect, and the speeding up of evolution: the computational roots of an illusion. *Journal of Theoretical Biology*, 371, pp.127–136.

Sauer, S. et al., 2002. Facile method for automated genotyping of single nucleotide polymorphisms by mass spectrometry. *Nucleic acids research*, 30(5), p.22.

Sawhney, V.K., 1983. The role of temperature and its relationship with gibberellic acid in the development of floral organs of tomato (*Lycopersicon esculentum*). *Canadian Journal of Botany*, 61(4), pp.1258–1265.

Schadt, E.E., Turner, S. & Kasarskis, A., 2010. A window into third-generation sequencing. *Human Molecular Genetics*, 19(R2).

Scheben, A., Batley, J. & Edwards, D., 2017. Genotyping-by-sequencing approaches to characterize crop genomes: choosing the right tool for the right application. *Plant Biotechnology Journal*, 15(2), pp.149–161.

Schmautz, Z. et al., 2016. Tomato productivity and quality in aquaponics: Comparison of three hydroponic methods. *Water (Switzerland)*, 8(11).

Schmieder, R. & Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, 27(6), pp.863–864.

Schmitz, G. et al., 2002. The tomato *BLIND* gene encodes a *MYB* transcription factor that controls the formation of lateral meristems. *Proceedings of the National Academy of Sciences of the United States of America*, 99(2), pp.1064–1069.

Schmitz, G. & Theres, K., 1999. Genetic control of branching in *Arabidopsis* and tomato. *Current Opinion in Plant Biology*, 2(1), pp.51–55.

Schubert, M., Lindgreen, S. & Orlando, L., 2016. AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. *BMC Research Notes*, 9(1).

Schwarz, D. et al., 2013. Rootstocks can enhance tomato growth and quality characteristics at low potassium supply. *Scientia Horticulturae*, 149, pp.70–79.

Scott, J.W., 2010. Phenotyping of tomato for SolCAP and onward into the void. *HortScience*, 45(9).

Selvin, P.R., 2000. The renaissance of fluorescence resonance energy transfer. *Nature Structural Biology*, 7(9), pp.730–734.

SGN, 2017. Sol Genomics Network. Available at: https://solgenomics.net/.

Shah, S.H. et al., 2016. Genetic Improvement of tomato (*Solanum lycopersicum*) with *AtDREB1A* gene for cold stress tolerance using optimized agrobacterium-mediated transformation system. *International Journal of Agriculture & Biology*, 18(3), pp.471–482.

Shannon, S. & Meeks-Wagner, D.R., 1991. A mutation in the *Arabidopsis TFL1* gene affects inflorescence meristem development. *Plant Cell*, 3(9), pp.877–892.

Sharp, R.E. & LeNoble, M.E., 2002. ABA, ethylene and the control of shoot and root growth under water stress. *Journal of Experimental Botany*, 53(366), pp.33–37.

Shi, J. et al., 2017. *ARGOS8* variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnology Journal*, 15(2), pp.207–216.

Shimada, Y. et al., 2001. Brassinosteroid-6-oxidases from *Arabidopsis* and tomato catalyze multiple C-6 oxidations in brassinosteroid biosynthesis1. *Plant Physiology*, 126(2), pp.770–779.

Shimizu-Sato, S. & Mori, H., 2001. Control of outgrowth and dormancy in axillary buds. *Plant Physiology*, 127(4), pp.1405–1413.

Shimizu-Sato, S., Tanaka, M. & Mori, H., 2008. Auxin–cytokinin interactions in the control of shoot branching. *Plant Molecular Biology*, 69(4), p.429.

Shirasawa, K. et al., 2010. SNP discovery and linkage map construction in cultivated tomato. *DNA Research*, 17(6), pp.381–391.

Signora, L. et al., 2001. ABA plays a central role in mediating the regulatory effects of nitrate on root branching in *Arabidopsis*. *Plant Journal*, 28(6), pp.655–662.

Sikder, S. et al., 2013. Induction of mutation in tomato (*Solanum lycopersicum L.*) by gamma irradiation and EMS. *Indian Journal of Genetics and Plant Breeding (The)*, 73(4), p.392.

Sim, S.-C. et al., 2015. Assessing the genetic variation in cultivated tomatoes (*Solanum lycopersicum* L.) using genome-wide single nucleotide polymorphisms. *Horticulture Environment and Biotechnology*, 56(6).

Sim, S.C., Durstewitz, G., et al., 2012. Development of a large SNP genotyping array and generation of high-density genetic maps in tomato. *PLoS ONE*, 7(7).

Sim, S.C., van Deynze, A., et al., 2012. High-density snp genotyping of tomato (*Solanum lycopersicum* L.) reveals patterns of genetic variation due to breeding. *PLoS ONE*, 7(9), pp.1–18.

Simons, J.L. et al., 2007. Analysis of the *DECREASED APICAL DOMINANCE* genes of Petunia in the control of axillary branching. *Plant Physiology*, 143(2), pp.697–706.

Skaar, J.R., Pagan, J.K. & Pagano, M., 2013. Mechanisms and function of substrate recruitment by F-box proteins. *Nat Rev Mol Cell Biol*, 14(6), pp.369–381.

Smet, I. et al., 2012. Analyzing lateral root development: how to move forward. *The Plant Cell*, 24(1), pp.15–20.

De Smet, I. et al., 2003. An abscisic acid-sensitive checkpoint in lateral root development of *Arabidopsis*. *Plant Journal*, 33(3), pp.543–555.

Snow, R., 1925. The correlative inhibition of the growth of axillary buds. *Annals of Botany*, os-39(4), pp.841–859.

Sokolov, B.P., 1990. Primer extension technique for the detection of single nucleotide in genomic DNA. *Nucleic Acids Research*, 18(12), p.3671.

Solanaceae project, 2017. International genome sequencing project (Feb). Available at: http://solgenomics.net/solanaceae-project/index.pl.

Soo, P.-C. et al., 2007. *Pirin* regulates pyruvate catabolism by interacting with the pyruvate dehydrogenase E1 subunit and modulating pyruvate dehydrogenase activity. *Journal of Bacteriology*, 189(1).

Sorefan, K. et al., 2003. *MAX4* and *RMS1* are orthologous dioxygenase-like genes that regulate shoot branching in *Arabidopsis* and pea. *Genes Dev*, 17.

Souer, E. et al., 2008. Patterning of inflorescences and flowers by the F-box protein DOUBLE TOP and the LEAFY homolog ABERRANT LEAF and FLOWER of Petunia. *Plant Cell*, 20(8).

Soyk, S. et al., 2017. Variation in the flowering gene *SELF PRUNING* 5G promotes day-neutrality and early yield in tomato. *Nature Genetics*, 49(1).

Spinelli, S. V et al., 2011. A mechanistic link between *STM* and *CUC1* during *Arabidopsis* development. *Plant Physiology*, 156(4), pp.1894–1904.

Sprink, T. et al., 2016. Regulatory hurdles for genome editing: process- vs. product-based approaches in different regulatory contexts. *Plant Cell Reports*, 35(7), pp.1493–1506.

Srivastava, V., Underwood, J.L. & Zhao, S., 2017. Dual-targeting by CRISPR-Cas9 for precise excision of transgenes from rice genome. *Plant Cell, Tissue and Organ Culture*, pp.1–8.

Steffens, B., Wang, J. & Sauter, M., 2006. Interactions between ethylene, gibberellin and abscisic acid regulate emergence and growth rate of adventitious roots in deepwater rice. *Planta*, 223(3), pp.604–612.

Steinert, J., Schiml, S. & Puchta, H., 2016. Homology-based double-strand break-induced genome engineering in plants. *Plant Cell Reports*, 35(7), pp.1429–1438.

Stirnberg, P., Zhao, S., et al., 2012. *FHY3* promotes shoot branching and stress tolerance in *Arabidopsis* in an *AXR1*-dependent manner. *The Plant Journal*, 71(6), pp.907–920.

Stirnberg, P., Liu, J.-P., et al., 2012. Mutation of the cytosolic ribosomal protein-encoding *RPS10B* gene affects shoot meristematic function in *Arabidopsis*. *BMC Plant Biology*, 12(1), pp.160-173.

Stofella, P.J., 1983. Root morphological characteristics of field-grown tomatoes. *HortScience*, 18(1), pp.70–72.

Suhl, J. et al., 2016. Advanced aquaponics: evaluation of intensive tomato production in aquaponics vs. conventional hydroponics. *Agricultural Water Management*, 178.

Sun, Y. et al., 2010. Efficiency of selective genotyping for genetic analysis of complex traits and potential applications in crop improvement. *Molecular Breeding*, 26(3), pp.493–511.

Sussex, I.M. & Kerk, N.M., 2001. The evolution of plant architecture. *Current Opinion in Plant Biology*, 4(1), pp.33–37.

Swartzberg, D., Hanael, R. & Granot, D., 2011. Relationship between hexokinase and cytokinin in the regulation of leaf senescence and seed germination. *Plant Biology*, 13(3), pp.439–444.

Swarup, R. et al., 2002. Auxin cross-talk: Integration of signalling pathways to control plant development. *Plant Molecular Biology*, 49(3–4), pp.411–426.

Szymkowiak, E.J. & Irish, E.E., 1999. Interactions between *JOINTLESS* and wild-type tomato tissues during development of the pedicel abscission zone and the inflorescence meristem. *Plant Cell*, 11(2), pp.159–175.

Tajima, R. & Kato, Y., 2013. A quick method to estimate root length in each diameter class using freeware ImageJ. *Plant Production Science*, 16(1), pp.9–11.

Tanaka, M. et al., 2006. Auxin controls local cytokinin biosynthesis in the nodal stem in apical dominance. *The Plant Journal*, 45(6), pp.1028–1036.

Tanaka, W. et al., 2013. Grass meristems II: Inflorescence architecture, flower development and meristem fate. *Plant and Cell Physiology*, 54(3), pp.313–324.

Teichmann, T. & Muhr, M., 2015. Shaping plant architecture. *Frontiers in Plant Science*, 6(APR), p.233.

Tezuka, T. et al., 2009. Construction of a linkage map and identification of DNA markers linked to *Fom-1*, a gene conferring resistance to *Fusarium oxysporum* f.sp. *melonis race 2* in melon. *Euphytica*, 168(2), pp.177–188.

TGRC, 2017. Tomato Genetics Resource Center. Available at: http://tgrc.ucdavis.edu/.

Thimann, K. V & Skoog, F., 1933. Studies on the growth hormone of plants: III. the inhibiting action of the growth substance on bud development. *Proceedings of the National Academy of Sciences of the United States of America*, 19(7), pp.714–716.

Thouet, J. et al., 2012. Repression of floral meristem fate is crucial in shaping tomato inflorescence. *PLoS ONE*, 7(2).

Tisdall, J.M. & Oades, J.M., 1982. Organic matter and water-stable aggregates in soils. *Journal of Soil Science*, 33(2), pp.141–163.

Tong, A.K. et al., 2001. Combinatorial fluorescence energy transfer tags for multiplex biological assays. *Nature Biotechnology*, 19(8), pp.756–759.

Tranchimand, S., Brouant, P. & Iacazio, G., 2010. The rutin catabolic pathway with special emphasis on quercetinase. *Biodegradation*, 21(6).

Truong, H.T.H. et al., 2015. Development of a SCAR marker linked to bacterial wilt (*Ralstonia solanacearum*) resistance in tomato line Hawaii 7996 using bulked-segregant analysis. *Horticulture Environment and Biotechnology*, 56(4), pp.506–515.

Turan, C. et al., 2015. New rapid DNA extraction method with Chelex from Venturia inaequalis spores. *Journal of Microbiological Methods*, 115.

Uitdewilligen, J.G.A.M.L. et al., 2013. A next-generation sequencing method for genotyping-by-sequencing of highly heterozygous autotetraploid potato. *PLOS ONE*, 8(5), pp.62355.

Umehara, M. et al., 2008. Inhibition of shoot branching by new terpenoid plant hormones. *Nature*, 455(7210), pp.195–200.

Vejchasarn, P., Lynch, J.P. & Brown, K.M., 2016. Genetic variability in phosphorus responses of rice root phenotypes. *Rice*, 9(1).

Vélez, J.M. et al., 2016. Current situation of wild *Solanum* spp. L. sect. Petota (*Solanum, Solanaceae*) in some Colombian regions. *Genetic Resources and Crop Evolution*, 63(1).

Venuprasad, R. et al., 2009. Identification and characterization of large-effect quantitative trait loci for grain yield under lowland drought stress in rice using bulk-segregant analysis. *Theoretical and Applied Genetics*, 120(1), pp.177–190.

Vercruyssen, L. et al., 2011. Combining enhanced root and shoot growth reveals cross talk between pathways that control plant organ size in *Arabidopsis*. *Plant Physiology*, 155(3), pp.1339–1352.

Vidyadhar, B. et al., 2015. Effect of growing conditions on growth, seed yield and quality attributes in cherry tomato (*Solanum lycopersicum* var cerasiferme). *Indian Journal of Agricultural Sciences*, 85(1), pp.114–117.

Vikram, P. et al., 2012. Bulk segregant analysis: “An effective approach for mapping consistent-effect drought grain yield QTLs in rice.” *Field Crops Research*, 134, pp.185–192.

del Viso, F. et al., 2012. Exon capture and bulk segregant analysis: Rapid discovery of causative mutations using high-throughput sequencing. *BMC Genomics*, 13(1).

Voland, M.L., 1987. Morphologic and genetic characterization of four tomato root mutants. New York: Cornell Unversity.

Voss-Fels, K. & Snowdon, R.J., 2016. Understanding and utilizing crop genome diversity via high-resolution genotyping. *Plant Biotechnology Journal*, 14(4), pp.1086–1094.

Wahid, P.A., 2000. A system of classification of woody perennials based on their root activity patterns. *Agroforestry Systems*, 49(2), pp.123–130.

Walsh, P.S., Metzger, D.A. & Higushi, R., 2013. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. BioTechniques 10(4): 506-13 (April 1991). *BioTechniques*, 54(3).

Wang, G., Zhang, G. & Wu, M., 2016. *CLE* peptide signaling and crosstalk with phytohormones and environmental stimuli. *Frontiers in Plant Science*, 6.

Wang, H. et al., 2012. An optimized DNA extraction method for costal cartilage by using protinase K and chelex-100. *Chinese Journal of Forensic Medicine*, 27(1).

Wang, J. et al., 2014. An Improved Simulation Model for Tomato Plant Based on Physiological Development Time. *2014 Seventh International Symposium on Computational Intelligence and Design*, pp.465–468.

Wang, K., Li, M. & Hakonarson, H., 2010. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16).

Wang, Q. et al., 2014. Auxin depletion from the leaf axil conditions competence for axillary meristem formation in *Arabidopsis* and tomato. *Plant Cell*, 26(5), pp.2068–2079.

Wang, Y. et al., 2014. The stem cell niche in leaf axils is established by auxin and cytokinin in *Arabidopsis*. *Plant Cell*, 26(5), pp.2055–2067.

Wang, Y. & Li, J., 2008. Molecular basis of plant architecture. *Annual Review of Plant Biology*, 59, pp.253–279.

Wang, Y. & Wu, W.-H., 2015. Genetic approaches for improvement of the crop potassium acquisition and utilization efficiency. *Current opinion in plant biology*, 25, pp.46–52.

Weigel, D. et al., 1992. *LEAFY* controls floral meristem identity in *Arabidopsis*. *Cell*, 69(5).

Weinig, C. & Delph, L.F., 2001. Phenotypic plasticity early in life constrains developmental responses later. *Evolution*, 55(5), pp.930–936.

Werner, T. et al., 2003. Cytokinin-deficient transgenic *Arabidopsis* plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem Activity. *Plant Cell*, 15(11), pp.2532–2550.

Werner, T. et al., 2001. Regulation of plant growth by cytokinin. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18), pp.10487–10492.

Widiatningrum, T. et al., 2015. A pirin-like protein from *Pseudomonas stutzeri* and its quercetinase activity. *Biochemistry and Biophysics Reports*, 3.

Widman, N. et al., 2014. Epigenetic differences between shoots and roots in *Arabidopsis* reveals tissue-specific regulation. *Epigenetics*, 9(2), pp.37–41.

Williamson, L.C. et al., 2001. Phosphate availability regulates root system architecture in *Arabidopsis*. *Plant Physiology*, 126(2), pp.875–882.

Wissuwa, M., Gamat, G. & Ismail, A.M., 2005. Is root growth under phosphorus deficiency affected by source or sink limitations? *Journal of Experimental Botany*, 56(417), pp.1943–1950.

Wolf, J.B. & Wade, M.J., 2009. What are maternal effects (and what are they not)? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1520), pp.1107–1115.

Xiao, W. & Oefner, P.J., 2001. Denaturing high-performance liquid chromatography: A review. *Human Mutation*, 17(6), pp.439–474.

Xu, C. et al., 2016. Control of inflorescence architecture in tomato by BTB/POZ transcriptional regulators. *Genes and Development*, 30(18).

Xu, J. & Zhang, S., 2015. Mitogen-activated protein kinase cascades in signaling plant growth and development. *Trends in Plant Science*, 20(1).

Yang, H. et al., 1999. Activation Tagging of the Floral Inducer FT. *Science*, 286(December), pp.1962–1965.

Yang, M. & Jiao, Y., 2016. Regulation of Axillary Meristem Initiation by Transcription Factors and Plant Hormones. *Frontiers in Plant Science*, 7, p.183.

Yoneyama, K. et al., 2012. How do nitrogen and phosphorus deficiencies affect strigolactone production and exudation? *Planta*, 235(6), pp.1197–1207.

Yoneyama, K. et al., 2013. Nitrogen and phosphorus fertilization negatively affects strigolactone production and exudation in sorghum. *Planta*, 238(5), pp.885–894. Available at: http://dx.doi.org/10.1007/s00425-013-1943-8.

Yoshida, A. et al., 2012. *ABERRANT SPIKELET AND PANICLE1*, encoding a TOPLESS-related transcriptional co-repressor, is involved in the regulation of meristem fate in rice. *The Plant Journal*, 70(2), pp.327–339.

Yu, B. et al., 2007. TAC1, a major quantitative trait locus controlling tiller angle in rice. *The Plant Journal*, 52(5), pp.891–898.

Žádníková, P. & Simon, R., 2014. How boundaries control plant development. *Current Opinion in Plant Biology*, 17, pp.116–125.

Zhang, D. & Yuang, Z., 2014. Molecular control of grass inflorescence development. *Annual Review of Plant Biology*, 65(1), pp.553–578.

Zhang, H. et al., 1999. Dual pathways for regulation of root branching by nitrate. *Proceedings of the National Academy of Sciences of the United States of America*, 96(11), pp.6529–6534.

Zhang, H. & Forde, B.G., 1998. An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. *Science*, 279(5349), pp.407–409.

Zhou, F. et al., 2013. *D14-SCFD3*-dependent degradation of D53 regulates strigolactone signalling. *Nature*, 504(7480), pp.406–410.

Zhou, Y.H. et al., 2009. Effects of phosphorus and chilling under low irradiance on photosynthesis and growth of tomato plants. *Biologia Plantarum*, 53(2), pp.378–382.

Zhu, P. et al., 2014. OTG-snpcaller: An optimized pipeline based on TMAP and GATK for SNP calling from Ion Torrent data. *PLoS ONE*, 9(5).

Zobel, R.W., 1973. Genetic and physiological investigations of two root mutants, University of Florida: Tomato Genetics Cooperative.

Zobel, R.W., 1972. Genetic and physiological investigations of two root mutants *in tomato, Lycopersion esculentum MILL.* California: University of California.

Zobel, R.W., 1991. Genetic control of root systems. In Y. Waisel, A. Eshel, & U. Kafkafi, eds. *Plant Roots, the hidden half.* Marcel Dekker, Inc., pp. 27–38.

Zobel, R.W., 2005. Primary and secondary root systems. In K. A. Barbarick, C. A. Roberts, & W. A. Dick, eds. *Roots and soil manangement: Interactions between roots and the soil.* American Society of Agronomy, Inc.; Crop Science Society of America, Inc; Soil Science Society of America, Inc; Madison, Wisconsin, USA., pp. 3–13.

Zook, J.M. et al., 2014. Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. *Nature Biotechnology*, 32(3), pp.246–251.

Zsögön, A. et al., 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.

APPENDICES

Appendix A :

Method for DNA extraction with Chelex 100 in 96-well format

A.1 Plate preparation

1. Prepare the extraction buffer (5% w/v Chelex): add 2.5 g of Chelex 100 (Bio-Rad, Hemel Hempstead, UK) into a 50 ml Falcon tube and make up to 50 ml with freshly obtained, deionized (MilliQ) water (autoclaving not required).
2. Add two 4 mm soda glass balls (Smith Scientific Ltd, Kent, UK) to each well of a 96-well deep-well plate (Starlabs, Milton Keynes, UK).

At room temperature, add a small piece (~20-30 mg) of young leaf tissue (e.g. small developing leaves from the apex which are less than 15 mm in length) to each well, placing it on top of the glass balls. Differences in the amount of plant material across the plate will affect final concentration (NB. Be careful with the labelling and plate orientation).

A.2 Tissue disruption and DNA extraction

1. Add 300 µl of the extraction buffer (Chelex 5% w/v) to each well;
2. Close the plate with a *sealing mat* (Starlabs) using a *seal applicator* (Starlabs). Alternatively, you can seal the plate with an adhesive seal (Starlabs), which is easier to use and is disposable.
3. Disrupt the tissue for 30 seconds at a frequency of 30 Hz;
4. Transfer 100 µl of supernatant to a PCR plate (Starlabs) and seal the wells using 8-strip PCR Caps (Starlabs).
5. Incubate the plate at 100^{o}C for 5 min in a PCR machine (MJ Research, Canada) with a heated lid to prevent evaporation.

6. Centrifuge (5430R, Eppendorf, Stevenage, UK) using a swing-out rotor at 3800 rpm for 4 minutes at ambient temperature or 20^{o}C.

Plates can be stored at 4^{o}C for short-term (< 1 month) or frozen at -20^{o}C for longer term. If using after storage, re-spin to bring down condensation and ensure the pellet is secure at the bottom of the tube.

Appendix B Transcriptomics data

	Gene ID	logFC	Average Expression	P value
Differently expressed gene in root tissues (*brt*NIL vs. MT)	Solyc12g014530	-0.5418533	2.076818	-0.004924884
	Solyc12g014540	-0.6101559	5.381991	-0.004924884
	Solyc12g014590	-0.8864737	3.076808	5.679906e-05
Differently expressed genes between root and leaf tissues (*brt*NIL vs. *brt*NIL)	Solyc12g014530	-1.536916	2.076818	2.292134e-06
	Solyc12g014540	0.4795581	5.381991	0.01629394
	Solyc12g014560	-0.8852245	2.643699	3.730434e-05
	Solyc12g014590	3.204554	3.076808	1.822179e-06
Differently expressed genes between root and leaf tissues (MT vs. MT)	Solyc12g014530	-0.7690019	2.076818	0.0005604215
	Solyc12g014540	0.8085703	5.381991	0.0003326477
	Solyc12g014560	-0.6221753	2.643699	0.0005505298
	Solyc12g014580	3.253546	-3.203736	0.004786132
	Solyc12g014590	4.082314	3.076808	1.410241e-07

Appendix C

C.1 Cold transfer experiment: summary of ANOVA analysis

	Flowers			
	1st truss		2nd truss	
	P	LSD (5%)	P	LSD (5%)
Genotype (G)	<.001	0.1120	<.001	0.0885
Treatment (T)	0.004	0.1120	0.446	0.0885
Transfer point (TP)	0.126	0.1771	0.180	0.1399
G x T	0.025	0.1584	0.838	0.1251
G x TP	0.574	0.2505	0.69	0.1978
T x TP	0.909	0.2505	<.001	0.1978
G x T x TP	0.611	0.3542	0.130	0.2798

	Branching points			
	1st truss		2nd truss	
	P	LSD (5%)	P	LSD (5%)
Genotype (G)	<.001	0.1431	<.001	0.1178
Treatment (T)	0.001	0.1431	0.549	0.1178
Transfer point (TP)	0.435	0.2263	0.610	0.1862
G x T	0.004	0.2024	0.635	0.1666
G x TP	0.646	0.3201	0.015	0.2633
T x TP	0.929	0.3201	<.001	0.2633
G x T x TP	0.749	0.4526	0.355	0.3724

Appendix D

D.1 Candidates for *BIFURICATE*

Gene	Polymorphisms	Apparent artefact	Expression data (TomExpress)
Solyc12g017930.1.1	Promoter	NO	-
Solyc12g017950.1.1	Intron	NO	-
Solyc12g017960.1.1	Exon	YES	-
Solyc12g017970.1.1	Exon	YES	-
Solyc12g018970.1.1	Exon	NO	-
Solyc12g018980.1.1	-	YES	-
Solyc12g018990.1.1	Promoter and Intron	NO	-
Solyc12g019000.1.1	Promoter and Intron	NO	-
Solyc12g019010.1.1	Intron and Exon	NO	Shoot; Flowers; Fruits
Solyc12g019020.1.1	Prom, Intr and Exon	NO	Flower buds; Mature fruits
Solyc12g019040.1.1	Prom, Intr and Exon	NO	Flowers; roots
Solyc12g019050.1.1	Promoter	NO	-
Solyc12g019060.1.1	Promoter	NO	-
Solyc12g019070.1.1	Promoter and Exon	YES	-
Solyc12g019080.1.1	Intron	NO	Flowers

Solyc12g019090.1.1	Intron and Exon	NO	Shoot; Fruit
Solyc12g019100.1.1	Prom, Intr and Exon	NO	Leaf; Fruit
Solyc12g019110.1.1	Intron	NO	-
Solyc12g019120.1.1	Promoter and Exon	NO	-
Solyc12g019130.1.1	Prom, Intr and Exon	NO	Root
Solyc12g019140.1.1	Prom, Intr and Exon	NO	Root
Solyc12g019150.1.1	Exon	YES	-
Solyc12g019160.1.1	Promoter and Intron	YES	-
Solyc12g019170.1.1	Prom, Intr and Exon	YES	-
Solyc12g019180.1.1	Prom, Intr and Exon	NO	-
Solyc12g019190.1.1	Prom, Intr and Exon	NO	Seeds; Fruit
Solyc12g019200.1.1	Prom, Intr and Exon	NO	Mature fruit
Solyc12g019210.1.1	Prom, Intr and Exon	YES	-
Solyc12g019220.1.1	Prom, Intr and Exon	NO	Root
Solyc12g019230.1.1	Prom, Intr and Exon	NO	Root
Solyc12g019240.1.1	Intron	NO	-
Solyc12g019260.1.1	Promoter and Intron	NO	-

Solyc12g019270.1.1	Prom, Intr and Exon	NO	-
Solyc12g019280.1.1	Prom, Intr and Exon	YES	-
Solyc12g019290.1.1	Promoter and Exon	YES	-
Solyc12g019300.1.1	Promoter	NO	-
Solyc12g019310.1.1	Prom, Intr and Exon	NO	Shoot; Root
Solyc12g019320.1.1	Prom, Intr and Exon	NO	Root; Leaf; Flower
Solyc12g019330.1.1	Promoter and Exon	YES	-
Solyc12g019340.1.1	Promoter and Exon	YES	-
Solyc12g019350.1.1	Promoter and Exon	NO	-
Solyc12g019360.1.1	Promoter and Exon	NO	-
Solyc12g019370.1.1	Promoter and Exon	YES	-
Solyc12g019380.1.1	Promoter and Exon	YES	-
Solyc12g019390.1.1	Promoter and Exon	YES	-
Solyc12g019400.1.1	Exon	YES	-
Solyc12g019410.1.1	Prom, Intr and Exon	NO	Root; LEAF
Solyc12g019420.1.1	Prom, Intr and Exon	NO	Flower; Fruit

Solyc12g019430.1.1	Prom, Intr and Exon	NO	Seed; Root tip; Fruit
Solyc12g019440.1.1	Prom, Intr and Exon	NO	-
Solyc12g019450.1.1	Prom, Intr and Exon	YES	-
Solyc12g019460.1.1	Prom, Intr and Exon	NO	Shoot and root
Solyc12g019470.1.1	Promoter and Exon	NO	-