(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 July 2012 (19.07.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/095841 A1

(51) International Patent Classification:
A01H 5/00 (2006.01) *C12N 15/82* (2006.01)
C07K 14/415 (2006.01)

(21) International Application Number:
PCT/IL2012/000016

(22) International Filing Date:
10 January 2012 (10.01.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11150498.1 10 January 2011 (10.01.2011) EP

(71) Applicant *(for all designated States except US)*: **STATE OF ISRAEL, MINISTRY OF AGRICULTURE AND RURAL DEVELOPMENT, A.R.O. - VOLCANI CENTER** [IL/IL]; P.O. Box 6, 50250 Beit Dagan (IL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SCHAFFER, Ari** [IL/IL]; Department of Vegetable Research, Agricultural Research Organization, Volcani Center, 50250 Bet Dagan (IL). **SHAMMAI, Arik** [IL/IL]; Department of Vegetable Research, Agricultural Research Organization, Volcani Center, 50250 Bet Dagan (IL).

(74) Agent: **DR. EYAL BRESSLER Ltd.**; 11 Tuval St., Lazrom House, 52522 Ramat Gan (IL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

[Continued on next page]

(54) Title: IMPROVED TOMATO PLANTS

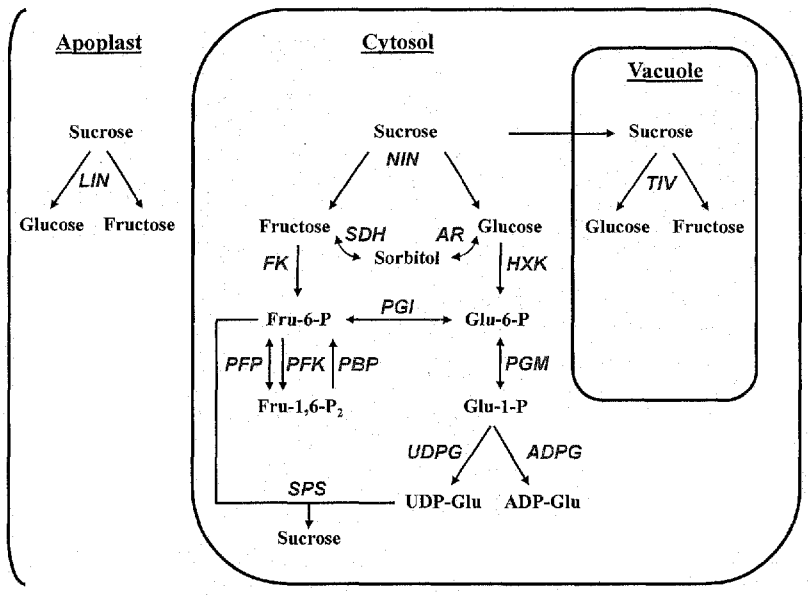


Figure 1

(57) Abstract: The present invention relates to a polynucleotide comprising a nucleic acid molecule encoding a MYB domain protein, a polynucleotide comprising a nucleic acid molcecule, an expression cassette comprising the polynucleotide, a vector molecule comprising the expression cassette and the use thereof. Moreover, the present invention provides for a method for improving fruit quality of tomato fruits, a transgenic tomato plant or transgenic tomato plant cell being genetically engineered, the use of the plant and a method of producing improved tomato fruits.

WO 2012/095841 A1

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

Improved tomato plants

The present invention generally relates to a polynucleotide encoding a protein which enables improved taste characteristics of tomato fruits. Moreover, method of producing the tomato fruits, method for improving fruit quality of tomato fruits, transgenic tomato plant and the use of the polynucleotide of the invention for producing tomato plants growing fruits with improved taste characteristics are also provided within the present invention.

The quality of tomato fruit is determined largely by its soluble sugar content, which, in the cultivated tomato (*Solanum lycopersicon*) is composed primarily of the hexoses, fructose and glucose, in near-equimolar quantities (Stevens et al., 1977; Davies and Hobson, 1981). Since fructose is approximately twice as sweet as glucose (Biester, 1925), genetically modifying hexose partitioning in the fruit towards increased fructose levels may contribute to the improvement of fruit taste.

The wild relatives of tomato have contributed, to date, a number of sugar modifying traits. For example, the trait of sucrose accumulation, determined by the *TIV* gene, has been transferred from *S. chmiliewskii* (Chetelat et al., 1995) and *S. habrochaites* (Hadas et al., 1995, Schaffer et al., 1998). Genetic traits of increased sugar levels were introduced from *S. pennellii* (*LIN5*, Fridman et al., 2000) and *S. habrochaites* (*AGPL1*, Petreikov, 2006, 2009). In the above examples, the biochemical and molecular differences between the alleles of the wild and cultivated species were identified, and the loci involved encoded for enzymes of carbohydrate metabolism, such as soluble acid invertase (*TIV,* Chetelat et al., 1995; Hadas et al., 1995; Miron et al., 2002), apoplastic acid invertase (*LIN5*, Fridman et al., 2000) and a subunit of ADP-glc pyrophosphorylase (*AGPL1,* Petreikov, 2006, 2009), the latter involved in starch synthesis.

The trait of high fructose to glucose ratio is extremely rare in fruits, as well as in nature, in general. Among the fruits, some cultivars of mango fruit stand out as accumulating high levels of fructose (Hubbard et al., 1991) and, outside of the plant world, high fructose accumulation characterizes the unique carbohydrate makeup of the seminal fluid of mammals, including humans (Frenette et al., 2006). The latter

phenomenon has been studied and it involves particular hexose transport mechanisms as well as a particular carbohydrate biochemistry, which includes the metabolism of glucose to fructose via the sugar alcohol intermediate, sorbitol, through the polyol pathway (Frenette et al., 2006)

S. habrochaites fruit accumulate primarily sucrose, with relatively low levels of hexose; however, its low hexose content is also characterized by an unusually high ratio of fructose to glucose (Davies, 1966). The trait of high fructose to glucose is independently inherited from that of sucrose accumulation (Stommel and Haynes, 1993; Schaffer et al., 1999) and hexose-accumulating tomatoes can be developed with an increased ratio of fructose to glucose (Schaffer et al., 2000). The trait was found to be determined by a major gene which is termed as *Fgr* (***f**ructose to **g**lucose **r**atio*) and breeding lines of tomato harboring the wild species *Fgr* allele had a significantly increased level of fructose, accompanied by a decrease in the glucose levels (Levin et al., 2000). The trait was mapped to the long arm of chromosome 4, in linkage to the *Adh1* locus (Levin et al., 2000).

Differential gene expression during tomato fruit development (using TOM1 microarrays) indicated that the sugar metabolism pathway in the developing fruit is modified in response to Fgr gene expression. The wild species allele led to increases in expression of two genes encoding for glucose metabolism enzymes, including a novel hexokinase (*HK5*) and the cytosolic phosphoglucomutase (PGM_{cyt}), and to a down-regulation of a gene encoding for a novel form of the fructose metabolism enzyme, pyrophosphate-dependent phosphofructokinase (*Pfpβ2*). Taken together, these concerted changes in transcription of the metabolic pathway of sugar metabolism can account for the function of the *Fgr* gene. The *Fgr* gene in tomato controls the ratio of the major soluble sugars, fructose to glucose, in the ripe fruit.

MYB genes controlling sugar metabolism, have been generally implicated in secondary metabolism, particularly phenylpropanoid and flavonoid pathways. For example, the analysis of the R2R3MYB families in Arabidopsis (Stracke et al., 2001; Yanhui et al., 2006), rice (Yanhui et al., 2006) and grape (Matus et al., 2008) show that of the more than 100 R2R3MYB genes present in each of the three genomes, only a small number have been functionally characterized. Many are involved in pigmentation patterns or other aspects of secondary metabolism (Du et al., 2009) or

in trichome and pollen development (Matus et al., 2008). However, the function of the MYB54 gene in Arabidopsis, for example, has not been reported down to the present day. It falls into the group II R2R3 MYBs which are characterized by an intron in the R2 sequence (Matus et al., 2008). However, homology is only at the R2R3 DNA binding region and there is little similarity between the C-term of the tomato *Fgr MYB* and other plant MYBs. *Fgr* in tomato, however, is the only characterized genetic trait for modified fructose accumulation in plants to date and, considering both the novelty of the fructose accumulation trait, as well as its potential importance in contributing to fruit quality, the identification and function of the *Fgr* gene is desirable. However, there are no reports of genetic variability for the particular increase in fructose levels in the cultivated tomato species. Further, the identity and function of the *Fgr* gene is unknown till the present day.

There is therefore an unmet need for providing extended options for producing fruits with improved taste characteristics based on a modification of fructose accumulation in plants.

This need is settled by providing herein the means for genetically modifying hexose partitioning in the fruit of tomato plants towards increased fructose levels. In particular, the present invention provides a polynucleotide comprising a nucleic acid molecule as defined by the features of independent claims 1 to 3, which polynucleotide encodes a MYB domain protein as defined by the features of independent claims 4 to 6. Further provided is a method for improving fruit quality of tomato fruits as defined by the features of independent claims 11 and 12, a transgenic tomato plant or transgenic tomato plant cell as defined by the features of independent claim 13 and a method of producing tomato fruit as defined by independent claim 15. Preferred embodiments are subject of the dependent claims. Moreover, it was now surprisingly found within the scope of the present invention that the *Fgr* alleles effects transcription of the sugar metabolism pathway in developing tomato plants and, thus, enables a control of a MYB transcription factor on primary sugar metabolism of tomato plants.

In a first embodiment, the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide

sequence encoding a polypeptide comprising the amino acid sequence as depicted in SEQ ID NO: 4.

In a second embodiment, the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence as depicted in SEQ ID NO: 3.

In a third embodiment, the invention relates to the polynucleotide of embodiment 2 comprising, consisting or consisting essentially of a nucleic acid sequence that has 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79% sequence identity to the sequence shown in SEQ ID NO: 3, which encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 4 of the invention relates to the polynucleotide of embodiment 2 comprising, consisting or consisting essentially of a nucleic acid sequence that has 80%; 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% sequence identity to the sequence shown in SEQ ID NO: 3, which encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 5 of the invention relates to the polynucleotide of embodiment 2 comprising, consisting or consisting essentially of a nucleic acid sequence that has 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% sequence identity to the sequence shown in SEQ ID NO: 3, which encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 6 of the invention relates to the polynucleotide of embodiment 2 comprising, consisting or consisting essentially of a nucleic acid sequence that has 95% sequence identity to the sequence shown in SEQ ID NO: 3, which encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 7 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence

the complementary strand of which hybridizes, particularly under stringent hybridization conditions, to the nucleic acid molecule of any of embodiments 1 to 6, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 8 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence that deviates from the nucleotide sequence defined in any of the embodiments 1 to 7 by the degeneracy of the genetic code, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 9 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence encoding a polypeptide comprising the amino acid sequence as depicted in SEQ ID NO: 8.

Embodiment 10 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence as depicted in SEQ ID NO: 6.

Embodiment 11 of the invention relates to the polynucleotide of embodiment 10 comprising, consisting or consisting essentially of a nucleotide sequence that has at least 60%; 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79% sequence identity to the sequence shown in SEQ ID NO: 6, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 12 of the invention relates to the polynucleotide of embodiment 10 comprising, consisting or consisting essentially of a nucleotide sequence that has at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% sequence identity to the sequence shown in SEQ ID NO: 6, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 13 of the invention relates to the polynucleotide of embodiment 10 comprising, consisting or consisting essentially of a nucleotide sequence that has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% sequence identity to the sequence shown in SEQ ID NO: 6, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 14 of the invention relates to the polynucleotide of embodiment 10 comprising, consisting or consisting essentially of a nucleotide sequence that has at least 95% sequence identity to the sequence shown in SEQ ID NO: 6, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 15 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence the complementary strand of which hybridizes, particularly under stringent hybridization conditions, to the nucleic acid molecule of any of embodiments 9 to 14, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 16 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence that deviates from the nucleotide sequence defined in any of the embodiments 9 to 15 by the degeneracy of the genetic code, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 17 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence encoding a polypeptide comprising the amino acid sequence as depicted in SEQ ID NO: 7.

Embodiment 18 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence as depicted in SEQ ID NO: 5.

Embodiment 19 of the invention relates to the polynucleotide of embodiment 18 comprising, consisting or consisting essentially of a nucleotide sequence that has at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% sequence identity to the sequence shown in SEQ ID NO: 5, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 20 of the invention relates to the polynucleotide of embodiment 18 comprising, consisting or consisting essentially of a nucleotide sequence that has at least 90% sequence identity to the sequence shown in SEQ ID NO: 5, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 21 of the invention relates to the polynucleotide of embodiment 18 comprising, consisting or consisting essentially of a nucleotide sequence that has at least 95% sequence identity to the sequence shown in SEQ ID NO: 5, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 22 of the invention relates to the polynucleotide of embodiment 18 comprising, consisting or consisting essentially of a nucleotide sequence that has at least 98% sequence identity to the sequence shown in SEQ ID NO: 5, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 23 of the invention relates to a polynucleotide comprising a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence the complementary strand of which hybridizes, particularly under stringent

hybridization conditions, to the nucleic acid molecule of any of the embodiments 18 to 23, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 24 of the invention relates to a nucleic acid molecule comprising, consisting or consisting essentially of a nucleotide sequence that deviates from the nucleotide sequence defined in any of the embodiments 18 to 24 by the degeneracy of the genetic code, which polynucleotide encodes a MYB domain protein and, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 25 of the invention relates to a MYB domain protein comprising a C-terminal protein interaction region comprising, consisting or consisting essentially of an amino acid sequence as depicted in SEQ ID NO: 8.

Embodiment 26 of the invention relates to the MYB domain protein of embodiment 25 comprising, consisting or consisting essentially of a C-terminal protein interaction region, which has 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59% sequence identity to the sequence shown in SEQ ID NO: 8, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 27 of the invention relates to the MYB domain protein of embodiment 26 comprising, consisting or consisting essentially of a C-terminal protein interaction region, which has 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69% sequence identity to the sequence shown in SEQ ID NO: 8, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 28 of the invention relates to the MYB domain protein of embodiment 26 comprising, consisting or consisting essentially of a C-terminal protein interaction region, which has 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79% sequence identity to the sequence shown in SEQ ID NO: 8, which protein, when

expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 29 of the invention relates to the MYB domain protein of embodiment 26 comprising, consisting or consisting essentially of a C-terminal protein interaction region, which has 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% sequence identity to the sequence shown in SEQ ID NO: 8, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 30 of the invention relates to the MYB domain protein of embodiment 26 comprising, consisting or consisting essentially of a C-terminal protein interaction region, which has 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% sequence identity to the sequence shown in SEQ ID NO: 8, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 31 of the invention relates to an amino acid sequence encoded by the polynucleotide of any of the embodiments 9 to 16.

Embodiment 32 of the invention relates to a MYB domain protein comprising a N-terminal DNA binding region comprising, consisting or consisting essentially of the amino acid sequence as depicted in SEQ ID NO: 7.

Embodiment 33 of the invention relates to the MYB domain protein of embodiment 34 comprising, consisting or consisting essentially of an N-terminal DNA binding region, which has 85%, 85%, 86%, 87%, 88%, 89% sequence identity to the sequence shown in SEQ ID NO: 7, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 34 of the invention relates to the MYB domain protein of any of the preceding embodiments comprising, consisting or consisting essentially of an N-terminal DNA binding region, which has 90%, 91%, 92%, 93%, 94%, 95%, 96%,

97%, 98%, 99% sequence identity to the sequence shown in SEQ ID NO: 7, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 35 of the invention relates to the MYB domain protein of any of the preceding embodiments comprising, consisting or consisting essentially of an N-terminal DNA binding region, which has 95% sequence identity to the sequence shown in SEQ ID NO: 7, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels..

Embodiment 36 of the invention relates to the MYB domain protein of any of the preceding embodiments comprising, consisting or consisting essentially of an N-terminal DNA binding region, which has 98% sequence identity to the sequence shown in SEQ ID NO: 7, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 37 of the invention relates to an amino acid sequence encoded by the polynucleotide of any of the embodiments 17 to 24.

Embodiment 38 of the invention relates to a MYB domain protein comprising, consisting or consisting essentially of an N-terminal DNA binding region, which has 90% sequence identity to the sequence shown in SEQ ID NO: 7 and a C-terminal protein interaction region, which has 70% sequence identity to the sequence shown in SEQ ID NO: 8, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 39 of the invention relates to a MYB domain protein of the preceding embodiment comprising, consisting or consisting essentially of a N-terminal DNA binding region, which has 95% sequence identity to the sequence shown in SEQ ID NO: 7 and a C-terminal protein interaction region, which has 90% sequence identity to the sequence shown in SEQ ID NO: 8, which protein, when expressed in a

tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 40 of the invention relates to the MYB domain protein of the preceding embodiment comprising, consisting or consisting essentially of a N-terminal DNA binding region, which has 98% sequence identity to the sequence shown in SEQ ID NO: 7 and a C-terminal protein interaction region, which has 95% sequence identity to the sequence shown in SEQ ID NO: 8, which protein, when expressed in a tomato plant and/or fruit from a polynucleotide encoding said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 41 of the invention relates to the MYB domain protein of embodiment 1, which has the sequence shown in SEQ ID NO: 2.

Embodiment 42 of the invention relates to the MYB domain protein of embodiment 1, which has the sequence shown in SEQ ID NO: 4.

Embodiment 43 of the invention relates to a polynucleotide comprising, consisting or consisting essentially of a nucleic acid sequence encoding a MYB domain protein of any of the preceding embodiments, which, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 44 of the invention relates to the polynucleotide, which has the sequence shown in SEQ ID NO: 1.

Embodiment 45 of the invention relates to the polynucleotide, which has the sequence shown in SEQ ID NO: 3.

Embodiment 46 of the invention relates to a polynucleotide comprising, consisting or consisting essentially of a nucleotide sequence, the complementary strand of which hybridizes, particularly under stringent hybridization conditions, to the polynucleotide of any one of embodiments 43 or 45, which polynucleotide, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 47 of the invention relates to the polynucleotide of any of the preceding embodiments comprising, consisting or consisting essentially of a nucleotide

sequence that deviates from the nucleotide sequence defined in any of the preceding embodiments by the degeneracy of the genetic code, which polynucleotide, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 48 of the invention relates to a polynucleotide comprising, consisting or consisting essentially of a nucleic acid sequence encoding a MYB domain protein of any of the preceding embodiments, which polynucleotide, upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 49 of the invention relates to an expression cassette comprising the polynucleotide of any of the preceding embodiments.

Embodiment 50 of the invention relates to the expression cassette of embodiment 49 comprising expression control sequences operably linked to said polynucleotide and capable of expressing in a plant a MYB domain protein of any of the preceding embodiments.

Embodiment 51 of the invention relates to a vector molecule comprising the expression cassette of embodiments 49 or 50.

Embodiment 52 of the invention relates to the use of a polynucleotide and/or a MYB domain protein of any of the preceding embodiments for modifying hexose partitioning in tomato fruits towards increased fructose levels.

Embodiment 53 of the invention relates to a method for improving fruit quality of tomato fruits by modifying hexose partitioning in the fruit towards increased fructose levels comprising transferring to and expressing in a tomato plant and/or fruit a polynucleotide, an expression cassette or a vector molecule of any of the preceding embodiments.

Embodiment 54 of the invention relates to the method of embodiment 53, wherein the polynucleotide, the expression cassette or the vector molecule of any of the preceding embodiments is transferred to the plant through plant transformation techniques.

Embodiment 55 of the invention relates to the method of embodiment 54 comprising (a) obtaining a first tomato plant verified to contain the polynucleotide of any of the preceding embodiments; (b) crossing said first tomato plant with a second tomato plant verified to lack the polynucleotide of any of the preceding embodiments; or to contain a polynucleotide comprising the sequence shown in SEQ ID NO. 1, and (c) identifying a tomato plant resulting from the cross exhibiting modified hexose partitioning in the fruit towards increased fructose levels and/or containing the polynucleotide of any of the preceding embodiments.

Embodiment 56 of the invention relates to the method of embodiment 55, wherein presence of the polynucleotide is verified by use of a molecular marker, particularly by a molecular marker physically located in a position on chromosome 4 that is within the genetic locus containing the polynucleotide.

Embodiment 57 of the invention relates to the method of embodiment 55, wherein presence of the polynucleotide is verified by use of molecular markers physically located in a position on chromosome 4 that is outside the genetic locus containing the polynucleotide.

Embodiment 58 of the invention relates to the method of any of embodiments 56 to 58, wherein the markers are selected from the group consisting of single nucleotide polymorphism (SNP) markers, indel (*i.e.*, insertions/deletions) markers, simple sequence repeat (SSR) markers, restriction fragment length polymorphism (RFLP) markers, random amplified polymorphic DNA (RAPD) markers, cleaved amplified polymorphic sequence (CAPS) markers, Diversity Arrays Technology (DArT) markers, and amplified fragment length polymorphisms (AFLPs) markers.

Embodiment 59 of the invention relates to a method for identifying in a tomato plant a polynucleotide of any of the preceding embodiments comprising using as a probe molecule a nucleotide sequence corresponding to a region which is part of the polynucleotide, particularly of the C-terminal part of the polynucleotide.

Embodiment 60 of the invention relates to the method of embodiment 59, wherein said polynucleotide probe is represented by at least one DNA molecule comprising, consisting or consisting essentially of a sufficient length of a stretch of contiguous nucleotides obtainable from a polynucleotide of any one of the preceding

embodiment, particularly a stretch of at least 10, particularly of at least 20, particularly of at lest 50 contiguous nucleotides, particularly from a polynucleotide having a sequence as shown in SEQ ID NO: 3, SEQ ID NO: 5, or SEQ ID NO: 6, and that hybridizes, particularly under high stringency hybridization conditions, with DNA of said polynucleotide of any one of the preceding embodiments, particularly with DNA of the polynucleotide as shown SEQ ID NO: 3, SEQ ID NO: 5, or SEQ ID NO: 6 and that does not hybridize under high stringency hybridization conditions with DNA of a polynucleotide having the sequence shown in SEQ ID NO: 1.

Embodiment 61 of the invention relates to a transgenic plant or transgenic plant cell which is genetically engineered with the polynucleotide, the expression cassette, or the vector of any of the preceding embodiments, and exhibits modified hexose partitioning in the fruit towards increased fructose levels.

Embodiment 62 of the invention relates to the plant or plant cell of embodiment 61, which is a tomato plant.

Embodiment 63 of the invention relates to the plant of embodiment 61 or 62, which is a hybrid plant, an inbred or a double haploid.

Embodiment 64 of the invention relates to propagation material or harvestable parts of the transgenic plant of any of embodiments 61 to 63 comprising the polynucleotide, the expression cassette, or the vector of any of the preceding embodiments, and exhibiting modified hexose partitioning in the fruit towards increased fructose levels.

Embodiment 65 of the invention relates to the propagation material of embodiment 64, which is plant seed.

Embodiment 66 of the invention relates to the propagation material of embodiment 65, which is hybrid seed.

Embodiment 67 of the invention relates to a tomato fruit grown on a plant according to embodiments 62 or 63 and exhibiting modified hexose partitioning in the fruit towards increased fructose levels.

Embodiment 68 of the invention relates to the use of a plant or of propagation material, particularly seeds, according to any of the preceding embodiments for

growing harvestable parts exhibiting modified hexose partitioning in the fruit towards increased fructose levels.

Embodiment 69 of the invention relates to the use of embodiment 68 for growing tomato fruits exhibiting modified hexose partitioning in the fruit towards increased fructose levels.

Embodiment 70 of the invention relates to a method of producing a plant with modified hexose partitioning in the fruit towards increased fructose levels comprising (a) obtaining a first plant verified to contain the polynucleotide of any of the preceding embodiments; (b) crossing said first tomato plant with a second tomato plant verified to lack the polynucleotide of any of the preceding embodiments; or to contain a polynucleotide comprising the sequence shown in SEQ ID NO. 1, and (c) identifying a tomato plant resulting from the cross exhibiting modified hexose partitioning in the fruit towards increased fructose levels and/or containing the polynucleotide of any of the preceding embodiments.

Embodiment 71 of the invention relates to the method of embodiment 70, wherein presence of the polynucleotide is verified by use of molecular markers physically located in a position on chromosome 4 that is within the genetic locus containing the polynucleotide.

Embodiment 72 of the invention relates to the method of embodiment 70, wherein presence of the polynucleotide is verified by use of molecular markers physically located in a position on chromosome 4 that is outside the genetic locus containing the polynucleotide.

Embodiment 73 of the invention relates to the method of any of embodiments 71 to 72, wherein the markers are selected from the group consisting of single nucleotide polymorphism (SNP) markers, indel (*i.e.*, insertions/deletions) markers, simple sequence repeat (SSR) markers, restriction fragment length polymorphism (RFLP) markers, random amplified polymorphic DNA (RAPD) markers, cleaved amplified polymorphic sequence (CAPS) markers, Diversity Arrays Technology (DArT) markers, and amplified fragment length polymorphisms (AFLPs) markers.

As described herein, there are several types of molecular markers that may be used in marker-based verification of the presence or absence of the polynucleotide within the invention.

In one embodiment (74), the invention relates to one or a plurality of probe molecules and/or to one or a plurality of primers, particularly one or a plurality of primer pairs, but especially one or a plurality of primer pairs consisting of a forward primer and a reverse primer, which primers are capable of annealing to a nucleotide sequence within a genomic region of the tomato genome that is genetically closely linked to the *fgr* gene, but particularly to a region within the *fgr* gene, and which comprises a polynucleotide according to the invention and as described herein before including an informative fragment thereof, wherein said fragment comprises a polymorphism, particularly a polymorphism that is based on an SNP, an SSR, a deletion or an insertion of at least one nucleotide, but especially a polymorphism based on an SNP, which polymorphism is diagnostic for the *fgr* allele and allows to discriminate between a tomato genotype containing the *fgr* allele from *Solanum lycopersicon* and a tomato genotype containing the *S. habrochaites* wild species allele.

In one embodiment (75) of the invention, a polynucleotide marker is provided which can be developed from a polynucleotide molecule or an informative fragment thereof selected from the group of polynucleotides as depicted in SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 6 and a polynucleotide encoding a polypeptide comprising a amino acid sequence as depicted in SEQ ID NO: 4, 7 or 8, wherein said polynucleotide comprises one or more polymorphisms, particularly a polymorphism that is based on an SNP, an SSR, a deletion or an insertion of at least one nucleotide, but especially a polymorphism based on an SNP, which polymorphism is diagnostic for the *fgr* allele and allows to discriminate between a tomato genotype containing the *fgr* allele from *Solanum lycopersicon* and a tomato genotype containing the *S. habrochaites* wild species allele.

In one embodiment (76) of the invention, the markers of the previously described embodiments can be used in an allelic discrimination assay or in a method for discriminating between different tomato genotypes containing the *fgr* allele from *Solanum lycopersicon* and the *S. habrochaites* wild species allele, respectively.

Embodiment (77) of the invention relates to a method of producing a plant with modified hexose partitioning in the fruit towards increased fructose levels comprising transforming a plant with the polynucleotide, the expression cassette, or the vector of any of the preceding embodiments, and expressing from said polynucleotide a MYB domain protein of any of the preceding embodiments.

Embodiment (78) of the invention relates to a method of producing tomato fruit with modified hexose partitioning in the fruit towards increased fructose levels comprising (a) growing a tomato plant according to any of the preceding embodiments; (b) allowing said tomato plant to set fruit; and (c) harvesting fruit of said tomato plant.

Embodiment (79) of the invention relates to a method of producing tomato seed capable of growing into a plant with modified hexose partitioning in the fruit towards increased fructose levels comprising (a) growing a tomato plant according to any of the preceding embodiments; (b) harvesting fruit of said tomato plant; and (c) extracting seed from said fruit.

In one embodiment (80) the invention relates to a method of suppressing the expression of an endogenous *fgr* gene of a tomato plant cell comprising introducing into said plant cell a first RNA strand and a second RNA strand, wherein said first RNA strand or, in the alternative, said second strand is sufficiently complimentary to at least a portion of an RNA strand of said endogenous *fgr* gene to hybridize or anneal to the RNA produced by the *fgr* gene such as to cause suppression of the expression of the endogeous fgr gene and said first RNA strand and said second RNA strand form a double stranded RNA, wherein said double stranded RNA participates in RNA interference of expression of said endogenous *fgr* gene.

Definitions

The technical terms and expressions used within the scope of this application are generally to be given the meaning commonly applied to them in the pertinent art of genetic engineering in plants, plant breeding and cultivation if not otherwise indicated herein below.

As used in this specification and the appended claims, the singular forms "a", "an", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a plant" includes one or more plants, and reference to "a cell" includes mixtures of cells, tissues, and the like.

A “cultivated tomato” plant is understood within the scope of the invention to refer to a plant that is no longer in the natural state but has been developed by human care and for human use and/or growing purposes and/or consumption. “Cultivated tomato plants” are further understood to exclude those wild-type species which comprise the trait being subject of this invention as a natural trait and/or part of their natural genetics.

An “allele” is understood within the scope of the invention to refer to alternative or variant forms of various genetic units identical or associated with different forms of a gene or of any kind of identifiable genetic element, which are alternative in inheritance because they are situated at the same locus in homologous chromosomes. Such alternative or variant forms may be the result of single nucleotide polymorphisms, insertions, inversions, translocations or deletions, or the consequence of gene regulation caused by, for example, by chemical or structural modification, transcription regulation or post-translational modification/regulation. In a diploid cell or organism, the two alleles of a given gene or genetic element typically occupy corresponding loci on a pair of homologous chromosomes.

An allele associated with a qualitative trait may comprise alternative or variant forms of various genetic units including those that are identical or associated with a single gene or multiple genes or their products or even a gene disrupting or

controlled by a genetic factor contributing to the phenotype represented by the locus.

As used herein, the term "marker allele" refers to an alternative or variant form of a genetic unit as defined herein above, when used as a marker to locate genetic loci containing alleles on a chromosome that contribute to variability of phenotypic traits.

As used herein, the term "breeding", and grammatical variants thereof, refer to any process that generates a progeny individual. Breedings can be sexual or asexual, or any combination thereof. Exemplary non-limiting types of breedings include crossings, selfings, doubled haploid derivative generation, and combinations thereof, which all are known techniques to the person skilled in the art.

"Backcrossing" is understood within the scope of the invention to refer to a process in which a hybrid progeny is repeatedly crossed back to one of the parents. Different recurrent parents may be used in subsequent backcrosses.

"Locus" is understood within the scope of the invention to refer to a region on a chromosome, which comprises a gene or any other genetic element or factor contributing to a trait.

As used herein, "marker locus" refers to a region on a chromosome, which comprises a nucleotide or a polynucleotide sequence that is present in an individual's genome and that is associated with one or more loci of interest, which may which comprise a gene or any other genetic element or factor contributing to a trait. "Marker locus" also refers to a region on a chromosome, which comprises a polynucleotide sequence complementary to a genomic sequence, such as a sequence of a nucleic acid used as probes.

For the purpose of the present invention, the term "segregation" or "co-segregation" refers to the fact that the allele for the trait and the allele(s) for the marker(s) tend to be transmitted together because they are physically close together on the same chromosome (reduced recombination between them because of their physical proximity) resulting in a non-random association of their alleles as a result of their proximity on the same chromosome. "Co-segregation" also refers to the presence

of two or more traits within a single plant of which at least one is known to be genetic and which cannot be readily explained by chance.

As used herein, the phrase "genetic marker" or "marker" refers to a feature of an individual's genome (*e.g.*, a nucleotide or a polynucleotide sequence that is present in an individual's genome) that is associated with one or more loci of interest. In some embodiments, a genetic marker is polymorphic in a population of interest, or the locus occupied by the polymorphism, depending on context. Genetic markers include, for example, single nucleotide polymorphisms (SNPs), indels (*i.e.*, insertions/deletions), simple sequence repeats (SSRs), restriction fragment length polymorphisms (RFLPs), random amplified polymorphic DNAs (RAPDs), cleaved amplified polymorphic sequence (CAPS) markers, Diversity Arrays Technology (DArT) markers, and amplified fragment length polymorphisms (AFLPs), among many other examples. Genetic markers can, for example, be used to locate genetic loci containing alleles on a chromosome that contribute to variability of phenotypic traits. The phrase "genetic marker" or "marker" can also refer to a polynucleotide sequence complementary to a genomic sequence, such as a sequence of a nucleic acid used as probes.

A "genetic marker" or "marker" can be physically located in a position on a chromosome that is within or outside of to the genetic locus with which it is associated (*i.e.*, is intragenic or extragenic, respectively). Stated another way, whereas "genetic markers" or "markers" are typically employed when the location on a chromosome of the gene or of a functional mutation, e.g. within a control element outside of a gene, that corresponds to the locus of interest has not been identified and there is a non-zero rate of recombination between the "genetic marker" or "marker" and the locus of interest, the presently disclosed subject matter can also employ "genetic markers" or "markers" that are physically within the boundaries of a genetic locus (*e.g.*, inside a genomic sequence that corresponds to a gene such as, but not limited to a polymorphism within an intron or an exon of a gene). In some embodiments of the presently disclosed subject matter, the one or more "genetic markers" or "markers" comprise between one and ten markers, and in some embodiments the one or more genetic markers comprise more than ten genetic markers.

As used herein, the term "genotype" refers to the genetic constitution of a cell or organism. An individual's "genotype for a set of genetic markers" includes the specific alleles, for one or more genetic marker loci, present in the individual's haplotype. As is known in the art, a genotype can relate to a single locus or to multiple loci, whether the loci are related or unrelated and/or are linked or unlinked. In some embodiments, an individual's genotype relates to one or more genes that are related in that the one or more of the genes are involved in the expression of a phenotype of interest. Thus, in some embodiments a genotype comprises a summary of one or more alleles present within an individual at one or more genetic loci of a quantitative trait. In some embodiments, a genotype is expressed in terms of a haplotype.

As used herein, the term "linkage", and grammatical variants thereof, refers to the tendency of alleles at different loci on the same chromosome to segregate together more often than would be expected by chance if their transmission were independent, in some embodiments as a consequence of their physical proximity.

The terms "protein," "peptide" and "polypeptide" are used interchangeably herein. The meaning of the terms “nucleic acid molecule“, “polynucleotide“, „polynucleotide sequence” and “polypeptide” is well known in the art, and the terms are, if not otherwise defined herein, used accordingly in the context of the present invention. For example, “polynucleotide sequence” as used herein refers to all forms of naturally occurring or recombinantly generated types of nucleic acids and/or nucleotide sequences as well as to chemically synthesized nucleic acids/nucleotide sequences. This term also encompasses nucleic acid analogs and nucleic acid derivatives such as, e. g., locked DNA, RNA, cDNA, PNA, oligonucleotide thiophosphates and substituted ribo-oligonucleotides. Furthermore, the term “polynucleotide sequence” also refers to any molecule that comprises nucleotides or nucleotide analogs. The phrase "nucleic acid" or “polynucleotide” refers to any physical string of monomer units that can be corresponded to a string of nucleotides, including a polymer of nucleotides (*e.g.*, a typical DNA, cDNA or RNA polymer), optionally containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers, modified oligonucleotides (*e.g.*, oligonucleotides comprising bases that are not typical to biological RNA or DNA, such as 2'-*O*-methylated oligonucleotides), and the like. The term "polynucleotide" is

understood herein to refer to polymeric molecule of high molecular weight which can be single-stranded or double-stranded, multi-stranded, or combinations thereof, composed of monomers (nucleotides) containing a sugar, phosphate and a base which is either a purine or pyrimidine. Unless otherwise indicated, a particular nucleic acid sequence of the presently disclosed subject matter optionally comprises or encodes complementary sequences, in addition to any sequence explicitly indicated.

Preferably, the term "polynucleotide sequence" refers to a nucleic acid molecule, i.e. deoxyribonucleic acid (DNA) and/ or ribonucleic acid (RNA). The "polynucleotide sequence" in the context of the present invention may be made by synthetic chemical methodology known to one of ordinary skill in the art, or by the use of recombinant technology, or may be isolated from natural sources, or by a combination thereof. The DNA and RNA may optionally comprise unnatural nucleotides and may be single or double stranded. "Polynucleotide sequence" also refers to sense and anti-sense DNA and RNA, that is, a polynucleotide sequence which is complementary to a specific sequence of nucleotides in DNA and/or RNA. Furthermore, the term "polynucleotide sequence" may refer to DNA or RNA or hybrids thereof or any modification thereof that is known in the state of the art (see, e.g., US 5525711, US 4711955, US 5792608 or EP 302175 for examples of modifications). The polynucleotide sequence may be single- or double-stranded, linear or circular, natural or synthetic, and without any size limitation. For instance, the polynucleotide sequence may be genomic DNA, cDNA, mRNA, antisense RNA, ribozymal or a DNA encoding such RNAs or chimeroplasts (Gamper, Nucleic Acids Research, 2000, 28, 4332 – 4339). Said polynucleotide sequence may be in the form of a plasmid or of viral DNA or RNA. "Polynucleotide sequence" may also refer to (an) oligonucleotide(s), wherein any of the state of the art modifications such as phosphothioates or peptide nucleic acids (PNA) are included.

A "polynucleotide fragment" is a fraction of a given polynucleotide molecule or of a "polynucleotide sequence". In higher plants, deoxyribonucleic acid (DNA) is the genetic material while ribonucleic acid (RNA) is involved in the transfer of information contained within DNA into proteins. A "genome" is the entire body of genetic material contained in each cell of an organism.

Unless otherwise indicated, a particular nucleic acid sequence of this invention also implicitly encompasses conservatively modified variants thereof (e.g. degenerate codon substitutions) and complementary sequences and as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, 1991; Ohtsuka *et al.*, 1985; Rossolini *et al.*, 1994). The term polynucleotide is used interchangeably with nucleic acid, nucleotide sequence and may include genes, cDNAs, and mRNAs encoded by a gene, etc.

The polynucleotide of the invention is understood to be provided in isolated form. The term "isolated" means that the polynucleotide disclosed and claimed herein is not a polynucleotide as it occurs in its natural context, if it indeed has a naturally occurring counterpart. Accordingly, the other compounds of the invention described further below are understood to be isolated. If claimed in the context of a plant genome, the polynucleotide of the invention is distinguished over naturally occurring counterparts by the insertion side in the genome and the flanking sequences at the insertion side.

As used herein, the term "gene" refers to any segment of nucleic acid associated with a biological function. Thus, genes include coding sequences and/or the regulatory sequences required for their expression. For example, gene refers to a nucleic acid fragment that expresses mRNA or functional RNA, or encodes a specific protein, and which includes regulatory sequences. Genes also include nonexpressed DNA segments that, for example, form recognition sequences for other proteins. Genes can be obtained from a variety of sources, including cloning from a source of interest or synthesizing from known or predicted sequence information, and may include sequences designed to have desired parameters.

"Marker-based selection" is understood within the scope of the invention to refer to e.g. the use of genetic markers to detect one or more nucleic acids from the plant, where the nucleic acid is associated with a desired trait to identify plants that carry genes for desirable (or undesirable) traits, so that those plants can be used (or avoided) in a selective breeding program. A "marker gene" encodes a selectable or screenable trait.

Suitable markers used within the invention may, for example, be selected from the group consisting of single nucleotide polymorphism (SNP) markers, indel (*i.e.*, insertions/deletions) markers, simple sequence repeat (SSR) markers, restriction fragment length polymorphism (RFLP) markers, random amplified polymorphic DNA (RAPD) markers, cleaved amplified polymorphic sequence (CAPS) markers, Diversity Arrays Technology (DArT) markers, and amplified fragment length polymorphisms (AFLPs) markers.

For example, RFLP involves the use of restriction enzymes to cut chromosomal DNA at specific short restriction sites, polymorphisms result from duplications or deletions between the sites or mutations at the restriction sites.

RAPD utilizes low stringency polymerase chain reaction (PCR) amplification with single primers of arbitrary sequence to generate strain-specific arrays of anonymous DNA fragments. The method requires only tiny DNA samples and analyses a large number of polymorphic loci.

AFLP requires digestion of cellular DNA with a restriction enzyme(s) before using PCR and selective nucleotides in the primers to amplify specific fragments. With this method, using electrophoresis techniques to visualize the obtained fragments, up to 100 polymorphic loci can be measured per primer combination and only small DNA sample are required for each test.

SSR analysis is based on DNA micro-satellites (short-repeat) sequences that are widely dispersed throughout the genome of eukaryotes, which are selectively amplified to detect variations in simple sequence repeats. Only tiny DNA samples are required for an SSR analysis. SNPs use PCR extension assays that efficiently pick up point mutations. The procedure requires little DNA per sample. One or two of the above methods may be used in a typical marker-based selection breeding program.

The most preferred method of achieving amplification of nucleotide fragments that span a polymorphic region of the plant genome employs the polymerase chain reaction ("PCR") (Mullis et al., Cold Spring Harbor Symp. Quant. Biol. 51:263 273 (1986)), using primer pairs involving a forward primer and a backward primer that are capable of hybridizing to the proximal sequences that define a polymorphism in

its double-stranded form. As disclosed herein, such primers may be used for fine mapping, map-based cloning and for expression analysis (see Tables 4 and 5).

Alternative methods may be employed to amplify fragments, such as the "Ligase Chain Reaction" ("LCR") (Barany, Proc. Natl. Acad. Sci. (U.S.A.) 88:189 193 (1991)), which uses two pairs of oligonucleotide probes to exponentially amplify a specific target. The sequences of each pair of oligonucleotides are selected to permit the pair to hybridize to abutting sequences of the same strand of the target. Such hybridization forms a substrate for a template-dependent ligase. As with PCR, the resulting products thus serve as a template in subsequent cycles and an exponential amplification of the desired sequence is obtained.

LCR can be performed with oligonucleotides having the proximal and distal sequences of the same strand of a polymorphic site. In one embodiment, either oligonucleotide will be designed to include the actual polymorphic site of the polymorphism. In such an embodiment, the reaction conditions are selected such that the oligonucleotides can be ligated together only if the target molecule either contains or lacks the specific nucleotide that is complementary to the polymorphic site present on the oligonucleotide. Alternatively, the oligonucleotides may be selected such that they do not include the polymorphic site (see, Segev, PCT Application WO 90/01069).

A further method that may alternatively be employed is the "Oligonucleotide Ligation Assay" ("OLA") (Landegren et al., Science 241:1077 1080 (1988)). The OLA protocol uses two oligonucleotides that are designed to be capable of hybridizing to abutting sequences of a single strand of a target. OLA, like LCR, is particularly suited for the detection of point mutations. Unlike LCR, however, OLA results in "linear" rather than exponential amplification of the target sequence.

Still another method that may alternatively be employed is the “Invader Assay” that uses a structure-specific flap endonuclease (FEN) to cleave a three-dimensional complex formed by hybridization of allele-specific overlapping oligonucleotides to target DNA containing a single nucleotide polymorphism (SNP) site. Annealing of the oligonucleotide complementary to the SNP allele in the target molecule triggers the cleavage of the oligonucleotide by cleavase, a thermostable FEN. Cleavage can be detected by several different approaches. Most commonly, the cleavage

product triggers a secondary cleavage reaction on a fluorescence resonance energy transfer (FRET) cassette to release a fluorescent signal. Alternatively, the cleavage can be detected directly by use of fluorescence polarization (FP) probes, or by mass spectrometry. The invasive cleavage reaction is highly specific, has a low failure rate, and can detect zeptomol quantities of target DNA. While the assay traditionally has been used to interrogate one SNP in one sample per reaction, novel chip- or bead-based approaches have been tested to make this efficient and accurate assay adaptable to multiplexing and high-throughput SNP genotyping.

Nickerson et al. have described a nucleic acid detection assay that combines attributes of PCR and OLA (Nickerson et al., Proc. Natl. Acad. Sci. (U.S.A.) 87:8923 8927 (1990)). In this method, PCR is used to achieve the exponential amplification of target DNA, which is then detected using OLA.

Schemes based on ligation of two (or more) oligonucleotides in the presence of a nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide, are also known (Wu et al., Genomics 4:560 569 (1989)), and may be readily adapted to the purposes of the present invention.

In one embodiment, a molecular marker is a DNA fragment amplified by PCR, e.g. a SSR marker or a RAPD marker. In one embodiment, the presence or absence of an amplified DNA fragment is indicative of the presence or absence of the trait itself or of a particular allele of the trait. In one embodiment, a difference in the length of an amplified DNA fragment is indicative of the presence of a particular allele of a trait, and thus enables to distinguish between different alleles of a trait.

In a specific embodiment of the invention simple sequence repeat (SSR) markers are used to identify invention-relevant alleles in the parent plants and/or the ancestors thereof, as well as in the progeny plants resulting from a cross of said parent plants. Simple sequence repeats are short, repeated DNA sequences and present in the genomes of all eukaryotes and consists of several to over a hundred repeats of a given nucleotide motif. Since the number of repeats present at a particular location in the genome often differs among plants, SSRs can be analyzed to determine the absence or presence of specific alleles.

"Microsatellite or SSRs (Simple sequence repeats) Marker" is understood within the scope of the invention to refer to a type of genetic marker that consists of numerous repeats of short sequences of DNA bases, which are found at loci throughout the plant's genome and have a likelihood of being highly polymorphic.

"PCR (Polymerase chain reaction)" is understood within the scope of the invention to refer to a method of producing relatively large amounts of specific regions of DNA or subset(s) of the genome, thereby making possible various analyses that are based on those regions.

"PCR primer" is understood within the scope of the invention to refer to relatively short fragments of single-stranded DNA used in the PCR amplification of specific regions of DNA.

"Phenotype" is understood within the scope of the invention to refer to a distinguishable characteristic(s) of a genetically controlled trait.

As used herein, the phrase "phenotypic trait" refers to the appearance or other detectable characteristic of an individual, resulting from the interaction of its genome, proteome and/or metabolome with the environment.

"Polymorphism" is understood within the scope of the invention to refer to the presence in a population of two or more different forms of a gene, genetic marker, or inherited trait or a gene product obtainable, for example, through alternative splicing, DNA methylation, etc.

"Probe" as used herein refers to a group of atoms or molecules which is capable of recognising and binding to a specific target molecule or cellular structure and thus allowing detection of the target molecule or structure. Particularly, "probe" refers to a labelled DNA or RNA sequence which can be used to detect the presence of and to quantitate a complementary sequence by molecular hybridization.

Polynucleotide sequences which are capable of to hybridizing with the polynucleotide sequences provided herein can for instance be isolated from genomic libraries or cDNA libraries of plants. Preferably, such polynucleotides are from plant origin, particularly preferred from a tomato plant such as *Solanum lycopersicum, Solanum lycopersicum var. cerasiforme, Solanum pimpinellifolium, Solanum cheesmaniae, Solanum neorickii, Solanum chmielewskii, Solanum*

habrochaites, Solanum pennellii, Solanum peruvianum, Solanum chilense, S. lycopersicoides, S. N peruvianum, S. corneliomuelleri, S. 'Callejon de Huaylas', S. galapagense a.d S. sitiens and *Solanum lycopersicum.* Alternatively, such nucleotide sequences can be prepared by genetic engineering or chemical synthesis.

Such polynucleotide sequences being capable of hybridizing may be identified and isolated by using the polynucleotide sequences described herein or parts or reverse complements thereof, for instance by hybridization according to standard methods (see for instance Sambrook and Russell (2001), Molecular Cloning: A Laboratory Manual, CSH Press, Cold Spring Harbor, NY, USA). Nucleotide sequences comprising the same or substantially the same nucleotide sequences as indicated in the listed SEQ ID NOs, or parts/fragments thereof, can, for instance, be used as hybridization probes. The fragments used as hybridization probes can also be synthetic fragments which are prepared by usual synthesis techniques, the sequence of which is substantially identical with that of a nucleotide sequence according to the invention.

The term "hybridize" as used herein refers to conventional hybridization conditions, preferably to hybridization conditions at which 5xSSPE, 1% SDS, 1xDenhardts solution is used as a solution and/or hybridization temperatures are between 35°C and 70°C, preferably 65°C. After hybridization, washing is preferably carried out first with 2xSSC, 1% SDS and subsequently with 0.2xSSC at temperatures between 35°C and 75°C, particularly between 45°C and 65°C, but especially at 59°C (regarding the definition of SSPE, SSC and Denhardts solution see Sambrook et al. loc. cit.). High stringency hybridization conditions as for instance described in Sambrook et al, supra, are particularly preferred. Particularly preferred stringent hybridization conditions are for instance present if hybridization and washing occur at 65°C as indicated above. Non-stringent hybridization conditions for instance with hybridization and washing carried out at 45°C are less preferred and at 35°C even less.

"Sequence Homology or Sequence Identity" is used herein interchangeably. The terms "identical" or percent "identity" in the context of two or more nucleic acid or protein sequences, refer to two or more sequences or subsequences that are the

same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection. For example, this term is used herein in the context of a nucleotide sequence or amino acid sequence/polypeptide/protein which has a homology, that is to say a sequence identity, of at least 50%, 55%, 60%, preferably of at least 70%, 75% more preferably of at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94, 95%, 96%, 97%, 98%, and even most preferably of at least 99% to another, preferably entire, nucleotide sequence or amino acid sequence, respectively. In other words, a polypeptide (being a MYB domain or fragment thereof) has at least 50%, 55% 60% preferably at least 70%, 75% more preferably at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94, 95%, 96%, 97%, 98%, and most preferably at least 99% identity/ homology to the polypeptide shown in SEQ ID No 2, 4, 5, 6, 7 and 8 respectively.

If two sequences which are to be compared with each other differ in length, sequence identity preferably relates to the percentage of the nucleotide residues of the shorter sequence which are identical with the nucleotide residues of the longer sequence. As used herein, the percent identity/homology between two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/ total # of positions x 100), taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm, as described herein below. For example, sequence identity can be determined conventionally with the use of computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive Madison, WI 53711). Bestfit utilizes the local homology algorithm of Smith and Waterman, Advances in Applied Mathematics 2 (1981), 482-489, in order to find the segment having the highest sequence identity between two sequences. When using Bestfit or another sequence alignment program to determine whether a particular sequence has for instance 95% identity with a

reference sequence of the present invention, the parameters are preferably so adjusted that the percentage of identity is calculated over the entire length of the reference sequence and that homology gaps of up to 5% of the total number of the nucleotides in the reference sequence are permitted. When using Bestfit, the so-called optional parameters are preferably left at their preset ("default") values. The deviations appearing in the comparison between a given sequence and the above-described sequences of the invention may be caused for instance by addition, deletion, substitution, insertion or recombination. Such a sequence comparison can preferably also be carried out with the program "fasta20u66" (version 2.0u66, September 1998 by William R. Pearson and the University of Virginia; see also W.R. Pearson (1990), Methods in Enzymology 183, 63-98, appended examples and http://workbench.sdsc.edu/). For this purpose, the "default" parameter settings may be used.

Another indication that two nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions. The phrase: "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA. "Bind(s) substantially" refers to complementary hybridization between a probe nucleic acid and a target nucleic acid and embraces minor mismatches that can be accommodated by reducing the stringency of the hybridization media to achieve the desired detection of the target nucleic acid sequence.

"Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments such as Southern and Northern hybridizations are sequence dependent, and are different under different environmental parameters. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes part I chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays" Elsevier, New York. Generally, highly stringent hybridization and wash conditions are selected to be about 5° C lower than the thermal melting point for the specific sequence at a

defined ionic strength and pH. Typically, under "stringent conditions" a probe will hybridize to its target subsequence, but to no other sequences.

The thermal melting point is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the melting temperature (T.sub.m) for a particular probe. An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formamide with 1 mg of heparin at 42°C., with the hybridization being carried out overnight. An example of highly stringent wash conditions is 0.1 5M NaCl at 72°C for about 15 minutes. An example of stringent wash conditions is a 0.2 times SSC wash at 65°C for 15 minutes (see, Sambrook, infra, for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, e.g., more than 100 nucleotides, is 1 times SSC at 45°C for 15 minutes. An example low stringency wash for a duplex of, e.g., more than 100 nucleotides, is 4-6 times SSC at 40°C for 15 minutes. For short probes (e.g., about 10 to 50 nucleotides), stringent conditions typically involve salt concentrations of less than about 1.0M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is typically at least about 30°C. Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide. In general, a signal to noise ratio of 2 times (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the proteins that they encode are substantially identical. This occurs, e.g. when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

A "plant" is any plant at any stage of development, particularly a seed plant.

A “plant part” as used herein refers to structural and/or functional sub-units of a plant including, but not limited to, plant cells, plant tissues, plant material, plant organs, harvestable plant parts, etc., as defined herein below.

A "harvestable plant part" is a part of a plant refers to those parts of the plant that are harvested at any suitable time and may be further processed for industrial use or consumption including flowers, fruits, leafs, seeds, fibers, etc.

A "plant cell" is a structural and physiological unit of a plant, comprising a protoplast and a cell wall. The plant cell may be in form of an isolated single cell or a cultured cell, or as a part of higher organized unit such as, for example, plant tissue, a plant organ, or a whole plant.

"Plant cell culture" means cultures of plant units such as, for example, protoplasts, cell culture cells, cells in plant tissues, pollen, pollen tubes, ovules, embryo sacs, zygotes and embryos at various stages of development.

"Plant material" or "plant material obtainable from a plant" refers to leaves, stems, roots, flowers or flower parts, fruits, pollen, egg cells, zygotes, seeds, cuttings, cell or tissue cultures, or any other part or product of a plant.

A "plant organ" is a distinct and visibly structured and differentiated part of a plant such as a root, stem, leaf, flower bud, or embryo.

"Plant tissue" as used herein means a group of plant cells organized into a structural and functional unit. Any tissue of a plant in planta or in culture is included. This term includes, but is not limited to, whole plants, plant organs, plant seeds, tissue culture and any groups of plant cells organized into structural and/or functional units. The use of this term in conjunction with, or in the absence of, any specific type of plant tissue as listed above or otherwise embraced by this definition is not intended to be exclusive of any other type of plant tissue.

As used herein, the term "population" means a genetically heterogeneous collection of plants sharing a common genetic derivation.

As used herein, the term "tomato" means any variety, cultivar, or population of *Solanum lycopersicum, Solanum lycopersicum var. cerasiforme, Solanum pimpinellifolium, Solanum cheesmaniae, Solanum neorickii, Solanum chmielewskii, Solanum habrochaites, Solanum pennellii, Solanum peruvianum, Solanum chilense, S. lycopersicoides, S. N peruvianum, S. corneliomuelleri, S. 'Callejon de Huaylas', S. galapagense a.d S. sitiens* and *Solanum lycopersicum.*

As used herein, the term "variety" or "cultivar" means a group of similar plants that by structural features and performance can be identified from other varieties within the same species.

In one embodiment, the present invention relates to a method for improving fruit quality of tomato fruits by modifying hexose partitioning in the fruit towards increased fructose levels comprising transferring to and expressing in a tomato plant and/or fruit a polynucleotide according to the present invention and as described herein but particularly a polynucleotide as shown in SEQ ID NOs: 1, 3, 5 and 6.

In one embodiment, this transfer is accomplished by using recombinant DNA technology and transformation techniques commonly known in the art.

As used herein, the term “plant transformation techniques” relates to the introduction of a transgene, conferring a specific trait, into the host plant, e.g. tomato plant. The transgene is incorporated into the host plant genome and stably inherited through future generations. The correct regulatory sequences are added to the gene of interest i.e. promoters and terminators, then the DNA is transferred to the plant cell culture using an appropriate vector. In some embodiments, the gene is attached to a selectable marker which allows selection for the presence of the transgene as described herein above. Once the plant tissue has been transformed, the cells containing the transgene are selected and regeneration back into whole plants is carried out.

Plant transformation can be carried out in a number of different ways for generating plant hosts with delegated functions required making plant host competent for expression of the sequence of interest, depending on the species of plant in question. For example, Agrobacterium mediated transformation may be used to transform plants according to the invention. Within this transformation method, plant or plant tissue (e.g. leaves) is cut into small pieces, e.g. 10x10mm, and soaked for 10 minutes in a fluid containing suspended Agrobacterium containing a Ti-plasmid vector carried by Agrobacterium (US 5,591,616; US 4,940,838; US 5,464,763). Placed on selectable rooting and shooting media, the plants will regrow.

Non-Agrobacterium techniques involve the uptake of exogenous genetic material directly by protoplasts or cells. These techniques include, but are not limited to, PEG or electroporation mediated uptake, particle bombardment-mediated delivery and microinjection. Examples of these techniques are described in Paszkowski et al., EMBO J 3, 2717-2722 (1984), Potrykus et al., Mol. Gen. Genet. 199,169-177 (1985), Reich et al., Biotechnology 4:1001-1004 (1986), and Klein et al., Nature 327,70-73 (1987). In each case, the transformed cells are regenerated to whole plants using standard techniques. For example, within particle bombardment particles of gold or tungsten are coated with DNA and then shot into young plant cells or plant embryos (US 05100792; EP 00444882B1; EP 00434616B1). Some genetic material will stay in the cells and transform them. This method may also be used to transform plant plastids. Further, electroporation technique may be used to transform plants according to the invention. During electroporation, transient holes are prepared in cell membranes using electric shock allowing DNA to enter the cell. Another transformation technique within the invention may be viral transformation (transduction). Here, the desired genetic material is packed into a suitable plant virus and allow this modified virus to infect the plant. If the genetic material is DNA, it can recombine with the chromosomes to produce transformed cells. However, genomes of most plant viruses consist of single stranded RNA which replicates in the cytoplasm of infected cell.

As used herein, the transformed plant or plant cell is capable of expressing or expresses the polypeptide encoded by the polynucleotide sequence of this invention. For example, the expression of a MYB domain protein from a polynucleotide according to the invention, lead to a modified hexose portioning in tomato fruits towards increased fruit levels and to tomato fruits with improved quality. In a specific embodiment, the "polypeptide" comprised in the said plant or said plant cell may be a heterologous with respect to the origin of the host cell. An overview of examples of different expression systems that may be used for generating the host cell of the present invention, for example the above-described particular one, is for instance contained in Glorioso et al. (1999), Expression of Recombinant Genes in Eukaryotic Systems, Academic Press Inc., Burlington, USA, Paulina Balbas und Argelia Lorence (2004), Recombinant Gene Expression:

Reviews and Protocols, Second Edition: Reviews and Protocols (Methods in Molecular Biology), Humana Press, USA.

The transformation or genetically engineering of the plant or plant cell with a nucleotide sequence or the vector according to the invention can be carried out by standard methods, as for instance described in Sambrook and Russell (2001), Molecular Cloning: A Laboratory Manual, CSH Press, Cold Spring Harbor, NY, USA. Moreover, the transgenic plant cell of the present invention is cultured in nutrient media meeting the requirements of the particular cell used, in particular in respect of the pH value, temperature, salt concentration, aeration, antibiotics, vitamins, trace elements etc.

The term "vector" or "vector molecule", as used herein, may comprise an expression cassette which may comprise expression control sequences operably linked to said polynucleotide and capable of expressing polypeptides such as MYB domain protein in a plant according to the present invention. The vector(s) may be in the form of a plasmid, and can be used alone or in combination with other plasmids, to provide transformed tomato plants, using transformation methods as described below to incorporate transgenes into the genetic material of the tomato plant(s). Further vectors may comprise cosmids, viruses, bacteriophages and other vectors commonly used in genetic engineering.

A "transgene" refers to a gene that has been introduced into the genome by transformation and is stably maintained. Transgenes may include, for example, genes that are either heterologous or homologous to the genes of a particular plant to be transformed. Additionally, transgenes may comprise native genes inserted into a non-native organism, or chimeric genes.

"Coding sequence" refers to a DNA or RNA sequence that codes for a specific amino acid sequence and excludes the non-coding sequences. It may constitute an "uninterrupted coding sequence", *i.e.*, lacking an intron, such as in a cDNA or it may include one or more introns bounded by appropriate splice junctions. An "intron" is a sequence of RNA which is contained in the primary transcript but which is removed through cleavage and re-ligation of the RNA within the cell to create the mature mRNA that can be translated into a protein.

The term “expression cassette”, as used herein, may be made up of one or more nucleotide sequences of the present invention in operable linkage with regulatory nucleotide sequences controlling their expression. As known in the art, an expression cassette may consist of a promoter sequence (promoter), an open reading frame encoding the MYB domain protein or a functional part thereof, a 3' untranslated region and a terminator sequence (terminator). The cassette may be part of a vector molecule as described herein above. Different expression cassettes can be transformed into plant or plant cells as long as the correct regulatory sequences are used.

As used herein, the term "promoter" refers to a nucleotide sequence, usually upstream (5') to its coding sequence, which controls the expression of the coding sequence by providing the recognition for RNA polymerase and other factors required for proper transcription. A “promoter” includes a minimal promoter that is a short DNA sequence comprised of a TATA box and other sequences that serve to specify the site of transcription initiation, to which regulatory elements are added for control of expression. Further, a "promoter" also refers to a nucleotide sequence that includes a minimal promoter plus regulatory elements that is capable of controlling the expression of a coding sequence or functional RNA. This type of promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a DNA sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue specificity of a promoter. It is capable of operating in both orientations (normal or flipped), and is capable of functioning even when moved either upstream or downstream from the promoter. Both enhancers and other upstream promoter elements bind sequence-specific DNA-binding proteins that mediate their effects. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even be comprised of synthetic DNA segments. A promoter may also contain DNA sequences that are involved in the binding of protein factors which control the effectiveness of transcription initiation in response to physiological or developmental conditions.

The "initiation site" is the position surrounding the first nucleotide that is part of the transcribed sequence, which is also defined as position +1. With respect to this site all other sequences of the gene and its controlling regions are numbered. Downstream sequences (*i.e.*, further protein encoding sequences in the 3' direction) are denominated positive, while upstream sequences (mostly of the controlling regions in the 5' direction) are denominated negative.

Promoter elements, particularly a TATA element, that are inactive or that have greatly reduced promoter activity in the absence of upstream activation are referred to as "minimal or core promoters". In the presence of a suitable transcription factor, the minimal promoter functions to permit transcription. A "minimal or core promoter" thus consists only of all basal elements needed for transcription initiation, *e.g.* a TATA box and/or an initiator.

According to the invention, the term "promoter" is a regulatory region of nucleic acid (e.g. DNA) driving the transcription of a MYB domain gene as depicted in SEQ ID NOS 1 and 3. Promoters are typically located near the genes they regulate, on the same strand and upstream (towards the 5' region of the sense strand). Several types of promoters are well known in the transformation arts, as are other regulatory elements that can be used alone or in combination with other promoters. A "plant promoter" is a promoter capable of initiating transcription in plant cells. Examples of promoters under developmental control include "Tissue-specific promoter" relating to regulated promoters that are not expressed in all plant cells but only in one or more cell types in specific organs (such as leaves or seeds), specific tissues (such as embryo or cotyledon), or specific cell types (such as leaf parenchyma or seed storage cells). These also include promoters that are temporally regulated, such as in early or late embryogenesis, during fruit ripening in developing seeds or fruit, in fully differentiated leaf, or at the onset of senescence.

A "cell type"-specific promoter or also called "inducible promoter" primarily drives expression in certain cell types in one or more organs, for example, vascular cells in roots or leaves by an external stimulus, such as a chemical, light, hormone, stress, or a pathogen.

"Constitutive expression" refers to expression using a constitutive or regulated promoter. "Conditional" and "regulated expression" refer to expression controlled by a regulated promoter.

"Constitutive promoter" refers to a promoter that is able to express the open reading frame (ORF) that it controls in all or nearly all of the plant tissues during all or nearly all developmental stages of the plant. Each of the transcription-activating elements do not exhibit an absolute tissue-specificity, but mediate transcriptional activation in most plant parts at a level of ?1% of the level reached in the part of the plant in which transcription is most active.

"Regulated promoter" refers to promoters that direct gene expression not constitutively, but in a temporally- and/or spatially-regulated manner, and includes both tissue-specific and inducible promoters. It includes natural and synthetic sequences as well as sequences which may be a combination of synthetic and natural sequences. Different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. New promoters of various types useful in plant cells are constantly being discovered, numerous examples may be found in the compilation by Okamuro *et al.* (1989). Typical regulated promoters useful in plants include but are not limited to safener-inducible promoters, promoters derived from the tetracycline-inducible system, promoters derived from salicylate-inducible systems, promoters derived from alcohol-inducible systems, promoters derived from glucocorticoid-inducible system, promoters derived from pathogen-inducible systems, and promoters derived from ecdysome-inducible systems.

"Operably-linked" refers to the association of nucleic acid sequences on single nucleic acid fragment so that the function of one is affected by the other. For example, a regulatory DNA sequence is said to be "operably linked to" or "associated with" a DNA sequence that codes for an RNA or a polypeptide if the two sequences are situated such that the regulatory DNA sequence affects expression of the coding DNA sequence (*i.e.*, that the coding sequence or functional RNA is under the transcriptional control of the promoter). Coding sequences can be operably-linked to regulatory sequences in sense or antisense orientation.

In the above context, the term "terminator" relates to a regulatory region of a DNA sequence that marks the end of gene on genomic DNA for transcription and, thus, initiates the termination of transcription of the respective gene.

"Expression", as used herein, refers to the transcription and/or translation of an endogenous gene, ORF or portion thereof, or a transgene in plants. For example, in the case of antisense constructs, expression may refer to the transcription of the antisense DNA only. In addition, expression refers to the transcription and stable accumulation of sense (mRNA) or functional RNA. Expression may also refer to the production of protein.

"Overexpression" relates to the level of expression in transgenic cells or organisms that exceeds levels of expression in normal or untransformed (nontransgenic) cells or organisms.

"Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of protein from an endogenous gene or a transgene.

"Gene silencing" refers to homology-dependent suppression of viral genes, transgenes, or endogenous nuclear genes. Gene silencing may be transcriptional, when the suppression is due to decreased transcription of the affected genes, or post-transcriptional, when the suppression is due to increased turnover (degradation) of RNA species homologous to the affected genes (English *et al.*, 1996). Gene silencing includes virus-induced gene silencing (Ruiz *et al.*, 1998).

As used herein, the term "BAC(s)" stands for bacterial artificial chromosome and defines a DNA construct, based on a functional fertility plasmid used for transforming and cloning in bacteria, such as *E. coli.* BACs may be used to sequence the genome of organisms such as plants. A short piece of the organism's DNA is amplified as an insert in BACs, and then sequenced. Finally, the sequenced parts are rearranged *in silico*, resulting in the genomic sequence of the organism.

In the description, examples and claims, reference is made to the following sequences:

Solanum lycopersicum DNA *[SEQ ID NO 1]*

ATGGCGGAAGGTGGTGGTTCCGGTGACGATACGACTAGAAGTAGCTGTCCACGTGGCCACTG
GCGACCGGCCGAAGACGAAAAACTCCGGCAGCTTGTAGAACAATATGGTCCTCAAAATTGGA
ACTCTATAGCTGAAAAACTTCAGGGAAGATCAGGGAAGAGTTGTAGATTAAGATGGTTTAAC
CAACTTGATCCAAGAATTAATAGAAGACCATTTAGTGAAGATGAAGAAGAGAAACTTATTGG
TGCTCATAGAATACATGGAAATAAATGGGCATTAATTTCAAGATTATTTCCAGGTAGAACTG
ATAATGCTGTGAAAAATCATTGGCATGTTTTAATGGCAAGAAAACAAAGGGAACAATCAAAG
ATTTGTGGTAAAAGAAGTTTATATCAACAACAACAACATGATAGTTTTCTAAGTGATTCTAA
ATCATCATGTTATGGTTTTCGACGAAGGAACAACAACAACAACAATAATAATATGAGAATAC
AAGAAGGTAATTATGGCTCCAAAATTAACTTTTTTGAATTTCAAAACCCTAATAAAGATAGG
GTTTTCTCAATGTCTACTACATATTCGTCTAGTTCACCTGAATTTTGTGGAAGAATTGGGAG
TCATTTGTTTAGAGAAAGCTCAATAGATCAAAAATCTTTACGTCAAAATAATTTAAGCTTTT
CAAGTCACGGAAGAGGAGGAGATAATAATAACTGCAAAAGGAGTACGACTTTTCAGAATCCT
TTCAGTTATGTTGATAGAAACGAATATGACAGTATAACTGAAAGAGTTGTGAATGTTAGCAA
TAGTACATTTTCCTTTGGAAAAATTCTCAAGGAAAACATTGAACAACAAAAATATGGAGAAG
AAGCAAGGGAAAAGAAAGATATTCCCTTTATAGATTTTCTTGGTGTGGGGATTTCTTCT

Solanum lycopersicum Protein *[SEQ ID NO 2]*

MAEGGGSGDDTTRSSCPRGHWRPAEDEKLRQLVEQYGPQNWNSIAEKLQGRSGKSCRLRWFN
QLDPRINRRPFSEDEEEKLIGAHRIHGNKWALISRLFPGRTDNAVKNHWHVLMARKQREQSK
ICGKRSLYQQQQHDSFLSDSKSSCYGFRRRNNNNNNNNMRIQEGNYGSKINFFEFQNPNKDR
VFSMSTTYSSSSPEFCGRIGSHLFRESSIDQKSLRQNNLSFSSHGRGGDNNNCKRSTTFQNP
FSYVDRNEYDSITERVVNVSNSTFSFGKILKENIEQQKYGEEAREKKDIPFIDFLGVGISS

Solanum habrochaites DNA *[SEQ ID NO 3]*

ATGGCGGAAGGTGGTGGTTCCGGTGACGATACAACTAGAAGTAGCTGTCCACGTGGCCACTG
GCGACCGGCCGAAGACGAAAAACTCCGGCAGCTTGTAGAACAATATGGTCCTCAAAATTGGA
ACTCTATAGCTGAAAAACTTCAAGGAAGATCAGGGAAGAGTTGTAGATTAAGATGGTTTAAC
CAACTTGATCCAAGAATTAATAGAAGACCATTTAGTGAAGATGAAGAAGAGAGACTTATTGG
TGCTCATAGAATACATGGAAATAAATGGGCATTAATTTCAAGATTATTTCCAGGTAGAACTG
ATAATGCTGTGAAGAATCATTGGCATGTTTTAATGGCAAGAAAACAAAGGGAACAATCAAAG
ATTTGTGGTAAAAGAAGTTTATATCAACAACAACAACATGATAATTTTCTAAGTGATTCTAA
ATCATCATGCTATGGTTTTCGACGAAGGAACAACAACAACATGAGAATACAAGAAGGTAAAA
TTAACTTTTTAGAATTTCAAAACCCTAATAAAGATAGGGTTTTCTCAATGTCTTCTAGTTCA
CCTGAATTTTGTGGAAGAATTGGGAGTCATTTGTTTAGAGAAAGCTCAATAGATCAAAAATC
TTTATGTCAAAATAATTTAAGCTTTGGAAGAGGAGGAGATAATGATAATTGCAAAAGGAGTA
TGACTTTTCAGAATCCTTTCAGTTATGTTGACAGTATAACTGAAAGAGTTGTGAATGTTAGC
AATAGTACATTTTCCTTTGCAAGAATTCTCAAGGAAAATATTCAACAACAAAAATATGGAGA
AGAAGCAAGGGAAAACAAAGATATTCCCTTTATAGATTTTCTTGGTGTGGGGATTTCTTCT

Solanum habrochaites protein *[SEQ ID NO 4]*

MAEGGGSGDDTTRSSCPRGHWRPAEDEKLRQLVEQYGPQNWNSIAEKLQGRSGKSCRLRWFN
QLDPRINRRPFSEDEEERLIGAHRIHGNKWALISRLFPGRTDNAVKNHWHVLMARKQREQSK
ICGKRSLYQQQQHDNFLSDSKSSCYGFRRRNNNNNMRIQEGKINFLEFQNPNKDRVFSMSSS
SPEFCGRIGSHLFRESSIDQKSLCQNNLSFGRGGDNDNCKRSMTFQNPFSYVDSITERVVNV
SNSTFSFARILKENIQQQKYGEEARENKDIPFIDFLGVGISS

N-Terminus of Solanum habrochaites DNA *[SEQ ID NO 5]*

ATGGCGGAAGGTGGTGGTTCCGGTGACGATACAACTAGAAGTAGCTGTCCACGTGGCCACTG
GCGACCGGCCGAAGACGAAAAACTCCGGCAGCTTGTAGAACAATATGGTCCTCAAAATTGGA
ACTCTATAGCTGAAAAACTTCAAGGAAGATCAGGGAAGAGTTGTAGATTAAGATGGTTTAAC
CAACTTGATCCAAGAATTAATAGAAGACCATTTAGTGAAGATGAAGAAGAGAGACTTATTGG
TGCTCATAGAATACATGGAAATAAATGGGCATTAATTTCAAGATTATTTCCAGGTAGAACTG
ATAATGCTGTGAAGAATCATTGGCATGTTTTAATGGCAAGAAAA

C-Terminus of Solanum habrochaites DNA *[SEQ ID NO 6]*

CAAAGGGAACAATCAAAGATTTGTGGTAAAAGAAGTTTATATCAACAACAACAACATGATAA
TTTTCTAAGTGATTCTAAATCATCATGCTATGGTTTTCGACGAAGGAACAACAACAACATGA
GAATACAAGAAGGTAAAATTAACTTTTTAGAATTTCAAAACCCTAATAAAGATAGGGTTTTC
TCAATGTCTTCTAGTTCACCTGAATTTTGTGGAAGAATTGGGAGTCATTTGTTTAGAGAAAG
CTCAATAGATCAAAAATCTTTATGTCAAAATAATTTAAGCTTTGGAAGAGGAGGAGATAATG
ATAATTGCAAAAGGAGTATGACTTTTCAGAATCCTTTCAGTTATGTTGACAGTATAACTGAA
AGAGTTGTGAATGTTAGCAATAGTACATTTTCCTTTGCAAGAATTCTCAAGGAAAATATTCA
ACAACAAAAATATGGAGAAGAAGCAAGGGAAAACAAAGATATTCCCTTTATAGATTTTCTTG
GTGTGGGGATTTCTTCT

N-Terminus of Solanum habrochaites protein *[SEQ ID NO 7]*

MAEGGGSGDDTTRSSCPRGHWRPAEDEKLRQLVEQYGPQNWNSIAEKLQGRSGKSCRLRWFN
QLDPRINRRPFSEDEEERLIGAHRIHGNKWALISRLFPGRTDNAVKNHWHVLMARK

C-Terminus of Solanum habrochaites protein *[SEQ ID NO 8]*

QREQSKICGKRSLYQQQQHDNFLSDSKSSCYGFRRRNNNNNMRIQEGKINFLEFQNPNKDRV
FSMSSSSPEFCGRIGSHLFRESSIDQKSLCQNNLSFGRGGDNDNCKRSMTFQNPFSYVDSIT
ERVVNVSNSTFSFARILKENIQQQKYGEEARENKDIPFIDFLGVGISS

The present invention is further described by reference to the following non-limiting figures, tables and examples.

Figure 1 shows sugar metabolism pathway in tomato fruit. Enzymes listed are LIN, apoplastic acid invertase; NIM, neutral invertase; FK, fructokinase; HK, hexokinase; TIV, soluble acid invertase; PGI, phosphoglucoisomerase; PGM, phosphoglucomutase; PFP, pyrophosphate-dependent phosphofructokinase; PFK, ATP- dependent phosphofructokinase; FBP, fructose-bisphosphatase; SPS, sucrose-phosphate synthase; SUS, sucrose synthase; UDPG, UDP-glc pyrophosphorylase; ADPG, UDP-glc pyrophosphorylase; AR, aldose reductase; SDH, sorbitol dehydrogenase; XI, xylose isomerase.

Figure 2 shows chromosome maps of tomato showing map locations of sugar metabolism genes. The red box indicates the map position of the Fgr trait.

Figure 3 shows fine mapping and positional cloning of the *Fgr* locus. *A*, Mapping of the Fgr gene using the initial markers TG208, Adh1, P74, U213812 and Asr2. Numbers to the left of the bars indicate the number of recombinant plants of each genotype and recombination distances are listed in cM. *B*, Fine mapping of the Fgr locus based on polymorphisms in the contig spanning TG208 to Adh1. Numbers on the left refer to the recombinant line which segregated for the gene. *C*, Region harboring the Fgr locus which did not show any recombinants in the recombination population of 6000 plants. White bars indicate the two ORFs present in the 39 Kb sequence.

Figure 4 shows fluorescence (left) and phase contrast (right) images of a tomato SC spread illustrating in situ hybridization with BACs LE_HBa0013P02 (green signal) and LE_HBa0208L16 (red signal) on the long arm of chromosome 4. On the right the fluorescent signals have been superimposed on the phase contrast image. The average positions of hybridization for the two BACs, 013P02 and 208L16, were calculated to be at 55.0% ± 1.6% and 56.7% ± 1.5% of the arm length away from the centromere, respectively. Scale bar 10 µm.

Figure 5 shows differential expression in tomato fruit of A) GRAS; B) R2R3MYB. Note the different scales in the two figures: the expression for the MYB gene is logarithmic. Numbers above bars indicate the expression relative to the expression of each gene at the IMG stage for the Fgr^{EE} genotype.

Figure 6 shows sequence alignment of R2R3MYB proteins from *S. lycopersicon* and *S. habrochaites*. Sequences are deposited as XX12345 and XX12346.

Figure 7 shows phylogenetic analysis of the DNA binding domain and the C-term domain of the R2R3Myb. Accession numbers of the protein sequences used are: Vitis vinifera; XP_002274206. Populus trichocarpa; XP_002307474. Pinus taeda; ABD60282. Picea glauca ABD60294. Ricinus communis; XP_002530307. Sorghum bicolor; XP_002450672. Zea mays; NP_001169060. Oryza sativa Japonica Group; BAD81105. Arabidopsis lyrata subsp. Lyrata; XP_002890254. Glycine max(soybean); ACU17230. Nicotiana Tabacum; FG154431, FG137759.

Figure 8 shows sugar content of the developing pericarp during fruit development. A), Total sugars; B), Glucose; C) Fructose; D) Ratio of fructose to glucose. Each data point is the average of three individual measurements of fruit from two near-isogenic lines.

Figure 9 shows QRT-PCR results for expression of *HK5*, *PGMcyt* and *Pfpb2* in tomato fruit pericarp at breaker stage. Plants were grown in different seasons and results presented are the averages and s.e. of three individual extractions for each bar value.

Figure 10 shows phylogenetic trees of HK proteins and PFP proteins from tomato and Arabidopsis.

Table 1 shows candidate genes of sugar metabolism in tomato fruit. Gene identification numbers refer to the TIGR database and chromosome positions are derived from V.1.5 of SGN.

Table 2 shows microarray results for sugar metabolism genes represented on the TOM1 microarray slides. (*) indicate genes for which at least three of the four replicate arrays at each stage showed consistent differential expression.

Table 3 shows expression data of the genes encoding for sugar metabolism enzymes in tomato plant. Data are derived from the Tomato Functional Genomics Database and number of ESTs is divided according to the following tissue libraries: vegetative (vgt), flower (flw), ovary (ovr), and fruit stages of IMG, MG, Br and Red.

Table 4 shows description of primers used for fine mapping and map-based cloning.

Table 5 shows primers used for expression analysis by quantitative on-line RT-PCR and their amplicon size.

The foregoing description will be more fully understood with the reference to the following Examples. Such Examples are, however, exemplary methods of practising the present invention and are not intended to limit the scope of the invention.

The following Examples illustrate the invention:

Example 1: Materials and Methods

A) Plant materials and measurements

Two independent *Fgr*^{HH} lines, 2591 and 3515, were used in this study for the initial mapping of the *FGR* gene. Both lines are derived from the interspecific cross of *L. hirsutum* (accession LA1777; available at the Tomato Genetics Resource Center (TGRC) of University of California, Davis, http://tgrc.ucdavis.edu/) and an *S. lycopersicon* male-sterile breeding line developed, as described in Levin *et al.* (2000), followed by five backcrosses to *S. lycopersicon* breeding lines. The two lines contain overlapping introgressions and this allowed for the gross mapping of the trait.

For fine-mapping of the *FGR* gene a segregating population of approximately 6,000 plants derived from selfs of BC_5F_2 plants was genotyped using DNA markers (Table 4). Twenty seven recombinants and their selfed populations were used for further mapping.

Plants were grown in the greenhouse under standard conditions, as previously described (Miron and Schaffer, 1991). Genomic DNA was extracted according to standard methods (Fulton et al., 2000).

B) Chromosome walking and fine mapping

For physical mapping of the *FGR* gene, sets of several tomato genomic libraries filters (Table 3) were screened with radioactive probes labeled using the NEBlot™ Kit (#N1500S) (New England BioLabs, Inc.) according to the supplier's instructions. Initially the Le_HBa BAC library set of filters (Clemson University Genomic Institute, USA) were screened and if these did not give useful results, additional libraries were

screened. Labeled BAC colonies on the filter were detected using a Fuji Film phosphoimager (FLA-5000). For the development of the contiguous BAC series, the selected BACs were end-sequenced using SP6 and T7 primers and a PCR product was developed from these end sequences. The purified PCR product was labeled as above and used as a probe for the identification of the contiguous BAC. Two BAC ends, LA483_92F2-T7 and LA483_83C22-T7, both from a *S. cheesmaniae* library, did not identify additional BACs. At this time the initial results of the sequencing project were made available and two scaffolds from the pre-version were identified. The end sequences of these scaffolds allowed for the identification of BAC Le_HBa213A03 which completed the contig.

C) Isolation of total RNA

For the production of RNA for both microarray analysis and quantitative RT-PCR, tomato fruit pericarp was ground to a fine powder in liquid nitrogen and stored at 80°C or immediately used for RNA extraction. Total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) and digested with DNase I enzyme (Qiagen) according to the manufacturer's instructions.

D) Microarray

Experimental design

Two biological and two technical replicates were performed for each of the three developmental stages studied (Immature green, Mature Green and Breaker). Independent RNA extracts were obtained from at least 4 fruits from 4 individual plants of each of the near-isogenic lines (two FGR^{EE} and two FGR^{HH}, described above). Sample labeling was performed with dye-swap (Cy3 and Cy5) to avoid dye bias, as described below. TOM1 microarray chips (Cornell University, Ithaca, NY, USA) were used for the hybridization experiments and results presented are derived from 12 high quality arrays.

E) Experiment procedure

For hybridization to microarrays, total RNA (3µg) was taken for mRNA amplification using the MessageAmp™ II amplification kit (Ambion, Austin, TX, USA). The amplified RNA (aRNA) was then used for cDNA synthesis, with 1 mM aminoallyl-

dUTP included for labelling. Traces of aRNA were removed by alkaline hydrolysis using 10µl NaOH (1M), incubated at 70°C for 15 min and neutralized by 10µl HCl (1M). The cDNA was then ethanol-precipitated using glycogen as a carrier, and the resultant pellet obtained post-centrifugation was air dried at room temperature. The aminoallyl-incorporated cDNA was labelled with cyanine-3 / cyanine-5 (Cy3/Cy5) fluorescent dye using the LabelStar ArrayKit (Amersham, Uppsala, Sweden). Post-labelling, the fluorescently labelled cDNA was cleaned of excess dye and nucleotides with a QIAquick PCR Purification Kit (Qiagen). The amount of incorporated fluorescent dye was calculated by measuring the absorbance at A550/A260 and A650/A260. The samples were stored at -20°C until further use. Microarray slides were incubated at 76°C for 2 min, overlaid with 60µl of pre-hybridization buffer (48% formamide, 3.2 X SSPE, 0.4% SDS, 2 X Denhardts and 0.2 µg/µl salmon sperm), covered with an RNase-free hybridslip and incubated for an additional 1 h at 50°C. The slides were washed with sterile dd water, then with 70% ethanol followed by 100% ethanol, and dried. Each chip was hybridized by mixing two samples of Cy3- and Cy5-labelled probes (each 100pM); the volume was reduced to 8.5µl using a Microcon YM30 filter (Millipore, Cambridge, UK). The following reagents were then added: 1.5µl 20 X SSPE, 1µl poly dA (15 µg/µl), 1µl yeast tRNA (4 µg/µl), 24µl hybridization buffer (62.8% formamide, 0.8% SDS, 4 X Denhardt's solution and 5 X SSPE). The probe solution was applied to the slide using HybriSlip covers after denaturation at 98°C for 3 min and centrifugation for 1 min at 13,000g. The slides were wrapped in aluminium foil and incubated at 42°C overnight. Slides were washed twice with 2X SSC / 0.1X SDS, followed by washing once with 0.2X SSC / 0.1X SDS, and dried.

F) Quantification and data analysis

The slides were scanned for fluorescence emission using a DNA microarray scanner (Agilent Technologies, Santa Clara, CA, USA). The spot intensities were quantified using ScanAlyze (Michael Eisen, Stanford University 1999). Background fluorescence values were calculated and subtracted from all feature intensities prior to ratio calculation. Data normalization was performed using the overall background by applying per-spot and per-chip normalization, and the default function (GeneSpring 5.1; Silicon Genetics, Redwood City, CA, USA). Filtration on

confidence was performed based on a one-sample t-test ($P \leq 0.05$), and the resulting gene lists were filtered for differentially expressed genes (above 1.5 and below 0.66 fold change in FGR HH vs. FGR EE samples). Raw data was collected and imported into Microsoft Excel for further analysis.

G) Quantitative RT-PCR

Total RNA was used as a template for first strand cDNA synthesis with the Super-script II pre-amplification system reverse transcriptase kit (Gibco BRL Life Technologies) at 42°C according to the supplier's instructions. Specific primers with short amplicons for on-line quantitative PCR were designed based on the sequence derived from genomic sequence of the ORFs and reference genes (Table 5). The cDNA was used as the template for quantitative PCR amplification using the Rotor-Gene 6000™ (Corbett Life Science, NSW, Australia) with QPCR SYBR® Green Mix (Thermo Fisher scientific Inc, UK), according to the manufacturer's instructions. Standards containing logarithmically increasing known levels of cDNA were run with each set of primers, in addition to the actin primers for normalization. All real-time PCR products were confirmed by sequencing.

H) FISH analysis

FISH analysis using labelled BAC clones was performed as in Chang et al. (2007)

I) Soluble sugar determination

Fruit pericarp portions of about 1 g fresh weight were extracted in hot 80% ethanol and analyzed by HPLC, as described in Petreikov et al. (2009). The HPLC column used for separation was a BioRad Fast Carbohydrate Column, used according to manufacturer's directions.

Example 2: Candidate genes of the metabolic pathway of hexose metabolism in tomato

The initial attempt to clone the *Fgr* gene was by the strategy of candidate gene mapping and co-localization with the trait. *Fgr* was previously mapped to the distal portion of chromosome 4, in proximity to the *Adh1* marker (Levin et al., 2000). The

candidate genes tested comprised those encoding for the enzymes of the sugar metabolism pathway, particularly of the reactions surrounding hexose metabolism (Fig. 1). In order to collect the complete array of sugar metabolism genes expressed in tomato fruit we identified 54 genes encoding for fifteen enzymatic reactions in the tomato genome (Table 1). The results in Table 1 indicate that most of the enzymes in the pathway are encoded by multiple genes. Some enzymes, such as the neutral invertases are encoded by as many as eight paralogues, and most of the remaining enzymes are encoded by 4-6 genes. Only five enzymes are encoded by single genes: PGM_{cyt} PGM_{pl}, PGI_{cyt}, PGI_{pl}, which are encoded by a single gene for each of the different compartmentalized enzyme forms (cytosolic and plastidic), and UGPase.

Analysis of the expression data for these genes derived from tomato ESTs (Table 3) indicates gene functionalization for many of the gene families. This can be seen in some flower specific genes, such as *LIN7*, *FK4* and the plastidic *FBP4*, as well as among gene families with paralogues whose expression has been recorded only, or overwhelmingly, in vegetative tissues. In addition, functionalization with respect to fruit development can also be seen in, for example, the differential expression patterns of the two SUS genes, one (SUS3) in early fruit development and one (SUS4) in later fruit development.

The 54 genes encoding for sugar metabolism enzymes were mapped to the tomato chromosome map, using a mixture of classical mapping techniques utilizing the *S. pennellii* ILs and the recently available database of the tomato genome sequence. Genes that were not observed in the fruit ESTs were also mapped since the trait of *Fgr* may be due to expression of a wild species allele encoding for a sugar metabolism enzyme. The metabolic pathway is randomly distributed among the 12 chromosomes with no indication of gene clustering, with the exception of the tandem LIN genes on chromosomes 9 and 10. However, none of the candidate genes encoding for sugar metabolism enzymes co-localized with the map position of the *Fgr* trait on chromosome 4 (Fig. 2).

Example 3: Map-based cloning of Fgr

In light of the absence of co-localization of the metabolism gene candidates, we identified the *Fgr* gene by a map-based cloning strategy, following a screen of over 12,000 recombination events between two flanking markers at a distance of 0.45 cM (*TG208* and *Asr1*) which revealed 24 recombinants (Fig. 3). The introgression of interest was narrowed to between *TG208* and *Adh1* based on the phenotype of the recombinants. This region, estimated at 0.2 cM was spanned by a set of 7 BACs identified by simultaneous chromosome walking, distally from *TG208* and proximally from *Adh1*, expanded by two scaffolds identified in the recently developed tomato genome sequence. The final BAC 213A03 connected the two contigs. In total, the physical length of the introgression harbouring *Fgr*, between the two markers was ~ 420 kb, or approximately 200kb per 0.1 cM. Analysis of the recombinant lines and their genotyping based on polymorphisms derived from the BAC ends and additional sequences in the region (Table 4) narrowed the region harboring the *Fgr* gene to 39 kb (Fig. 3c).

In order to ascertain direction during the walking procedure FISH was performed using probes made from selected BACs and synaptonemal complex (SC) spreads of tomato chromosomes. Fig. 4 illustrates FISH results for simultaneous hybridization with the BAC harboring TG208 (LE_HBa0013P02) and the BAC containing Adh1 (LE_HBa0208L16). The difference between the average positions of hybridization signals for these two BACs on the long arm of chromosome 4 was 1.7% of the arm length. This translates to a distance of 0.26 ?m (based on average length of 15.2 ?m for this arm in SC spreads). At an estimated DNA density of 1.54 Mbp/?m of euchromatic SC (Stack et al., 2009), the separation between the two markers is calculated to be approximately 400 kb, in good agreement with the sequence length.

Example 4: FGR encodes a MYB domain protein

The 39kb region harbors two open reading frames (ORFs), one for a member of the GRAS transcription factor family and one for a member of the R2R3MYB

transcription factor family (Fig 3c). Comparative expression analysis of the two ORFs in developing tomato fruit of the two genotypes showed that only the *R2R3MYB* gene was differentially expressed between the two genotypes (Fig 5).

In addition, comparison of the protein sequences indicated that the GRAS protein was identical in the two genotypes, whereas the R2R3MYB was only 90% identical (Fig 6). Most of the amino acid differences were located in the C-term protein interaction region while the R2R3 DNA binding region was highly conserved. These results strongly indicate that *Fgr* is encoded by the *R2R3MYB* gene.

Sequence comparisons of the cDNA and genomic DNA of the two MYB alleles indicated that the R2R3 DNA binding region is encoded by two exons interrupted by an intron of dissimilar length in the two species (1,182 and 604 bp in the cultivated and wild species allele, respectively). The R2 protein sequence is identical and the R3 protein sequence differs only in a single conserved amino acid change, from R to K. The intron/exon arrangement places the gene in the class II R2R3 plant MYBs, as classified by Jiang et al (2004). The phylogenetic trees of the MYB protein and its closest relatives are presented in Fig. 7. Based on the R2R3 DNA binding domain the closest Arabidopsis MYB protein (AtMYB54) is ~85% identical, but based on the c-term sequence the Solanaceae protein shows little similarity to any known MYB protein.

Example 5: Differential gene expression of the sugar metabolism pathway

In order to determine the effect of the *Fgr* gene on the sugar metabolism pathway, we carried out a TOM1 microarray comparison of genes expressed in the developing fruit flesh of near-isogenic lines. The analysis of sugar content of the developing pericarp showed that the phenotypic differences in fructose and glucose contents, together with differences in the ratio between them, were clearly evident at the breaker stage (Fig. 8).

cDNA of fruit pericarp of FGR^{EE} and FGR^{HH} lines, at two stages prior to the observed phenotype (IMG and MG) and at breaker, were compared. Of the twenty genes of sugar metabolism represented on the TOM1 chip, only three were differentially expressed (Table 2). These results were corroborated by quantitative

RT-PCR of cDNA of breaker stage fruit, repeated on fruit grown in two, or three, different seasons (Fig. 9; Table 5).

The results indicate that two genes in the pathway are up-regulated in response to FGR^{HH}, while one gene is down-regulated. The up-regulated genes include an uncharacterized hexokinase, HK5, and the gene encoding for cytosolic PGM. The down-regulated gene encodes for a novel paralogue of the PPi-dependent PFP β subunit. The differential expression of these three genes can account for the different phenotype, as will be discussed.

The *HK5* and *PFPβ2* genes are novel members of their respective gene families that have not been studied. *HK5* is represented only as a single EST derived from an IMG fruit library and *PFPβ2* expression has been reported only in vegetative issue libraries (Table 3). Figure 10 shows the relationship of the HK5 and PFPβ2 proteins to the other members of their respective families.

In the description and examples, reference is made to the following tables:

Table 1

Enzyme Name (E.C.)	*Gene*	Gene ID	Chr.	cM	Scaffold
Soluble Invertase (3.2.1.26)	*TIV1*	TC204740	3	74.1	00353
	TIV2 (LIN9)	TC191483	8	55.3	01291
Cell wall Invertase (3.2.1.26)	*LIN5*	TC191468	9	37	00033
	LIN6	TC191459	10	73	01162
	LIN7	TC216266	9	37	00033
	LIN8	TC195389	10	73	01162
Neutral invertase (3.2.1.26)	*NIN1*	TC208528	3	157.5	00008
	NIN2	TC203385	6	44	00077
	NIN3	AW650866	1	165	00425
	NIN4	TC191616	4	118	00192
	NIN5	AW930853	11	14	01386
	NIN6	TC200436	1	31.6	03223
	NIN7	TC205840	1	100.5	00106
	NIN8	TC206783	11	61	00395
Sucrose synthase (2.4.1.13)	*SUS3*	TC217342	7	31	00014
	SUS4	TC217174	12	36	00061
Sucrose Phosphate Synthase (2.4.1.14)	*SPS1*	TC217140	9	90	02760
	SPS2	TC235113	3	157.5	00134
	SPS3	TC233541	11	49	00162
Hexokinase (2.7.1.1)	*HK1*	TC224027	3	134	01172
	HK2	TC218301	2	112	01403
	HK3	TC217200	12	36	00061
	HK4	TC242310	4	118	00192
	HK5	EST408442	2	112	01403
Fructokinase (2.7.1.4)	*FXK1*	TC197906	3	30	01216
	FK2	TC211288	6	69	00823
	FK3	TC207359	2	112	01403
	FK4	TC224873	3	157.5	01537
	FK5 (Putative)	TC195256	9	37	00033
	FK6 (Putative)	TC233563	5	62.8	00243
Glucose 6-	*PGI cyt*	TC191429	12	54.5	01370

Phosphate Isomerase (5.3.1.9)	*PGI pl*	TC191924	4	118	00192
Phosphogluco mutase (5.4.2.2)	*PGM cyt*	TC197381	4	56	00584
	PGM pl	TC191924	3	30	01216
Fructose 1,6 bisphosphatase (3.1.3.11)	*F1,6 BPase1 cyt*	TC220146	9	37	00033
	F1,6 BPase2 cyt	TC219646	12	66	00427
	F1,6 BPase3 cyt	TC219509	4	74	00103
	F1,6 BPase4 pl	TC232750	10	73	01162
	F1,6 BPase5 pl	TC204006	1	140	02184
PPi-dependent Phospho fructokinase (2.7.1.90)	*PFP α1*	TC207279	12	96	01049
	PFP α2	TC200686	4	118	00192
	PFP β1	TC201053	2	75.5	00215
	PFP β2	TC205117	7	41.5	00370
ATP dependent PFK (2.7.1.11)	*PFK1 (pl)*	TC210350	3	134	01172
	PFK2 (cyt)	AW041366	7	38	00296
	PFK3 (cyt)	TC217833	12	96	01049
	PFK4 (cyt)	TC198787	4	74	00103
	PFK5	AW154869	12	-	-
	PFK6	TC223937	6	44	00077
	PFK7 (Putative)	TC200865	3	134	01172
	PFK8 (Putative)	TC230273	6	44	00077
UDPG pyrophosphorylase (2.7.7.9)	*UDPG*	TC201548	3	134	01172
α-Glucosidase (3.2.1.10)	*AGL*	TC191386	4	56	00146

Xylose isomerase	5.3.1.5	*XI*	TC197169	7	1 million
Sorbitol dehydrogenase	1.1.1.14	*SDH*	TC221731	1	31 million
Aldehyde Reducrase	1.1.1.21	*AR*	TC219479	1	88 million

Table 2

		Relative expression		
Gene	**ID**	**IMG**	**MG**	**Breaker**
TIV 1	SGN-U143525	0.91	1.36	1.12
LIN 5	SGN-U147087	0.84	0.96	1.05
HK 5	SGN-U156546	1.51*	1.15	1.7*
FK 2	SGN-U144275	1.06	1.19	1.02
FK 3	SGN-U151013	1.11	1.06	1.1
PGI cyt	SGN-U145142	1.1	1.03	1.06
PGM cyt	SGN-U148507	1.4	1.28	1.5*
F 1,6 BPase 1	SGN-U145762	1.05	0.9	1.01
F 1,6 BPase 2	SGN-U144122	0.95	0.93	1.02
F 1,6 BPase 3	SGN-U144121	0.98	1.05	1.06
F 1,6 BPase 4	SGN-U145518	1.11	1.07	1.18
PFK 1	SGN-U152482	1	0.99	0.88
PFK 2	SGN-U156526	0.94	0.91	0.94
PFP α 1	SGN-U145076	0.94	0.91	0.94
PFP β 1	SGN-U144883	1	0.99	0.88
PFP β 2	SGN-U152529	0.88	0.46*	0.49*
SUS 3	SGN-U213119	0.99	0.94	0.95

Table 3

	Gene	TIGR (TED)	Number of ESTs						
			Vgt	Flw	Ovr	Img	MG	Br	Red
1	TIV1 [2]	TC204740 (TC123892 TC123890 TC123821 TC123825 TC123829)	0	0	0	4	6	470	146
2	TIV 2 (LIN 9)	AM050394	-	-	-	-	-	-	-
3	LIN 5 [2]	TC191468 (TC125260)	0	1	10	0	0	1	0
4	LIN 6 [2]	TC191459 (TC125748)	7	0	0	0	0	0	0
5	LIN 7 [2]	TC216266 (TC115985 TC115986)	0	162	0	0	0	0	0
6	LIN 8 [2]	TC195389 (TC126778)	4	0	0	0	0	0	0
7	NIN 1 [2]	TC208528 (TC128730)	2	0	0	0	0	0	0
8	NIN 2	TC203385 TC201966 (TC122872)	2	2	0	0	0	2	0
9	NIN 3	AW650866 TC	2	0	0	0	1	0	0
10	NIN 4	TC191616	13	1	2	0	7	1	4
11	NIN 5	AW930853	24	21	0	0	1	3	1
12	NIN 6	TC200436	2	0	0	0	0	0	1
13	NIN 7	TC205840 (TC131080)	2	0	0	0	0	0	0
14	NIN 8	TC206783	1	0	0	0	1	0	0
15	Sus3 [2]	TC191561 (TC11630)	50	3	0	1	1	2	0
16	Sus 4 [2]	TC191406 (TC118038)	1	3	1	2	0	0	0
17	SPS 1 [2]	TC191376 (TC130060 TC126090)	1	0	0	0	1	3	3
18	SPS 2 [3]	TC208497	1	0	0	0	0	0	1
19	SPS 3 [2]	TC207103	3	0	0	0	0	0	0
20	HK 1 [2]	TC199071 (TC120665)	2	0	0	0	0	1	0
21	HK 2 [2]	TC193529 (TC119716)	1	1	2	0	0	0	0

22	HK 3 [3]	TC191431	2	0	0	0	0	0	0
23	HK 4 (pl) [3]	TC191465	0	0	0	0	0	1	0
24	HK 5 [3]	EST408442	0	0	0	1	0	0	0
25	FK 1 [2]	TC197906 (TC119482 TC119483)	1	2	0	0	0	1	0
26	FK 2 [2]	TC211288 (TC116377)	24	1	7	6	5	2	2
27	FK 3 (pl) [2]	TC207359 (TC120789)	6	19	0	0	1	15	0
28	FK 4 [2]	TC199858 (TC116279)	0	64	0	0	0	0	0
29	FK 5 (Putative) [2]	TC195256 (TC128151)	1	3	0	0	0	0	0
30	FK 6 (Putative) [3]	TC207113	2	0	0	0	2	0	1
31	PGI (pl) [2]	TC191429 (TC125642 TC128592)	8	3	0	0	2	0	1
32	PGI (cyt) [2]	TC191443 (TC115953)	15	1	0	1	4	6	1
33	PGM 2 (cyt) [2]	TC191924 (TC126701 TC126523 TC127617)	6	5	0	3	1	3	0
34	PGM (pl) [2]	TC197381 (TC123362)	1	0	0	1	1	1	0
35	F1,6BPase1 (cyt) [2]	TC191916 (TC124956 TC124956)	11	1	0	2	0	0	0
36	F1,6BPase2 (cyt) [2]	TC192926 (TC124675)	1	1	2	2	0	3	0
37	F1,6BPase3 (cyt) [2]	TC192145 (TC124676 TC124677)	6	3	1	3	3	2	0
38	F1,6BPase4 (pl) [2]	TC206426 (TC116944)	10	9	1	0	0	0	0
39	F1,6BPase5 (pl) [2]	TC204006 (TC118259)	3	3	0	0	2	0	0
40	PFP α-subunit 1 [2]	TC207279 (TC124642)	12	2	3	0	0	1	0
41	PFP α-subunit 2 [2]	TC200686 (TC124641)	5	1	0	4	0	2	0
42	PFP β-subunit 1 [2]	TC201053 (TC116691)	8	12	2	3	3	2	1
43	PFP β-subunit 2 [2]	TC205117 (TC121321)	3	0	0	0	0	0	0
44	PFK 1 (pl) [2]	TC210350 (TC125963)	2	1	0	0	0	1	0

45	PFK 2 (cyt)	AW041366	-	-	-	-	-	-	-
46	PFK 3 (cyt) [2]	TC193001 (TC118103)	2	0	0	0	0	6	0
47	PFK 4 (cyt) [3]	TC198787	1	0	0	0	0	0	1
48	PFK 5	AW154869	-	-	-	-	-	-	-
49	PFK 6 [2]	TC198995 (TC125962)	4	2	0	0	0	6	0
50	PFK 7 (Putative) [2]	TC200865 (TC119192)	1	3	0	1	0	1	0
51	PFK 8 (Putative) [3]	TC205519	4	0	0	0	0	0	0
52	UDPG [2]	TC201548 (TC124052)	52	24	1	17	9	16	4
53	XI	(TC115963)	12	5	2	0	1	4	1
54	SDH	(TC127176)	4	0	0	0	1	0	0
55	AR	(TC129208)	3	0	0	0	0	0	0

Table 4

Marker	Primers	Method	Assay	Polymorphism	
				Habrochites	*Lycopersicum*
TG208	5`AAGCCTCCAACTTTCACCAGCT 5`TTATGAAAGACCGTCGTTGCGG	Restriction reaction	EcoRI	1000bp	800bp +200bp
ADH1	5`GAGTGTACTGGAAATGTTAACG 5`CAACCCTGTAAATCAAGTATGT	size polymorphism	Agarose gel 2%	130bp	150bp
54I7-T7	5`CTGCAACCATTGTCCATTTG 5`GGCGACAACATAGTGCAAGA	Sequencing	SNPs	A/T/G/A/G	C/G/T/T/C
U213812	5`GTGCACTCAAGGCAGTTTGA 5`TGAAGCGAGACAAGGAGGTT	Sequencing	SNPs	G/A/G/AA	A/C/A/GT
ASR2	5`CACCATCACCATTTGTTCCA 5`CACAAACGGCGTAAGAAATTA	size polymorphism	Agarose gel 2%	247bp	230bp

83C22-T7-f	CATTCCACCAGCCATTTCTT	
83C22-T7-r	CATTCGTGTAGATGGCAACG	390
83C22-sp6-f	TGCTAAATGTGGAGCAGTCG	
83C22-sp6-r	CATGCACTGAGCAATGTCCT	402
92F2-T7-f	CCTTGCTTGGAAAGTGTGGT	
92F2-T7-r	GCTTCGCTTCTTGCTTCTGT	466
Scaffold6658-B-f	ataaacagcccaccagcatc	
Scaffold6658-B-r	attgcaattgccaagaatcc	636
Scaffold5665-B-f	ggcctcatcctgaatgaaaa	
Scaffold5665-B-r	atcaaaacaccctggagtgg	812
Scaffold5665-E-f	aaagttggggcaagtttgtg	
Scaffold5665-E-r	agattggttggttcgctcac	642
72F6-T7-f	GAACCAGCACCTTCAACCAT	
72F6-T7-r	TAGGCTTATTCCACGCAACC	690
72F6-sp6-f	TCGAAACCTGGAAAATGGAG	
72F6-sp6-r	TTCCAGGGAAGTGCAAAATC	818
213E23-t7-f	TTTCTGAGTACGTGCGTTGC	
213E23-t7-r	CACCCAAACTTTTTCTTTCTTCA	735
213E23-SP6-f	GCAGTGCTCATAAGCCAAGA	
213E23-SP6-r	GCTCCCAGTTTGGTACGTTT	500
213A3-T7-f	AGCCATGATCAAAAGCAAGG	
213A3-T7-r	TCAGTGCCTGTGCTAGATGC	1100
69J22-T7-f	GGGGGCATGTTGTTACATCT	
69J22-T7-r	TGCCAATCCTTTACACAGCA	487
69J22-sp6-f	AACATGCACATTGGGTGAAC	
69J22-sp6-r	TCATGGTCACTAATGCCCTCT	440
13P2-sp6-f	ACGCTGACCTTACCCCTAGC	
13P2-sp6-r	AAGGAAGCTTCTCCGACCTC	384
13P2-t7-f	TGCAAGCTTATGCCTTGTGT	
13P2-t7-r	TTGCACATCGCCAGAAGA	320

Table 5

Gene	Forward primer	Reverse primer	Product
Actin	CACCATTGGGTCTGAGCGAT	GGGCGACAACCTTGATCTTC	251 bp
PFP-β2	ACAGGCTCCTGGTGGGCACAA	TCAAAGCCACCTTGATTTCTG	190 bp
PGM cyt	ATATGGGCTGTTTTGGCTTG	TTGCAGCTTCACCAAATGAG	200 bp
HXK 5	AATCCCTACGATGCCAGCTC	ACAGGTTTTTGCTGGCTACA	151 bp
GRAS 6	GCTCCTCAAACTCCAAAGCA	ATTCGAGCTGCAAGAGCTTC	170 bp
MYB-R2R3	TTGTTCCTTCGTCGAAAACC	GGAAGATCAGGGAAGAGTTG	320 bp

5

References

Alba R, Fei Z, Payton P, Liu Y, Moore S, Debbie P, Cohn J, D'Ascenzo M, Gordon J, Rose J, Martin G, Tanksley S, Bouzayen M, Jahn M, Giovannoni J (2004). ESTs, cDNA Microarrays, and Gene Expression Profiling: Tools for Dissecting Plant Physiology and Development. Plant Journal 39:697-714.

Beruter J (1985) Sugar accumulation and changes in the activities of related enzymes during development of the apple fruit. J Plant Physiology 121:331-341.

Biester AM (1925) Carbohydrate studies. I. Relative sweetness of pure sugars. Amer J Physiol 73: 387-400.

Carlisle SM, Blakeley SD, Hemmingsen SM, Trevanion SJ, Hiyoshi T, Kruger NJ, Dennis DT (1990) Pyrophosphate-dependent phosphofructokinase. Conservation of protein sequence between the alpha- and beta-subunits and with the ATP-dependent phosphofructokinase. J Biol Chem. 265:18366-71.

Causse M ,Duffe P ,Gomez MC ,Buret M ,Damidaux R ,Zamir D ,Gur A ,Chevalier C ,Lemaire-Chamley M ,Rothan C . 2004. A genetic map of candidate genes and QTLs involved in tomato fruit size and composition. J Exp Bot. 55:1671-85.

Chang, S-B., L.K. Anderson, J.D. Sherman, S.M. Royer, and S.M. Stack (2007) Predicting and testing physical locations of genetically mapped loci on tomato pachytene chromosome *1*. Genetics **176**:2131-2138.

Chetelat R., T. Klann E., Deverna J.W., Yelle S., Bennett A.B. (1993) Inheritance and genetic mapping of fruit sucrose accumulation in *Lycopersicon chmielewskii*. Plant Journal, 4: 643-650.

Chetelat RT, DeVerna JW, Bennett AB (1995) Introgression into tomato (Lycopersicon esculentum) of the *Lycopersicon chmielewskii* sucrose accumulator gene (*sucr*) controlling fruit sugar composition. Theor Appl Genet 91: 327-333.

Cincarevsky F (1999) High resolution mapping of DNA markers associated with fructose to glucose ratio in the mature tomato fruit. M.Sc Thesis, The Hebrew University of Jerusalem.

Davies JN (1966) Occurrence of sucrose in the fruit of some species of *Lycopersicon*. Nature, 209: 640-641.

Davies JN, Hobson GE (1981) The constituents of tomato fruit - the influence of environment, nutrition and genotype. CRC Crit Rev Food Sci Nutr 15: 205-280

Du et al., Biochemical and molecular characterization of plant MYB transcription factor family. Biochemistry (Moscow) 74:1-11.

Eshed Y and Zamir D (1994) Genomic library of *Solanum pennellii* in *L.esculentum*: a tool for fine maping of genes. Euphytica 79:175-179.

Fei Z, Tang X, Alba R, White J, Ronning C, Martin G, Tanksley S, Giovannoni J (2004). Comprehensive EST Analysis of Tomato and Comparative Genomics of Fruit Ripening. Plant Journal 40:47-59.

Frenette G, Thabet M, Sullivan R. (2006) Polyol pathway in human epididymis and semen. J Androl. 2:233-9.

Fridman E, Pleban T, Zamir D (2000) A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. Proc Natl Acad Sci USA 97: 4718-4723.

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13: 207-209

Granot D (2008) Putting plant hexokinases in their proper place. *Phytochemistry* 69:2649-54.

Hadas R, Schaffer AA, Miron D, Fogelman M, Granot D (1995) PCR-generated molecular markers for the invertase gene and sucrose accumulation in tomato. Theor Appl Genet 90: 1142-1148.

Harel-Beja R, Tzuri G, Portnoy V, Lotan-Pompan M, Lev S, Cohen S, Dai N, Yeselson L, Meir A, Libhaber SE, Avisar E, Melame T, van Koert P, Verbakel H, Hofstede R, Volpin H, Oliver M, Fougedoire A, Stalh C, Fauve J, Copes B, Fei Z, Giovannoni J, Ori N, Lewinsohn E, Sherman A, Burger Y, Tadmor Y, Schaffer AA, Katzir N (2010) A genetic map of melon highly enriched with fruit quality QTLs and EST markers, including sugar and carotenoid metabolism genes. Theor Appl Gen Hubbard N.L., Pharr D.M. and Huber, S.C. (1991), Sucrose phosphate synthase and other sucrose metabolizing enzymes in fruits of various species, Physiol. Plant. **82** pp. 191–196.

Jiang, C., Gu, X. and Peterson, T. 2004b. Identification of conserved gene structures and carboxy-terminal motifs in the MYB gene family of Arabidopsis and Oryza sativa L. ssp. indica. Genome Biol. 5(7): R46.

Karve A, Moore BD. 2009 Function of Arabidopsis hexokinase-like1 as a negative regulator of plant growth. J Exp Bot. 60:4137-49.

Karve A, Rauh BL, Xia X, Kandasamy M, Meagher RB, Sheen J, Moore BD. 2008 Expression and evolutionary features of the hexokinase gene family in Arabidopsis. Planta. 228(3):411-25.

Levin I, Gilboa N, Cincarevsky F, Oguz I, Petreikov M, Yeselson E, Shen S, Bar M and Schaffer A.A (2006) Epistatic interaction between two unlinked loci derived from introgressions from *Lycopersicon hirsutum* further modulates the fructose to glucose ratio in mature tomato fruits. Israel Journal of Plant Science 54:215-222.

Levin I, Gilboa N, Yeselson E, Shen S and Schaffer AA (2000) *Frg*, a major locus that modulates fructose to glucose ratio in mature tomato fruit. Theor Appl Genet, 100: 256-262.

Liu YS, Gur A, Ronen G, Causse M, Damidaux R, Buret M, Hirschberg J, Zamir D. 2003 There is more to tomato fruit colour than candidate carotenoid genes. P lant Biotechnol J. 1:195-207.

Matus JT, Aquea F, Arce-Johnson P. 2008 Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes.BMC Plant Biol. 8:83.

Miron D. and Schaffer A. A. (1991) Sucrose phosphate synthase, sucrose synthase and invertase activities in developing fruit of *Lycopersicon esculentum* Mill. and the sucrose accumulating *Lycopersicon hirsutum* Humb. and Bonpl. Plant Physiology. 95: 623–627.

Miron D., Petreikov M., Carmi N., Shen S., Levin I., Granot D., Zamski E., Schaffer A. A. (2002) Sucrose uptake, invertase localization and gene expression in developing fruit of *Lycopersicon esculentum* and the sucrose. Physiologia Plantarum, 115: 35-47.

Nielsen, T.H., Rung, J.H., and Villadsen, D. (2004). Fructose-2,6- bisphosphate: a traffic signal in plant metabolism. Trends Plant Sci. 9:556-563.

Petreikov M, Dai N, Granot D, Schaffer A.A (2001) Characterization of native and yeast-expressed tomato fruit fructokinase enzymes. Phytochemistry 58: 841-847.

Petreikov M, Shen S, Yeselson Y, Levin I, Bar M, Schaffer A.A (2006) Temporally extended gene expression of the ADP-Glc pyrophosphorylase large subunit (*AgpL1*) leads to increased enzyme activity in developing tomato fruit. Planta 224: 1465-1479.

Petreikov M, Yeselson L, Shen S, Levin I, and Schaffer AA. (2009) Carbohydrate Balance and Accumulation during Development of Near-isogenic Tomato Lines Differing in the AGPase-L1 Allele. J.Amer. Soc. Hort. Sci. 134:134–140.

Ruan Y-L, Patrick JW, Brady C (1997) Protoplast hexose carrier activity is a determinate of genotypic difference in hexose storage in tomato fruit. Plant Cell Environ 20: 341–349.

Schaffer A.A and Petreikov M (1997) Inhibition of fructokinase and sucrose synthase by cytosolic levels of fructose in young tomato fruit undergoing transient starch synthesis. Physiologia Plantarum, 101:800-806.

Schaffer AA, Levin I, Oguz I, Petreikov M, Cincarevsky F, Yeselson E, Shen S, Gilboa N and Bar M (2000) ADPglucose pyrophosphorylase activity and starch accumulation in immature tomato fruit: the effect of a *Lycopersicon hirsutum*-derived introgression encoding the large subunit. Plant Sci., *152: 135-144.*

Schaffer AA, Petreikov M, Miron D, Fogelman M, Spiegelman M, Bnei-Moshe Z, Shen S, Granot D, Hadas R, Dai N, Levin I, Bar M, Friedman M, Pilowsky M, Gilboa N, Chen L (1998) Modification of carbohydrate content in developing tomato fruit. Hortscience 34: 1024-1027.

Schauer N, Semel Y, Roessner U, Gur A, Balbo I, Carrari F, Pleban T, Perez-Melis A, Bruedigam C, Kopka J, Willmitzer L, Zamir D, Fernie AR. 2006 Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement.Nat Biotechnol. 24:447-54.

Stack, S.M., S.M. Royer, L.A. Shearer, S.B. Chang, J.J. Giovannoni, D.H. Westfall, R.A. White, and L.K. Anderson (2009) Role of fluorescence in situ hybridization in sequencing the tomato genome. Cytogenet. Genome Res. **124**:339-350.

Stevens, M.A. (1986) Inheritance of tomato fruit quality components. Plant Breed Rev 4: 273–311.

Stevens R, Buret M, Duffé P, Garchery C, Baldet P, Rothan C, Causse M. (2007) Candidate genes and quantitative trait loci affecting fruit ascorbic acid content in three tomato populations. Plant Physiol. 143:1943-53.

Stitt M (1990) Fructose-2,6-bisphosphate as a regulatory molecule in plants. Ann Rev Plant Physiol Molec Biol 41:153-185.

Stitt M (1998) Pyrophosphate as an energy donor in the cytosolof plant cells: an enigmatic alternative to ATP. Bot. Acta 111:167-175.

Stommell JR, Haynes KG (1993) Genetic control of fruit sugar accumulation in a *Lycopersicon esculentum* X *L. hirsutum* cross. J Amer Soc Hort Sci 118: 859-863

Stracke R, Werber M, Weisshaar B. 2001) The R2R3-MYB gene family in Arabidopsis thaliana. Curr Opin Plant Biol. 4:447-56.

Tanksley SD (1979) Linkage, chromosomal association, and expression of *Adh-1* and *Pgm-2* in tomato. Biochemical Genetics 17: 1159-1167.

van der Merwe MJ, Groenewald JH, Stitt M, Kossmann J, Botha FC. (2010) Downregulation of pyrophosphate: D-fructose-6-phosphate 1-phosphotransferase activity in sugarcane culms enhances sucrose accumulation due to elevated hexose-phosphate levels. Planta. 231:595-608.

Wong HW, Kiss F, Wu M-X, Buchanan BB (1990) Pyrophosphate fructose-6-P 1-phosphotransferase from tomato fruit. Plant Physiol 94:499-506.

Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q, Yunping S, Li Z, Xiaohui D, Jingchu L, Xing-Wang D, Zhangliang C, Hongya G, Li-Jia Q 2006. The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol Biol. 60:107-24.

Claims

1. A polynucleotide comprising a nucleic acid molecule selected from the group consisting of

   a. a nucleic acid molecule comprising a nucleotide sequence encoding a polypeptide comprising the amino acid sequence as depicted in SEQ ID NO: 4;

   b. a nucleic acid molecule comprising a nucleotide sequence as depicted in SEQ ID NO: 3;

   c. a nucleic acid molecule comprising a nucleotide sequence that has at least 60% sequence identity to the sequence shown in SEQ ID NO: 3 and encodes a MYB domain protein;

   d. a nucleic acid molecule comprising a nucleotide sequence that has at least 80% sequence identity to the sequence shown in SEQ ID NO: 3 and encodes a MYB domain protein;

   e. a nucleic acid molecule comprising a nucleotide sequence that has at least 90% sequence identity to the sequence shown in SEQ ID NO: 3 and encodes a MYB domain protein;

   f. a nucleic acid molecule comprising a nucleotide sequence that has at least 95% sequence identity to the sequence shown in SEQ ID NO: 3 and encodes a MYB domain protein;

   g. a nucleic acid molecule comprising a nucleotide sequence the complementary strand of which hybridizes to the nuclei acid molecule of any of a) – f);

   h. a nucleic acid molecule comprising a nucleotide sequence that deviates from the nucleotide sequence defined in any of a) – g) by the degeneracy of the genetic code:

   wherein said nucleic acid molecule as defined in any of a) – h) upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

2. A polynucleotide comprising a nucleic acid molecule selected from the group consisting of

a. a nucleic acid molecule comprising a nucleotide sequence encoding a polypeptide comprising the amino acid sequence as depicted in SEQ ID NO: 8;

b. a nucleic acid molecule comprising a nucleotide sequence as depicted in SEQ ID NO: 6;

c. nucleic acid molecule comprising a nucleotide sequence that has at least 60% sequence identity to the sequence shown in SEQ ID NO: 6 and encodes a MYB domain protein;

d. nucleic acid molecule comprising a nucleotide sequence that has at least 80% sequence identity to the sequence shown in SEQ ID NO: 6 and encodes a MYB domain protein;

e. nucleic acid molecule comprising a nucleotide sequence that has at least 90% sequence identity to the sequence shown in SEQ ID NO: 6 and encodes a MYB domain protein;

f. nucleic acid molecule comprising a nucleotide sequence that has at least 95% sequence identity to the sequence shown in SEQ ID NO: 6 and encodes a MYB domain protein;

g. nucleic acid molecule comprising a nucleotide sequence the complementary strand of which hybridizes to the nuclei acid molecule of any of a) – f);

h. nucleic acid molecule comprising a nucleotide sequence that deviates from the nucleotide sequence defined in any of a) – g) by the degeneracy of the genetic code:

wherein said nucleic acid molecule as defined in any of a) – h) upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

3. A polynucleotide comprising a nucleic acid molecule selected from the group consisting of

a. a nucleic acid molecule comprising a nucleotide sequence encoding a polypeptide comprising the amino acid sequence as depicted in SEQ ID NO: 7;

b. a nucleic acid molecule comprising a nucleotide sequence as depicted in SEQ ID NO: 5;

c. a nucleic acid molecule comprising a nucleotide sequence that has at least 85% sequence identity to the sequence shown in SEQ ID NO: 5 and encodes a MYB domain protein;

d. a nucleic acid molecule comprising a nucleotide sequence that has at least 90% sequence identity to the sequence shown in SEQ ID NO: 5 and encodes a MYB domain protein;

e. a nucleic acid molecule comprising a nucleotide sequence that has at least 95% sequence identity to the sequence shown in SEQ ID NO: 5 and encodes a MYB domain protein;

f. a nucleic acid molecule comprising a nucleotide sequence that has at least 98% sequence identity to the sequence shown in SEQ ID NO: 5 and encodes a MYB domain protein;

g. a nucleic acid molecule comprising a nucleotide sequence the complementary strand of which hybridizes to the nuclei acid molecule of any of a) – f);

h. a nucleic acid molecule comprising a nucleotide sequence that deviates from the nucleotide sequence defined in any of a) – g) by the degeneracy of the genetic code:

wherein said nucleic acid molecule as defined in any of a) – h) upon expression in a tomato plant and/or fruit, leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

4. A MYB domain protein comprising a C-terminal protein interaction region comprising an amino acid sequence selected from the group consisting of

a. the amino acid sequence as depicted in SEQ ID NO: 8

b. an amino acid sequence which has 40% sequence identity to the sequence shown in SEQ ID NO: 8.

c. an amino acid sequence, which has 60% sequence identity to the sequence shown in SEQ ID NO: 8.

d. an amino acid sequence, which has 70% sequence identity to the sequence shown in SEQ ID NO: 8.

e. an amino acid sequence, which has 80% sequence identity to the sequence shown in SEQ ID NO: 8.

f. an amino acid sequence, which has 90% sequence identity to the sequence shown in SEQ ID NO: 8;

g. an amino acid sequence encoded by the polynucleotide of claim 2,

wherein said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

5. A MYB domain protein comprising a N-terminal DNA binding region comprising an amino acid sequence selected from the group consisting of

a. the amino acid sequence as depicted in SEQ ID NO: 7;

b. an amino acid sequence which has 85% sequence identity to the sequence shown in SEQ ID NO: 7;

c. an amino acid sequence, which has 90% sequence identity to the sequence shown in SEQ ID NO: 7;

d. an amino acid sequence, which has 95% sequence identity to the sequence shown in SEQ ID NO: 7;

e. an amino acid sequence, which has 98% sequence identity to the sequence shown in SEQ ID NO: 7;

f. an amino acid sequence encoded by the polynucleotide of claim 3,

wherein said protein leads to a modified hexose partitioning in tomato fruits towards increased fructose levels.

6. A MYB domain protein comprising a N-terminal DNA binding region, which has 90% sequence identity to the sequence shown in SEQ ID NO: 7 and a C-terminal protein interaction region, which has 70% sequence identity to the sequence shown in SEQ ID NO: 8.

7. A polynucleotide comprising a nucleic acid sequence encoding a MYB domain protein of any of the preceding claims.

8. An expression cassette comprising the polynucleotide of any of the preceding claims.

9. A vector molecule comprising the expression cassette of claim 8.

10. Use of a polynucleotide, an expression cassette or a vector molecule of any of the preceding claims for modifying hexose partitioning in tomato fruits towards increased fructose levels.

11. A method for improving fruit quality of tomato fruits by modifying hexose partitioning in the fruit towards increased fructose levels comprising transferring through plant transformation techniques to and expressing in a tomato plant and/or fruit a polynucleotide, an expression cassette or a vector molecule of any of the preceding claims.

12. A method for improving fruit quality of tomato fruits by modifying hexose partitioning in the fruit towards increased fructose levels comprising:

a. obtaining a first tomato plant verified to contain the polynucleotide of any of the preceding claims;
b. crossing said first tomato plant with a second tomato plant verified to lack the polynucleotide of any of the preceding claims; and
c. identifying a tomato plant resulting from the cross exhibiting modifed hexose partitioning in the fruit towards increased fructose levels.

13. A transgenic tomato plant including parts thereof or a transgenic tomato plant cell which plant, plant part or plant cell is genetically engineered with and contains the polynucleotide, the expression cassette or the vector molecule of any of the preceding claims and exhibits modified hexose partitioning in the fruit towards increased fructose levels.

14. Use of a plant according to claim 13 for growing harvestable parts exhibiting modified hexose partitioning in the fruit towards increased fructose levels.

15. A method of producing tomato fruit with modified hexose partitioning in the fruit towards increased fructose levels comprising:

a. growing a tomato plant according to claim 13;

b. allowing said tomato plant to set fruit; and

c. harvesting fruit of said tomato plant.

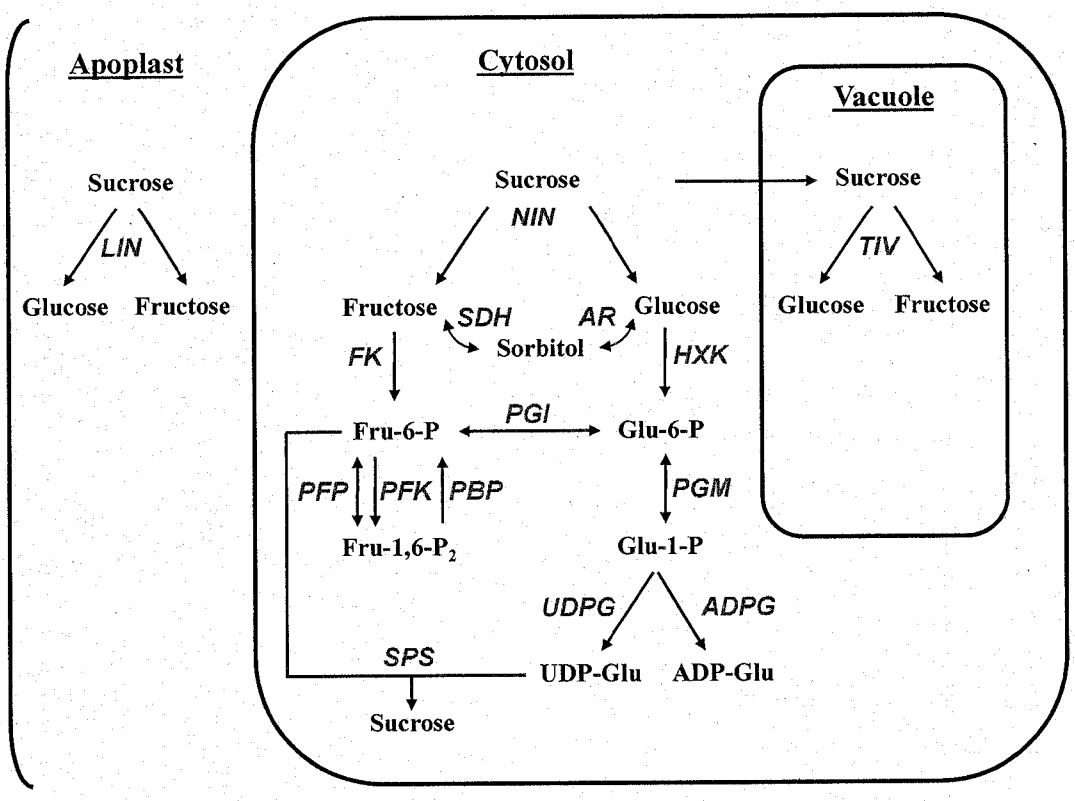

Apoplast
Sucrose
LIN
Glucose
Fructose
Cytosol
Sucrose
NIN
Fructose
SDH
AR
Glucose
Sorbitol
FK
HXK
Fru-6-P
PGI
Glu-6-P
PFP
PFK
PBP
PGM
Fru-1,6-P2
Glu-1-P
UDPG
ADPG
SPS
UDP-Glu
ADP-Glu
Sucrose
Vacuole
Sucrose
TIV
Glucose
Fructose

Figure 1

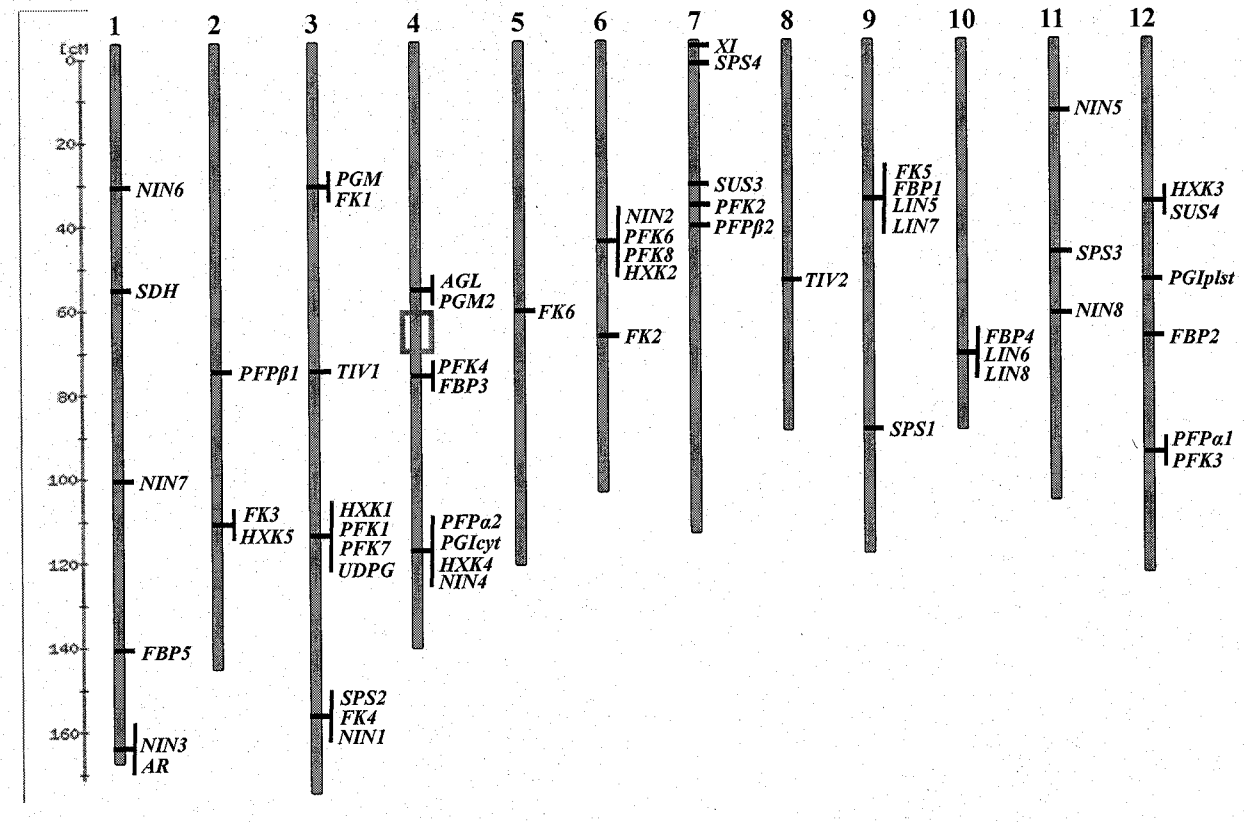

cM
0
20
40
60
80
100
120
140
160
1
NIN6
SDH
NIN7
FBP5
NIN3
AR
2
PFPβ1
FK3
HXK5
3
PGM
FK1
TIV1
HXK1
PFK1
PFK7
UDPG
SPS2
FK4
NIN1
4
AGL
PGM2
PFK4
FBP3
PFPα2
PGIcyt
HXK4
NIN4
5
FK6
6
NIN2
PFK6
PFK8
HXK2
FK2
7
XI
SPS4
SUS3
PFK2
PFPβ2
8
TIV2
9
FK5
FBP1
LIN5
LIN7
SPS1
10
FBP4
LIN6
LIN8
11
NIN5
SPS3
NIN8
12
HXK3
SUS4
PGIplst
FBP2
PFPα1
PFK3

Figure 2

5

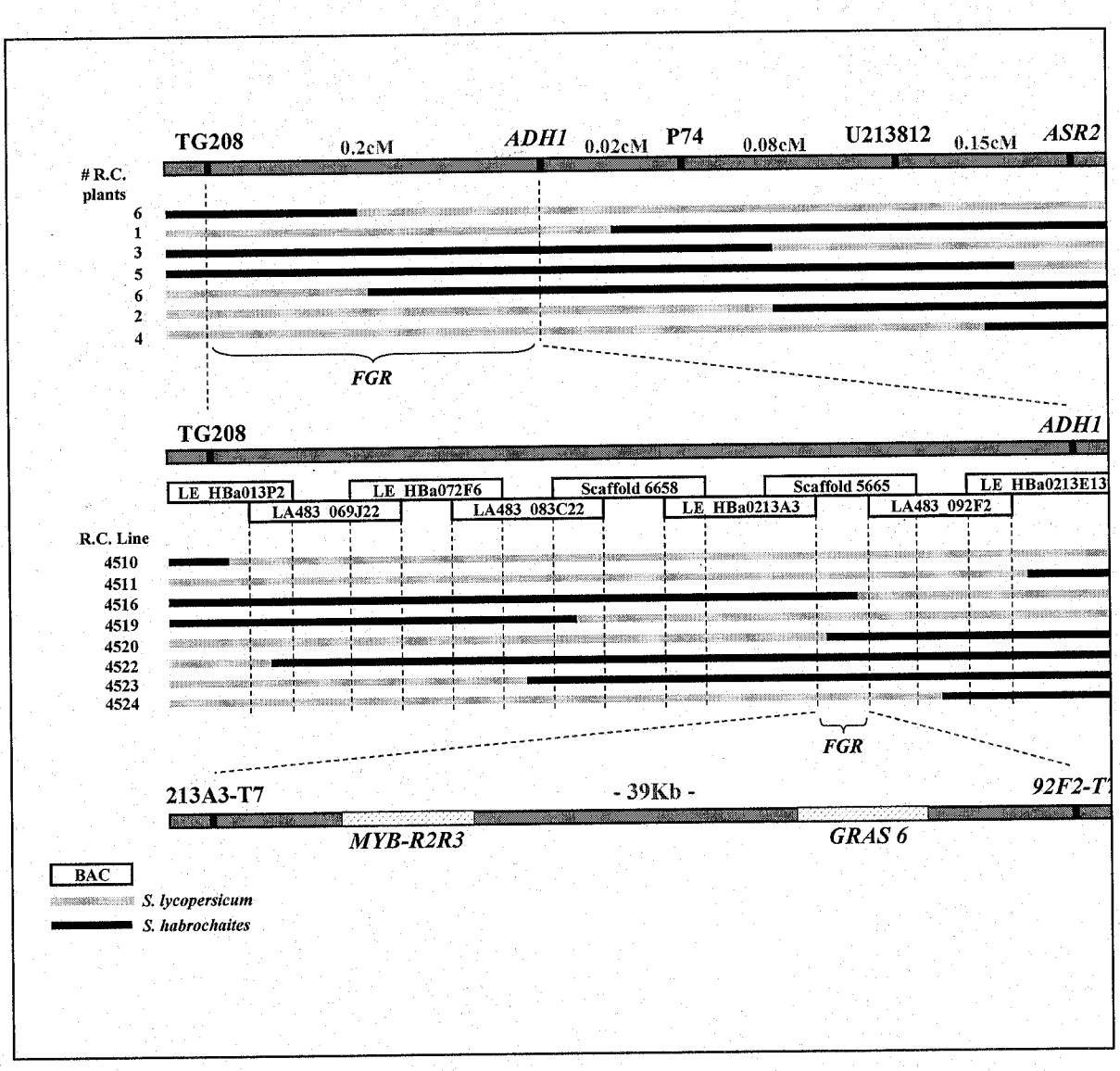

TG208
0.2cM
ADH1
0.02cM
P74
0.08cM
U213812
0.15cM
ASR2
R.C. plants
6
1
3
5
6
2
4
FGR
TG208
ADH1
LE_HBa013P2
LE_HBa072F6
Scaffold 6658
Scaffold 5665
LE_HBa0213E13
LA483_069J22
LA483_083C22
LE_HBa0213A3
LA483_092F2
R.C. Line
4510
4511
4516
4519
4520
4522
4523
4524
FGR
213A3-T7
- 39Kb -
92F2-T
MYB-R2R3
GRAS 6
BAC
S. lycopersicum
S. habrochaites

Figure 3

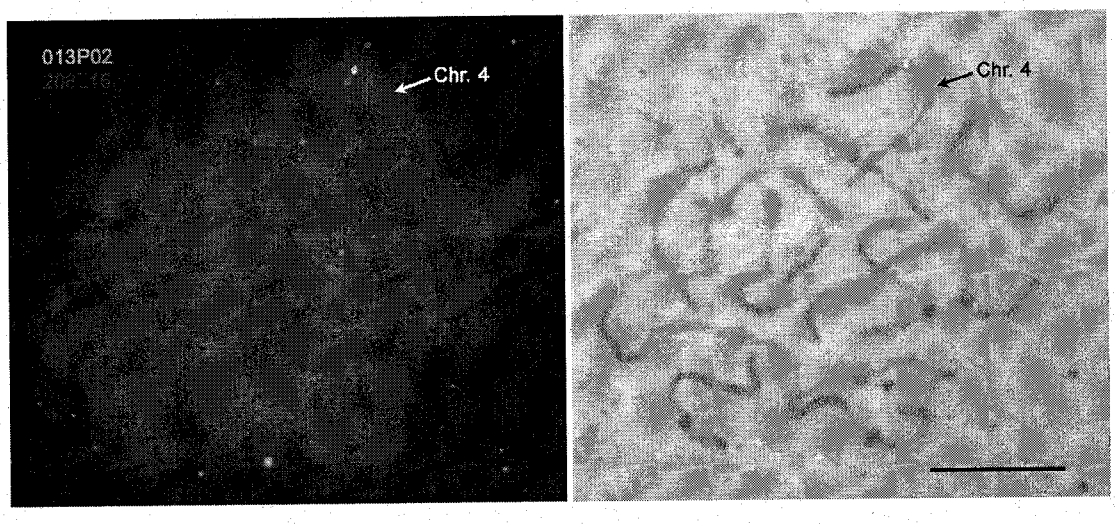


Figure 4

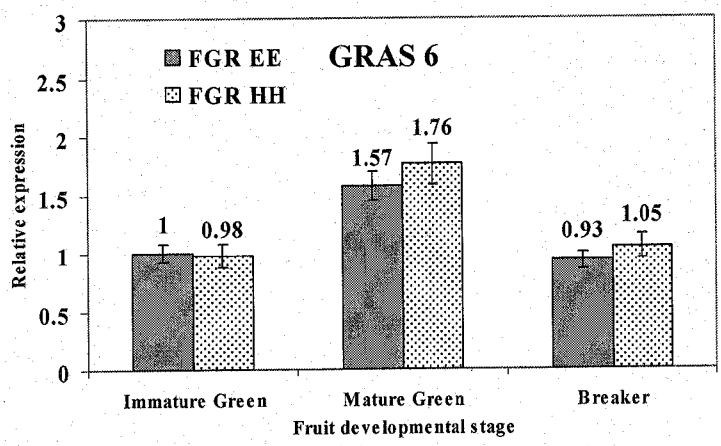


Figure 5

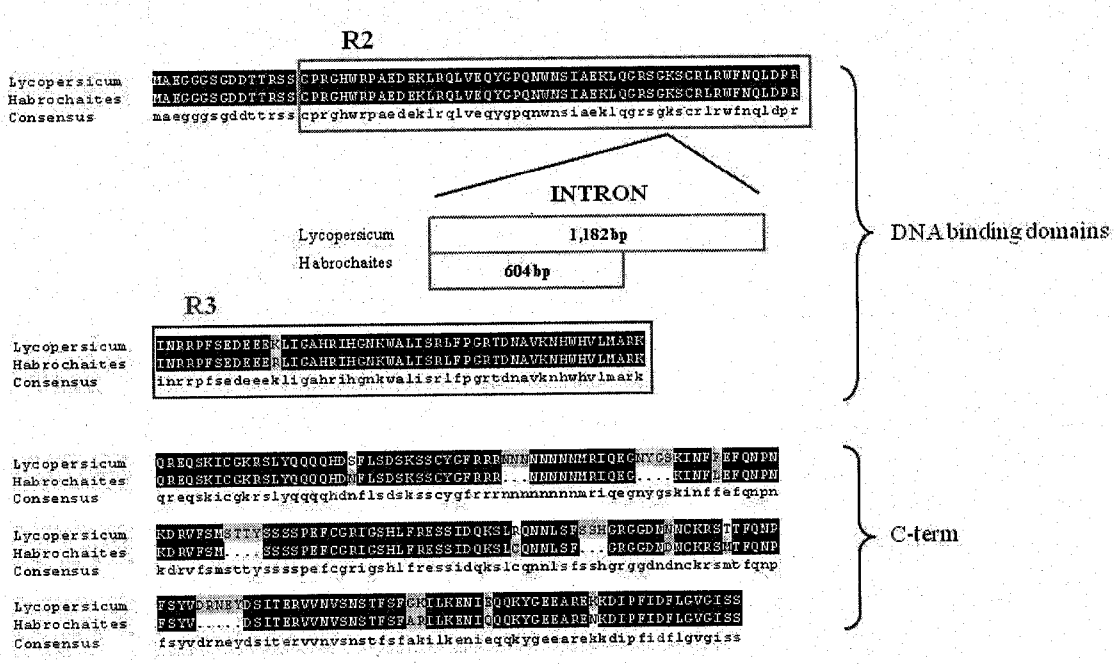

R2
Lycopersicum
Habrochaites
Consensus
INTRON
Lycopersicum
Habrochaites
1,182bp
604bp
DNA binding domains
R3
C-term

Figure 6

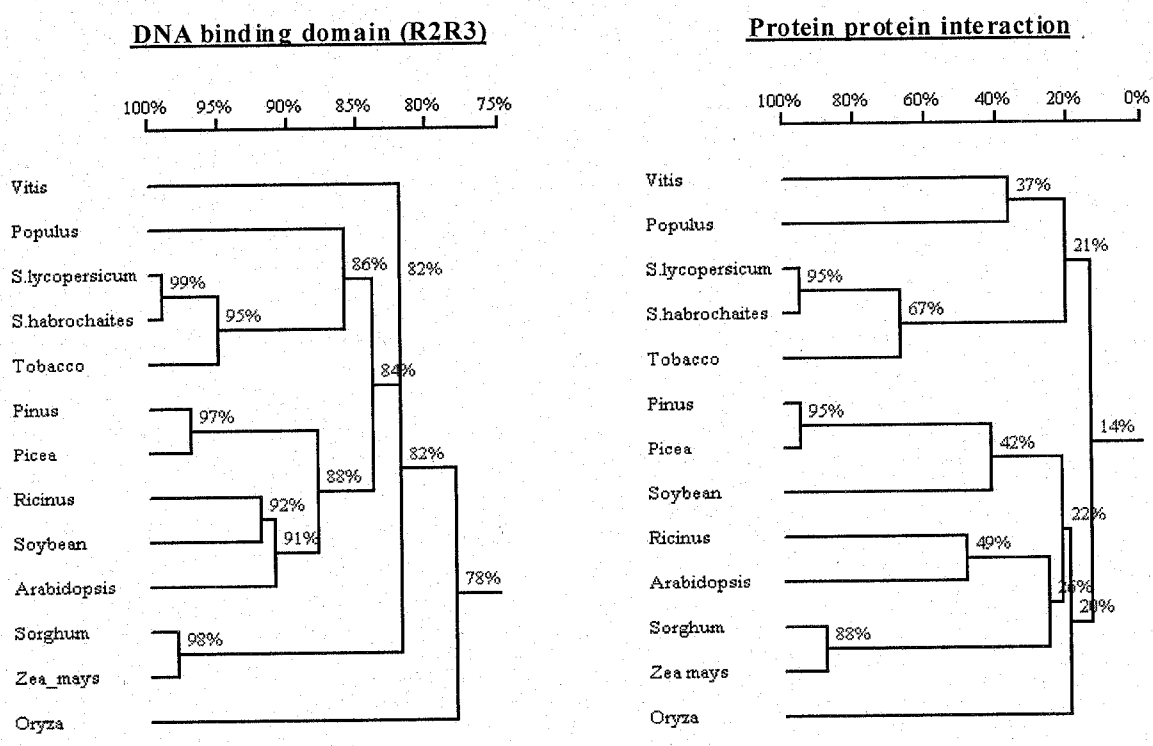

DNA binding domain (R2R3)
100% 95% 90% 85% 80% 75%
Vitis
Populus
S.lycopersicum
S.habrochaites
Tobacco
Pinus
Picea
Ricinus
Soybean
Arabidopsis
Sorghum
Zea_mays
Oryza
99%
95%
86%
82%
84%
97%
82%
88%
92%
91%
78%
98%
Protein protein interaction
100% 80% 60% 40% 20% 0%
Vitis
Populus
S.lycopersicum
S.habrochaites
Tobacco
Pinus
Picea
Soybean
Ricinus
Arabidopsis
Sorghum
Zea mays
Oryza
37%
21%
95%
67%
95%
42%
14%
22%
49%
88%

Figure 7.

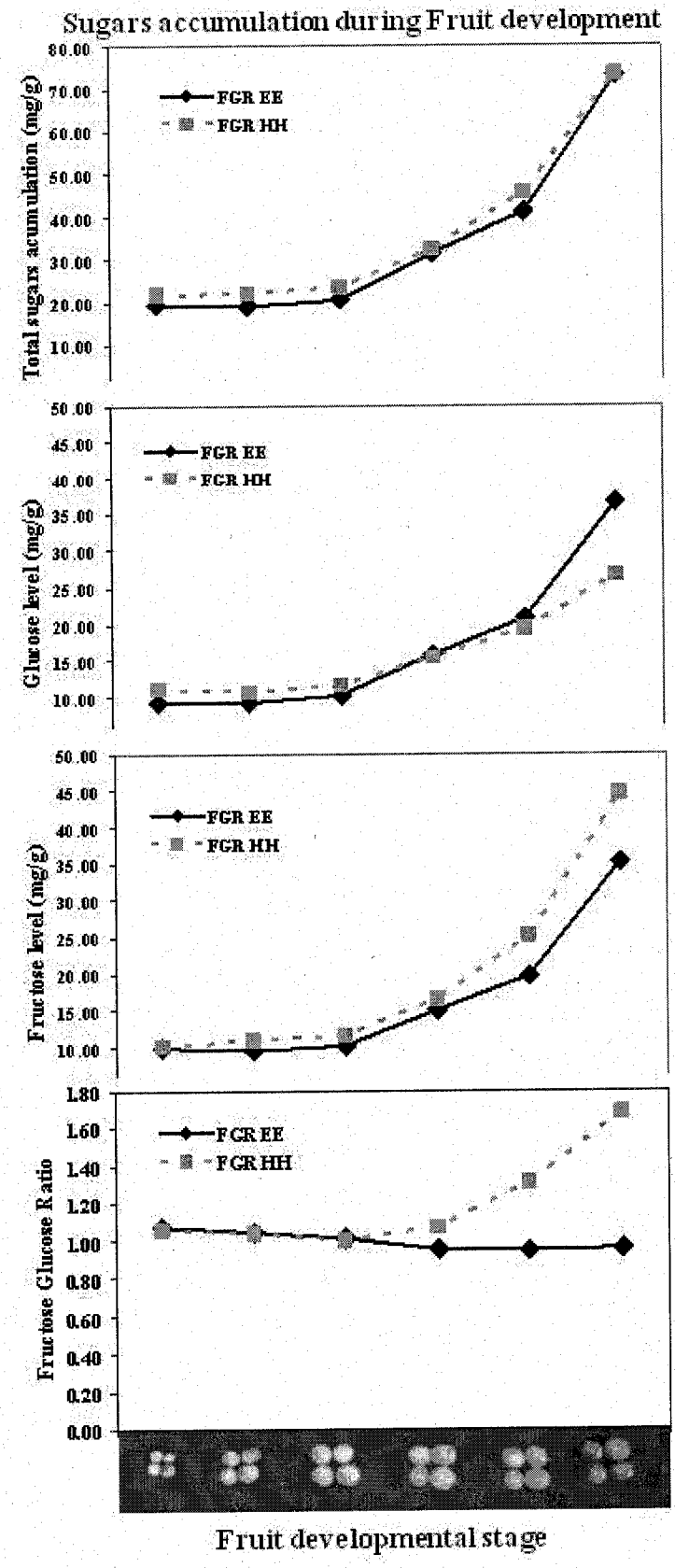
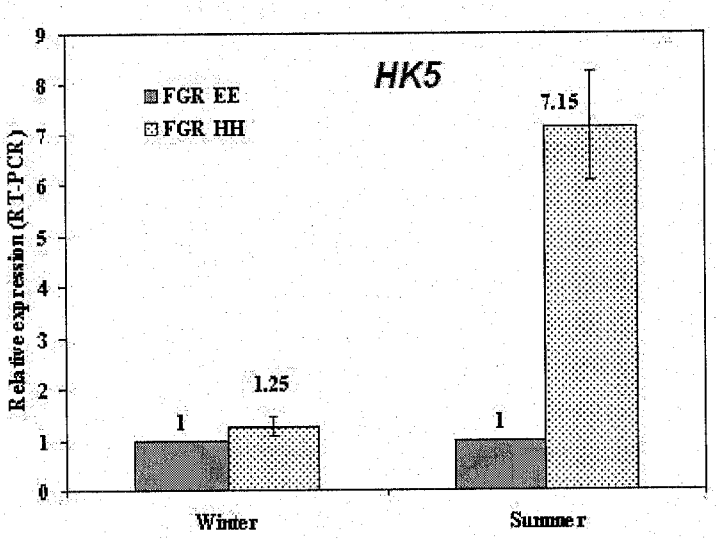
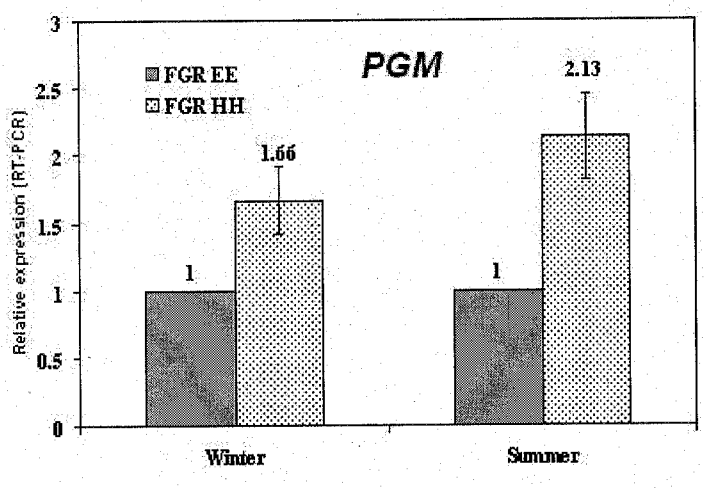

Sugars accumulation during Fruit development
Total sugars acumulation (mg/g)
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
FGR EE
FGR HH
Glucose level (mg/g)
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
FGR EE
FGR HH
Fructose level (mg/g)
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
FGR EE
FGR HH
Fructose Glucose Ratio
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
FGR EE
FGR HH
Fruit developmental stage

Figure 8


HK5
FGR EE
FGR HH
Relative expression (RT-PCR)
9
8
7
6
5
4
3
2
1
0
1
1.25
1
7.15
Winter
Summer


PGM
FGR EE
FGR HH
Relative expression (RT-PCR)
3
2.5
2
1.5
1
0.5
0
1
1.66
1
2.13
Winter
Summer

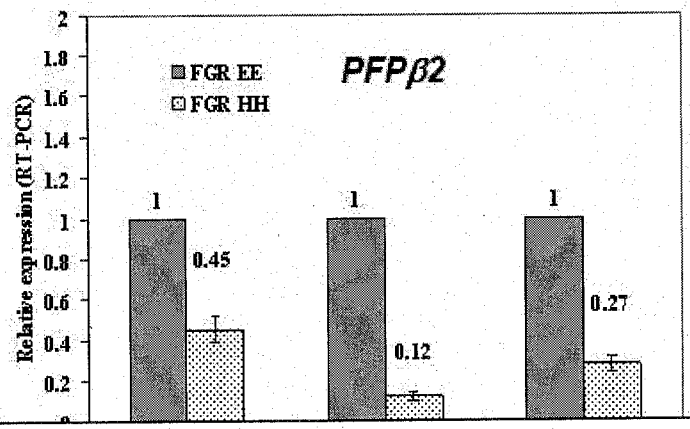
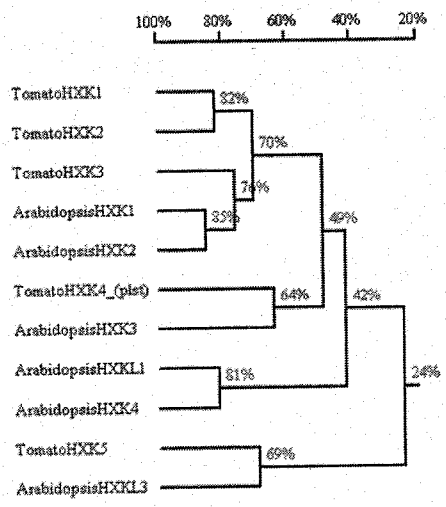

PFPβ2
FGR EE
FGR HH
Relative expression (RT-PCR)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
1
0.45
1
0.12
1
0.27

Figure 9

Hexokinase and hexokinase like proteins from tomato and Arabidopsis



PFP α and β subunits from tomato and Arabidopsis

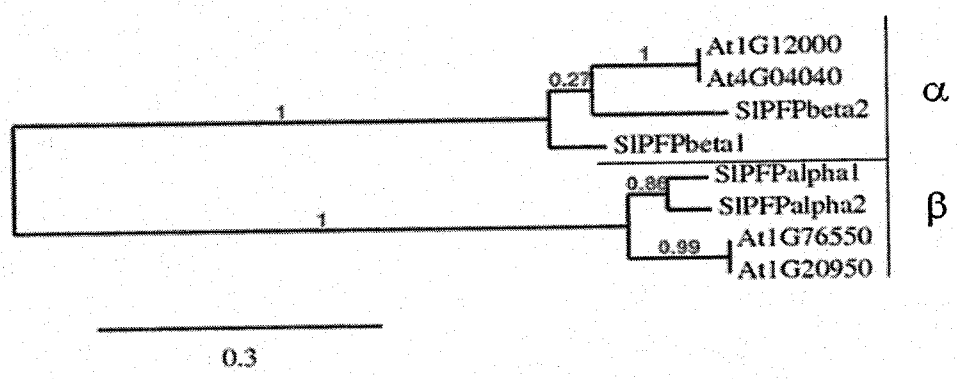


Figure 10

INTERNATIONAL SEARCH REPORT

International application No
PCT/IL2012/000016

A. CLASSIFICATION OF SUBJECT MATTER
INV. A01H5/00 C07K14/415 C12N15/82
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A01H C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, WPI Data, Sequence Search, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LEVIN I ET AL: "Fgr, a major locus that modulates the fructose to glucose ratio in mature tomato fruits", THEORETICAL AND APPLIED GENETICS, vol. 100, no. 2, January 2000 (2000-01), pages 256-262, XP002673514, ISSN: 0040-5752 cited in the application abstract page 257, left-hand column, line 19 - line 33 page 262, left-hand column, line 20 - line 34 ----- -/--	1-15

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 April 2012	09/05/2012

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Mundel, Christophe

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99/04621 A1 (ISRAEL STATE [IL]; LEVIN ILAN [IL]; SHAFFER ARTHUR A [IL]) 4 February 1999 (1999-02-04) abstract page 4, line 20 - line 23 page 6, line 4 - line 5 page 9, line 8 - line 12 page 10, line 25 - line 27	1-15
X	DATABASE P-PSDB [Online] 24 March 2005 (2005-03-24), Bloksberg et al.: "Eucalyptus grandis transcription factor cDNA MYB family", XP002673515, retrieved from ebi Database accession no. ADW16614 the whole document & WO 2005/001050 A2 (ARBORGEN LLC [US]; BLOKSBERG LEONARD N [US]; BRYANT CATHERINE [US]; CO) 6 January 2005 (2005-01-06)	1-3,7-9
X	DATABASE UNILIB [Online] 4 November 2009 (2009-11-04), Salin et al.: "Quercus robur xylem EST clone LG0AAB23YJ16RM1 similar to MYB transcription factor R2R3 type", XP002673516, retrieved from ebi Database accession no. FP036688 the whole document	1-3,7
X	DATABASE EMBL [Online] 24 March 2009 (2009-03-24), Tuskan et al.: "MYB158 Populus trichocarpa", XP002673517, retrieved from EBI Database accession no. B9H7M6 the whole document	5

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/IL2012/000016

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9904621 A1	04-02-1999	AU 8354898 A	16-02-1999
		EP 0998191 A1	10-05-2000
		JP 2001510685 A	07-08-2001
		WO 9904621 A1	04-02-1999