Eco-TILLING on tomato core collections for brix: Implementing the methodology and first results

Master Thesis Plant Breeding

– Specialization Plant Breeding and Genetic Resources

Code: PBR-80436

Author	Guillaume Bauchet
Registration No:	820707038060
Supervisors:	Dr. Christian Bachem/ Antoine Gady
Examiners:	Dr. Christian Bachem/ Dr. Sjaak van Heusden
Chair Group	Plant Breeding
Starting - End time	14-05-2009 - 14-12-2009

Table of Content

Acknowledgements

First of all, I would like to give a particular thanks to Antoine Gady for his guidance all along this thesis. Antoine has always supported my work and has been very comprehensive and patient during difficult moments.
As well, special thanks to Christian Bachem for giving this opportunity to work on this Eco-TILLING topic. This master thesis gave me a chance to discover and learn about lab research. It also gave the great chance to work and interact with people from different backgrounds and countries.
I am grateful to Richards Finkers, who introduced me to the world of bioinformatics and who always answered to my requests on the BreeDB database and Anne Korststee who shared with me her knowledge on suc to starch regulation pathways.
Thanks to Arno Bovy and Sjaak van Heusden for their comments and for sharing their plant material and data.
My thanks to Rene Klein Lankorst and to the EU-SOL project for giving the opportunity to intend the phenotyping workshop in Akko.
Many thanks to Roni Tadmor, Jacob Tadmor and Saleh for their help and their warm welcoming in Israël.
Finally, a special thanks to Ping Ping Huang who was an amazing labmate. Thanks for your help, support and great food!

I want to dedicate this work to my dad who left us few months ago

Abstract

Single nucleotide polymorphisms (SNPs) are the most common form of DNA sequence variation. Once discovered, SNPs can be used in a broad way: for germplasm/sequence characterization, define haplotype of specific gene; saturate a genetic map.

In the present study, Eco TILLING methodology using High Resolution Melting (HRM) analysis was tested and implemented on the EU-SOL and CBSG tomato core collections. The two core collections were screened for polymorphism within conserved domains for a gene set (20 genes) related to the brix (total sugar content). Thus, collections were phenotyped for brix content and other agronomic traits. Two subsets were made, a CBSG subset (94 accessions, one PCR plate of genomic DNA), EU-SOL subset (258 accessions, GenomiPhi DNA, dispatched in three PCR plates).

In a first step, the HRM screening method efficiency and accuracy were tested. Moneymaker was used as a reference line to form heteroduplexes. GenomiPhi DNA was revealed as a source of unexpected variation.

In a second step screening was performed. On a subset of 352 accessions, about 1126 Kb were screened. 188 polymorphisms were discovered including 110 SNPs and 79 Indels. Method showed 82% efficiency in comparison to whole plate sequencing through Sanger method. One SNP induces an amino acid change next to an active site. The SNP is located in Sucrose Invertase conserved domain inducing a change from Lysine (K) to Glutamic acid (E). Change is predicted as non deleterious by SIFT software.

Total observed natural polymorphism was lower than expected; targeting a selection of conserved domains within a gene set of highly conserved proteins may explain this result.

With those primary results, we demonstrate the applicability of Eco-TILLING using HRM screening technique for SNP discovery in tomato. We compared obtained results with a similar Eco-TILLING project using endonuclease digestion and LICOR gel as screening technique. Detected polymorphism show similar results in both experiments in coding and non coding regions. As well, differences of polymorphism are observed between EU-SOL and CBSG subsets. This may be an evidence of changes in genetic diversity induced by breeding.

Method using GenomiPhi DNA can be improved with gained experience. Further investigations can be implemented concerning coding SNP. Confirm with re-sequencing, perform enzyme activity test to measure potential changes and evaluate induced changes on protein structure with 3D modeling.

Abbreviations

AGPase: ADP glucose pyrophosphorylase
CC: core collection
CD: conserved domain
Frk: fructokinase
Hxk: hexokinase
Icl: isocitrate lyase
Lin: invertase
LS: Lightscanner™
Mls: malate synthase
GPI: Glucose-6-phosphate isomerase
HRM: high resolution melting
PCR: polymerase chain reaction
SIFT: sorting intolerant from tolerant
SNP: single nucleotide polymorphism
Susy: sucrose synthase
TILLING: target induced lesion in genomes
UGPase: UDP-glucose-pyrophosphorylase
WGA: Whole genome amplification

1. Introduction

1.1 Tomato

Tomato (*Solanum lycopersicum*) is a major cultivated vegetable in the world originated from western South America (Jenkins 1948). Diploid and autogamous plant, tomato and its wild relatives (*Solanum* section *lycopersicon*) belong to the large family of Solanaceae which include other important cultivated crops such as eggplant pepper and potato. Tomato has an important nutritional value and sugar content which makes it a desirable crop for the human diet. Tomato is a model organism for Solanaceous and more broadly for fleshy-fruited plants (Kimura and Sinha 2008).
As most of the cultivated crops, tomato has faced an erosion of its genetic diversity through centuries of breeding and selection. Tomato breeding has mainly focused on productivity and resistance to biotic stresses as tomato breeders have mainly selected for yield, uniformity and disease resistance (Zamir 2001).

To preserve genetic and phenotypic diversities and make resources available for breeding purposes, core collections (CC) have been developed. Improvement of agronomical traits using this natural diversity becomes critical (Gur and Zamir 2004). This improvement is supported by a combination of scientific and technologic tools from the genomic era (Bai and Lindhout 2007) such as web tools and databases. In tomato a particular focus is done on nutritional and organoleptic traits. Taste and flavor are major concerns for the consumer (Heuvelink 2005). Taste is a good example of complex trait as it relies on many components that are needed in appropriate amounts. Major components of the taste are sugars and acids and their ratio within the fruit. Most of the modern tomato varieties are derived from the domestication of the Peruvian wild cherry types, brought to Mexico by pre Hispanic civilizations and spread over Europe in the XVIth century (Luckwill 1943). Increasing tomato flavor through an increase of sugar content and titrable acids is under investigation for several decades already (Stevens, Kader et al. 1979). During precolombian era tomato fruit was mainly consumed as a dessert. As a consequence, selection was focused on sweetness (up to 60% of sugars in the dry matter weight). Unfortunately, through domestication the trait has often been negatively selected. This is notably due to weight-sugar antagonism (Lecomte, Duffe et al. 2004) and negative linkage between fruit size shape and sugar content. Explore core collections for their genetic and phenotypic resources for targeted traits a major tool for plant improvement (Zamir 2008). A straightforward method to evaluate sugar content of a tomato fruit is to measure its brix value.

1.2 Brix in tomato: importance and metabolism

Fruit metabolism has a vast complexity many components, including primary and secondary metabolites. Brix or soluble solids content is the measure of the dissolved sugar-to-water mass ratio of the fruit juice, a major trait to evaluate the taste and measure its sweetness component.

Phenotypically brix can be easily measured by the use of refractometer. A more accurate characterization of flavor components can be performed using sugar dosage and metabolite profiling (Bovy, Schijlen et al. 2007). Genetically, its heritability and expression measurements are more complex to evaluate, as sugars synthesis relies on signaling cascades and enzymatic reactions, mainly produced by the central carbon metabolism of the fruit. Those sugars vary qualitatively and quantitatively through their type and their ratio within the fruit. In fruits from domesticated tomato, sucrose, glucose and fructose are the major sugars present. In wild accessions accumulated sugars are mostly sucrose (Kortstee, Appeldoorn et al. 2007). The observed variation in relative levels of hexose and sucrose is influenced by relative activities of enzymes responsible for the degradation of those sugars. At green stage, the tomato fruit undergo a period of transient starch accumulation characterized by developmental changes in the activities of key enzymes in the sucrose to starch pathway (Zamski and Schaffer 1996). It is hypothesized that this transient starch may contribute to the soluble hexose level in the mature fruit (Schaffer and Petreikov 1997).

As there is a known difference between the temporal metabolic control of starch synthesis and its spatial control, the investigation of existing genetic diversity for the related coding sequence of those key enzymes may contribute to explain observed differences in brix value and their impact on the plant (for example on the 3D structure of the protein). Sucrose invertases and sucrose synthases play a key role in the loading/unloading capacity of sucrose in the initial stages of fruit development. Fridman *et al*. (2000) uncovered the molecular basis of this variation of invertase *lin5* by characterizing a 484 bp sequence derived from a wild accession, *Solanum pennelli* (Fridman, Pleban et al. 2000). It was confirmed that the wild-species allele increased glucose and fructose contents in cultivated tomato fruits in various genetic backgrounds and environments (Zanor, Osorio et al. 2009). An exhaustive investigation of genes involved in sucrose metabolic pathways (see figure 1) and their sequence may reveal other allele variations of interest. Majors genes described in literature are listed below.

-*ADP glucose pyrophosphorylase* (large and small subunit): catalyzes the synthesis of ADP-Glc in starch synthesizing tissues. The AGPase family present four genes.
(Zamski and Schaffer 1996).

-*Fructokinase*: catalyzes the phosphorylation of fructose by ATP at C1 to form fructose-1-phosphate. Four fructokinases are identified. (Pego and Smeekens 2000).

-*Glucose-6-phosphate isomerase*: occurs in the glycosis, Glucose-6-phosphate isomerase, is also called phosphoglucose isomerase (PGI). By isomerization of an aldose to a ketose it converts G6P to fructose-6-phosphate (Schaffer and Petreikov 1997).

-*Hexokinases*: Catalyses the reaction of transfer of a phosphoryl group from ATP to glucose to form glucose-6-phosphate. Four isoforms are listed (Claeyssen and Rivoal 2007).

-*Invertases*: Hydrolysis of sucrose to D-glucose and D-fructose is accompanied by a change in optical rotation from dextrogire to levogire. Invertase mediates the hydrolytic cleavage of sucrose into the hexose monomers. Consequently, hydrolyzed sucrose is sometimes called invert sugar and the enzyme that catalyzes this process, α-D-glucosidase, also named invertases. Three types of invertases exist: apoplast, cytoplasm and vacuolar. If invertases are involved in carbohydrate partitioning they are also involved in growth and development signaling. Tomato invertase include: *lin*5, *lin*6, *lin*7, *lin*8 and *tiv*1 (also named *SI*) (Roitsch and Gonzalez 2004).

-*Sucrose synthase*: Carries out the kinetically reversible transglycosilation reaction between UDPsucrose and UDP fructose (Schaffer and Petreikov 1997)
-*UDP-glucose-pyrophosphorylase:* Part of the glycogen metabolism, UDP-glucose-pyrophosphorylase catalyzes the reaction of UTP and G1P (Schaffer and Petreikov 1997).
-Isocitrate lyase: Cleaves the isocitrate to succinate and glyoxylate.
-Malate synthase: Condenses glyoxylate with a second molecule of acetylCoA to form malate.
Those two last enzymes are part of the glyoxylate cycle that produces glucose from fatty acids (Voet and Voet 2004).

As genomic or coding DNA sequences are available for those genes, a reverse genetic technique such as a candidate gene approach can be implemented to discover potential allelic variants (Pflieger, Lefebvre et al. 2000). Active sites within sequences are predominant in gene function, thus, their sequences are well conserved. A base pair change within a conserved domain has a great chance to influence the transcription product and its translated protein. A change in the protein structure can induce major changes in metabolism.

1.3 Eco-TILLING

Among reverse genetic techniques, TILLING (Targeted Induced Local Lesions IN Genomes) for which several mutant tomato populations and screening platform have been developed (Gady, Hermans et al. 2009). This successful reverse genetic technique uses chemical mutagenesis and SNP discovery method for point mutation identification in mutant population (Till, Zerr et al. 2006). Nevertheless, in studies of variation, elucidation of characteristics cannot be identified using only induced genetic mutations (Gazzani, Gendall et al. 2003). Enzymatic mismatch cleavage method used for TILLING is applicable to any heteroduplex DNA (Till, Zerr et al. 2006). A method with similar approach as TILLING can be applied to collections of natural variants: Eco-TILLING. Eco-TILLING aimed to be a fast and accurate tool to explore genetic diversity gathered in core collections (Comai, Young et al. 2004). Through massive screening of core collections for targeted genes Eco-TILLING can enrich SNPs database and provides tool to compare lines (i.e. cultivars against wild relatives)

Several SNP screening methods exist: the classical Eco-TILLING method is based on endonuclease enzyme (Cel1, Endo1), which cleaves at mutation point by recognizing mismatches in double strand DNA molecules. It can be adequately replaced by a SNP screening method using High Resolution Melting curve analysis (HRM) with Light scanner instrument (Idaho Technology Inc.) as implemented by Gady et al. for TILLING. Improvements in cost and time can be expected (Parry, Madgwick et al. 2009).

Differently to TILLING, the pooling strategy is not used in order to unambiguously assign haplotypes. DNA is extracted from core collections and set of DNA plates is constructed. To increase mismatch detection sensitivity, in each well the tested accession is bulked with a reference line (i.e. a known cultivar), to form DNA heteroduplexes. Heteroduplexes formation is performed following Polymerase Chain Reaction (PCR): products are denatured and reannealed to form heteroduplexes between the sequences of tested line and reference line.

The HRM screening method uses the LC Green dye (Idaho Technology Inc.), which binds into

the double stranded DNA and emits fluroescence. During the melting step, LC-green release is recorded allowing the visualization of sample specific melting pattern. Comparison of these profiles allows the detection of heteroduplexes that formed during PCR (Liew, Pryor et al. 2004). Two main patterns are expected:
-The reference and tested accession show a same melting curve pattern (homoduplex), no putative SNP or INDEL are predicted.
-The reference and tested accessions show a different melting curve pattern (heteroduplex), meaning presence of putative SNP or INDEL.

PCR amplification is performed for the targeted fragment and the product for each accession is analyzed by High Resolution Melting with Lightscanner™ instrument. Melting curves are displayed as illustrated in Figure 2 (see figure 2).

The selected positive line that contains putative allelic variant is sequenced by Sanger method to identify the polymorphism. Polymorphism can lead to amino acid change and it may affect protein function. Programs such as SIFT (Sorting Intolerant From Tolerant) (Ng and Henikoff 2003) or CUPsat (Parthiban, Gromiha et al. 2006) can be implemented to investigate whether or not an amino acid substitution is predicted as affecting protein function.

1.4 Goal

In this thesis we propose to combine Eco-TILLING to High Resolution Melting on two distinct tomato core collections for a set of genes involved in Brix. Eco-TILLING aims to be a straightforward and cost efficient method to reveal natural polymorphism in targeted genes. Implementing High Resolution Melting technology using a LightScanner™ aims to improve the screening through put. Two goals were assigned:
-Implementation of the technique and efficiency assessment.
-Screening of the two core collections for coding allelic variants discovery in genes involved in brix and find potential links between observed phenotype and genotypic data collected.

2. Material and Methods

2.1 The EU-sol and CBSG core collections

EU-SOL is a European collaborative project on the genetic and genomic analysis of tomato and potato (http://www.eu-sol.net/). Project's module "screening natural diversity" focuses on the development of a core collection and the use of the project online BreeDB database (Finkers 2009). The core collection is composed of about 7000 accessions, including cultivars and domesticated along with representative wild relatives. This germplasm was provided by gene banks, institutes and plant scientists. In the summer 2007, the EU-SOL Core Collection was planted in Akko, Israel. In 2008 and 2009 extensive phenotyping on several agronomic traits

including brix was performed. The map displays accessions region of origin (see figure 3).

The Centre for BioSystems Genomics (CBSG) is public private consortium of major Dutch and international companies and academic plant scientists working on tomato (http://www.cbsg.nl/tomato.aspx), potato, *Arabidopsis* and other *brassica*. The CBSG tomato program aims to implement research on the genetic basis of tomato quality traits, notably to screen tomato varieties for the natural biodiversity of taste characteristics (van Berloo, van Heusden et al. 2008). For this research purpose a set of 94 cultivars known for their agronomic qualities has been designed. Metabolite profiling was performed on the set, including flavonoids, sugars and acids (Tikunov, Lommen et al. 2005).

2.2 Phenotyping core collections for Brix: two different sets

Two collections with two different backgrounds: The two collections were phenotyped under two distinct field trials in different locations. EU-SOL set in open field in Akko, Israel, CBSG set in greenhouses in Wageningen, Netherlands.

-EU-SOL Core Collection

This collection aims to gather the existing natural genetic diversity within the *lycopersicum* gender and wild relatives. The collection is under characterization (phenotyping, genotyping for population structure, metabolomic profiling). Based on available phenotyping data, subsets for specific traits are currently designed, notably for brix. In 2008, in Akko, brix was measured on 4500 accessions on one single fruit stored overnight in a cold room. Summer 2009, subsets of the collection were selected according to a list of traits of interest such as fruit color, firmness, epidermis, size, and brix. The brix phenotyping panel consisted of 330 accessions, determinate and indeterminate types, which showed high or low brix value in 2008. The Collection was evaluated according to a phenotypic catalog of key descriptors (Tadmor 2009). Collecting data for brix, fruit weight and fruit size was the priority. For each accession, 5 fruits were harvested, weighted, measured for their height and diameter. Brix was measured with the standard refractometer method. A data set was obtained (see table 1). Based on these data a subset of extremes (150 accessions) from the 330 accessions BRIX panel was repeated for the sugar content measurement. Fruit sampling was performed on the 150 accessions for metabolomic profiling and sugar dosage (fructose,fructose 6 P, glucose, glucose 6 P, sucrose).

-CBSG set

This set is a group of top cultivars (hybrids) from breeding companies; it was extensively characterized for a large panel of traits including brix, metabolites profiling sugars (glucose, fructose, sucrose), citric and malic acid (Tikunov, Lommen et al. 2005). Brix was previously measured. CBSG used two grams of fruit flesh for each accessions, 100 µl of supernatant was used on refractometer for measurement. The related data set was provided by CBSG.

2.3 Genotyping candidate genes for Brix: Eco-TILLING on core collections

A brix set of four 96-wells PCR plates was designed. Three plates for EU-SOL core collection accessions and one plate from CBSG including the 94 accessions. (See 2.4)
Method assessment:
Before screening for conserved domains in the candidate genes, method was assessed for different criteria: capability to use GenomiPhi DNA for HRM and method accuracy as well as throughput of the HRM method in comparison with other screening method currently used in Eco-TILLING.
Thus, two testing experiments were implemented as well as a literature study:

- A comparative study between HRM polymorphism detection and Sanger sequencing on a whole plate.
- A comparison between HRM polymorphism detection and massive parallel sequencing technique KeyPoint™ (Rigola, van Oeveren et al. 2009) method from Keygene N.V. (Wageningen, the Netherlands)
-Eco-TILLING using HRM was also compared to Eco-TILLING using endonuclease through comparison of results obtained by Mosca et al (2009).

2.3.1 DNA material

EU-SOL Core collection DNA material was isolated, increased with Illustra GenomiPhi V2 amplification kit (GE Healthcare, Little Chalfont, Buckinghamshire, UK) and provided by Keygene N.V. for 2250 accessions. Dosage was performed using Lambda DNA (Invitrogen, Carlsbad, California, USA). Genomiphi DNA amplification or Whole genome amplification (WGA) use a bacteriophage polymerase Phi29 polymerase that performs DNA high fidelity amplification of the complete template genome in 10 kb fragments (Sato, Ohtsuka et al. 2005)

CBSG core collection DNA material (Genomic DNA with a 30 ng/µl concentration) was provided by Sjaak van Heusden.
For the production of reference DNA samples, Moneymaker seeds were provided by Fien Meijer and grown in greenhouse during two weeks. Tissue samples were ground with Retsch grinder (Retsch GmbH, Haan, Germany), DNA isolation was performed with King Fisher maxi prep protocol (Fisher Thermoscientific, Waltham, Massachusetts, USA).

-In the three EUSOL PCR plates, a reference control-I (Genomic DNA of tomato cultivar Moneymaker), a reference control-II (heteroduplex of genomic and GenomiPhi DNA from Moneymaker), a second reference line (Heinz 1706) and a negative control (MilliQ water) were added on EU-SOL plates, in respective position A1;C1;B1;D1 (see figure 4).
Comment:
Reference control-I was added to check potential contamination of the DNA, Moneymaker DNA strand against itself should not show variants.

Reference control-II was added to control GenomiPhi DNA sequence fidelity. GenomiPhi DNA is an amplification product from bacteriophage. Thus, a mismatch can occur due to polymerase lack of fidelity and can be checked. Reference control-II is a duplex between the template (genomic DNA, from Moneymaker cultivar) and a tested line (GenomiPhi Moneymaker).
-In the CBSG plate only two controls are present: thereference line (the tomato cultivar Moneymaker) and a negative control (MilliQ water) as CBSG core collection DNA is genomic. Reference and control are in position A1 and H12 respectively (see figure 4).

2.3.2 Gene sequence and annotation

Candidate genes were selected according to literature. Most of the candidates are belonging to multigenic families with several orthologs or paralogs. Depending on available data, genomic DNA (gDNA) or coding DNA (cDNA) sequences of targeted gene were obtained from NCBI (www.ncbi.nlm.nih.gov). When targeted gene sequence was not available on tomato, its homolog from Arabidopsis thaliana was taken and a tBLASTn analysis on the SOL and TIGR databases (www.solgenomics.net;www.compbio.dfci.harvard.edu/tgi/plant.html) was performed to find putative orthologs. Potential conserved domains were searched and identified with inter pro scan online tool, using the EMBL-EBI database (www.ebi.ac.uk/interpro/) to give a prediction on functional domains occurrence.
Intron positioning on cDNA sequences was performed using the SGN intron finder tool (www.solgenomics.net/tools/intron_detection/find_introns.pl) that predicts putative intron positions on tomato gene sequence based on available information from homolog gene in Arabidopsis. Each candidate was annotated according to the pipeline define above (see figure 5).
Selected sequences were annotated for coding regions/non coding region as well; conserved domains were indicated using APE software (www.biology.utah.edu/jorgensen/wayned/ape).

2.3.3 Primer design and PCR reaction

Primers pairs were design to amplify each conserved domain (CD) within each targeted gene (one to three CD per gene). Primers design for HRM includes several parameters: a high Tm (Tm>60°) and amplicon size higher than 200 bp for sequencing purposes and smaller than 400 bp (up to 500 bp in some cases) to fit with Lightscanner™ instrument (Idaho Technology Inc.) requirements. For multigenic families, DNA sequence alignment of orthologs was done to detect non homolog regions using NCBI BLAST tool. Non homolog regions were targeted to increase primer specificity. In case were cDNA was available only, the positions of putative introns obtained from the SGN intron finder were taken into account in order to minimize the potential size of the amplified fragment. Also, several pairs of primers were design for a same amplicon in order to maximize chances to obtain a targeted fragments fitting with the different constrains. Each primer pair was tested and characterized using gradient PCR with a Tm range from 55° to 65°.
PCR reaction mix was 10 µl, with 0.25 µl both forward and reverse primers (10µM), 1 µl LC green

plus dye (Idaho Technology Inc), 0.4 µl dNTPs (5µM), 2 µl 5x Phire™ Reaction Buffer, 0.1 µl Phire™ Hot Start DNA Polymerase (Finzymes) and sterilized water. PCR reaction settings were 40 cycles, 5 seconds at 94°C, 20 seconds at Tm (Based on the tested primer pair Tm was ranging from 60 to 65°) and 20 seconds at 72°C followed by a desannealing – reannealing step for heteroduplex formation: 94 °C, 30 seconds and cooling down to 25 °C, 30 seconds.

2.3.4 Screening for allelic variants

Allelic variant screening and analysis were performed immediately after PCR reaction. This procedure was based on High-Resolution Melt analysis, using Lightscanner™ System (Idaho Technology, Inc). The melting temperature range was set from 75°C to 95°C. Melting curves with high fluorescent intensity (higher than 800) were chosen for analysis. Putative polymorphic genotypes were distinguished by plotting the fluorescence difference between reference line and tested accessions normalized melting curves. Positive wells showing lower or higher melting temperature compare to reference were selected. After screening and analysis, PCR products of selected lines were sequenced.

2.3.5 DNA sequencing

Sequencing method is based on Sanger sequencing with ddNTPs as chain terminator. Sequencing mix contains 4µl DETT mix (1µl DETT dye, 3µl DETT buffer, Amersham), 1 µl primer (1µM), 100ng PCR products, and sterilized water for a total volume of 10 µl. PCR reaction for sequencing was 25 cycles, 20 s. at 94°C, and 15 s. at 50°C and 60 s. at 60°C. The sequencing runs were performed by Greenomics (www.greenomics.wur.nl/UK) with ABI Prism 3700 sequencer (Applied Biosystems, Carlsbad, California, USA). To improve sequencing quality, samples were purified on Sephadex™ G-50 96 microspin plate (GE Healtcare).

2.3.6 SNPs analysis

DNA sequences were analyzed by using SeqMan module of Lasergene software package (DNAstar Inc, Madison, Wisconsin, USA) to characterize the putative allelic variants observed from the HRM screening. The DNA sequence of the amplicon carrying SNPs was translated into amino acids sequence using the ExPASy website's using online translating tool (http://www.expasy.ch/). Alignments of modified protein sequence against reference was performed using BLASTp (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). Amino acid changes were positioned and annotated on the reference were sequence. Prediction effect of non synonymous SNPs on amino acid sequence and protein structure were analyzed by online prediction SIFT software (http://sift.jcvi.org/) to observe whether the amino acid substitution affects protein function. Tomato lines showing to be affected (predicted as "Deleterious") by SNP change were listed.

3. Results

3.1 Phenotyping the EU-SOL Core Collection

EU-SOL Core collection phenotyping was performed in the field. Brix, fruit size and fruit weight were measured on more than 330 accessions during the summer 2009, data was collected and added to BreeDB database. Figure 6 shows a comparison between 2008 and 2009 data sets with a low correlation (r^2= 0.34), indicating a strong environmental effect. Phenotyping conditions and protocol were different between EU-SOL and CBSG core collections (see 2.2). For both set, metabolic data and sugar dosages were performed, information to related accessions were added when available. As well, descriptive statistics were calculated (Mean value, Min value, Max value, Standard deviation)
Figure 7 shows distribution of brix values for three sets: CBSG (94 accessions), EU-SOL Akko 2008 (293 accessions) and EU-SOL Akko 2009 (139 accessions, 2 measurements averaged).
CBSG shows a distribution of its accession brix value (3 to 8) narrower than for EU-SOL 2008 (1 to 10) and 2009 (3 to 10). As well, CBSG brix set mean (5.0) is lower than EU-SOL brix set mean (5.5). Standard deviation is also lower for CBSG (0.89) than EU-SOL (2.12) as summarized in as referred in table 7 and table 8.

3.2 Bioanalysis and genotypic data

Use of GenomiPhi DNA
The first PCR results with GenomiPhi DNA were inconsistent. PCR products on gel (agarose 1%, 80 v, 45 min) gave SMIRs. A too high DNA concentration was suspected. Thus concentration was checked with Nanodrop spectrophotometer (Nanodrop ND-1000, Thermoscientific, Wilmington, Delaware, USA) at first. Concentration measurements were inconsistent. Thus Lambda DNA (Invitrogen) was used and sample loaded on gel was for a better quantification. Dilution rate had to be dramatically increased, a 160x dilution of GenomiPhi DNA was observed as being the best.

Polymorphism detection accuracy and through put efficiency
-Eco-TILLING with HRM screening method was tested for efficiency by comparing its output to complete Sanger sequencing. Both techniques were performed using CBSG set (genomic DNA). Conserved domain of an invertase, *lin8*, was targeted.
-With whole plate sequencing (94 lines), 72 polymorphisms corresponding to 4 variants were detected with Sanger sequencing.
-With Eco-TILLING, 51 lines showed putative variants and were sequenced. 59 polymorphisms corresponding to 3 variants were observed. Eco-TILLING covered 82% of the total polymorphism observed by sequencing 48% of total lines.

-Eco-TILLING vs Keypoint

Eco-TILLING was also tested with same plant material and PCR fragment as Rigola *et al* did to assess Keypoint technology (Rigola, van Oeveren et al). Rigola *et al* used 92 accessions genomic DNA from the EU-SOL core collection and found 6 haplotypes related to eIF4E gene by resequencing three amplicons. With the same primers we performed amplification and screened with HRM the 92 accessions using GenomiPhi DNA. Experiment is detailed in annex 1. If polymorphism was found for amplicon 1, E1A1, (represented with a blue star), we could not identify properly the other polymorphism in E2A2 (red star) and E3A3 (orange star).

-Eco-TILLING HRM vs EcoTILLING LICOR

Eco-TILLING HRM was also compared to Eco-TILLING using endonuclease and LICOR gel from first results obtained by Mosca et al. (2009). With Eco-TILLING endonuclease ENDO 1 method, Mosca *et al* screened six large genomic DNA fragments (1000-1500 bp each) corresponding to genes involved in carotenoids pathway on 89 tomato lines. Each fragment includes intronic and exonic DNA. A total of 144 polymorphism (114 in introns, 30 in exons) were detected, 8 non synonymous SNPs in exonic region (Mosca, Minoia et al. 2009). A total of 667 Kb was screened, a frequency of 1 polymorphism for 4.6 Kb and 1 exonic SNP for 22 Kb (see table 2).

As a comparison, Eco-TILLING with HRM screening platform for conserved domains (located in exons) in invertase gene family: *lin*6*, lin*7*, lin*8, *tiv*1 (*lin*5 was not included due to poor sequencing data). On 370 individuals 188 polymorphisms were detected with 26 SNPs in exonic region for a total of 670 Kb screened, a frequency of 1 polymorphism for 3.3 Kb and 1 exonic SNP for 24 Kb (see table 2).

-Screening and sequencing for SNPs

Twenty five candidate genes were listed and 38 conserved domains localized (see table 3). 65 primer pairs were designed and tested, 30 were successful in amplification and fit with required amplicon length after testing with gradient PCR and migration on agarose gel (see table 4). Twenty pairs were used to screen Invertases (*lin*), fructokinase (*frk*), malate synthase (*mls*), sucrose synthase (*susy*) , UGPase and isocitrate lyase (*Icl*), for a total of 80 PCR plates screened. Polymorphism was identified for Invertase 7 (*lin*7) and Invertase 8 (*lin8*) and acid invertase (*tiv1*), Isocitrate Lyase (*Icl*) and Fructokinase2 (*Frk*2) Different issues were encountered with other candidate genes: no amplification (*GPI*, *PGM*) lack of product specificity (*AGPase*), low quality of sequenced products (*Fxk*, *Hxk*) (see overview: table 5). Finally, a total of 1126 Kb sequence was screened. 188 polymorphisms were identified for those fragments, among them, 110 SNPs and 78 Indels forming 14 haplotypes Thirty one SNPs are non redundant distributed on 52 accessions (16 accessions from EUSOL; 37 accessions from CBSG). All polymorphisms are listed and described in table 6.

CBSG panel presents 79 SNPs displayed on 41 accessions with variants corresponding to 8 haplotypes for two genes, *lin*7 and *lin*8. Fourteen SNPs are non redundant (see table 7).

For the same experiment, the EU-SOL panel presents 31 SNPs on 19 accessions with variants corresponding to 9 haplotypes for six genes including *lin*8 but also *SI*, *Icl*, *Frk*2 and *Susy*3. Seventeen SNPs are not redundant (see table 8).

All Indels are in intronic regions. Seven SNPs are in exonic region, one is coding in soluble invertase (*tiv*1) sequence in core collection accession 2701 (*Solanum penellii*). Allelic variant

induces an amino acid change from Lysine (K) to Glutamic acid (E). This variant is predicted as non deleterious by SIFT software. The variant is located in position 307 on protein sequence, neighboring an active site for glycosylation (WECVDF) as previously described by Klann (Klann, Yelle et al. 1992; Klann, Chetelat et al. 1993).

4. Discussion

-Method assessment: Genomic vs GenomiPhi DNA.

Comparison of the work performed with the Keypoint protocol (Rigola, van Oeveren et al. 2009), allowed detection of Identical SNPs; nevertheless some haplotypes could not be re-identified. As well important rate of false positive occurred. This lack of accurate detection and false positives could find answers in the dosage of GenomiPhi DNA. Indeed, if GenomiPhi DNA as proven to be HRM compatible (Cho, Ciulla et al. 2008) it has a constraining aspect concerning its accurate dosage. As proposed by the manufacturer a supplemental purification step of the GenomiPhi product should have been be performed (GE_Healthcare 2006). This is recommended in the case of dosage method using UV absorbtion. It can be adequately replaced by a titration dye such as Pico Green assay (Invitrogen, Carlsbad, California, USA).

Moreover, if our titration with lambda DNA gave satisfying result for a good amplification it may have induced some variation in the results. This dosage aspect may have an effect in a method using DNA heteroduplex. Indeed, heteroduplex construction requires adding a reference template to tested genotype in an equimolar way. HRM standard working DNA quantity ranges from 30 to 50 µg (Gady, Hermans et al. 2009). An unbalanced mixture due to low accuracy dosage may induce false positive and negative at the HRM screening step as underlined by Cho *et al*.

Also Sephadex™ G-50 96 microspin plate (GE Healtcare) was used for purification, if quality improvement could be seen at the sequencing step, potential bias at the detection step could not be corrected as the purification was performed after the screening.

- Method assessment: SNPs detection

Eco-TILLING using HRM implementation allowed detection of polymorphism. It showed a polymorphism detection of 82% compare to extensive Sanger sequencing on genomic DNA from CBSG set. This result is encouraging for further investigation of polymorphism and the use of HRM as screening method for Eco-TILLING. Indeed, the screening step reduced by 50% the amount of lines to sequence, reducing de facto sequencing and labor cost. This positive result obtained with genomic DNA also tends to confirm our hypothesis about GenomiPhi DNA and related issue.

-Screening for polymorphism, comparison of techniques:

This can be illustrated if we compare our results for the screening of invertases to the results from Mosca *et al* using Eco-TILLING with ENDO 1 endonuclease (Mosca, Minoia et al. 2009). With a similar amount of screened sequences (667 and 670 kb), total polymorphism found by Mosca *et al* is relatively similar (159 and 144 respectively) and frequency as well (1/4635 bp and 1/3393 bp). Concerning SNPs in coding sequences, results are similar: 30 are found by Mosca *et al* and 26 by

Eco-TILLING with HRM. Qualitatively, they found height non-synonymous SNPs and one was identified in our study. Notable differences between the two experiments can be seen in the amount of individuals screened and the type of targeted fragments. Difference of non synonymous SNPs between the two experiments may find an explanation in targeted regions. Mosca *et al* focused on whole genes (introns + exons), our objective was focused on conserved domains (mainly exons) which show a lower variability in average.

We selected conserved domains of key enzymes of the carbohydrate metabolism. As those proteins are critical for plant's development they show a high level of conservation in tomato but also a comparable evolutionary constrains in related species such as potato (Fridman and Zamir 2003). Thus, selection of targeted regions for polymorphism discovery is critical in an autogamous species such as tomato.

-Screening for polymorphism: a EUSOL-CBSG collection comparison

A brief comparison of the brix data of respective groups shows a narrower brix distribution for CBSG compare to EU-SOL (see figure 7). As well, we may expect a different outcome concerning the proportion of observable variability within sequences of targeted genes. Based on compared polymorphism between EU-SOL and CBSG sets (see results), an intuitive hypothesis for a narrower genetic background could be formulated for the CBSG set compare to EU-SOL. Less candidate genes are polymorphic in CBSG compare to EU-SOL. Three haplotypes showing the highest brix values (7.6, 7.8 and 9.5) are three wild relatives, *S. cerasiforme*, *S. pimpinelifolium* and *S. pennelli*.

Differences for allelic variation and frequency can be seen as well between the EU-SOL and CBSG subsets. The CBSG set presents a high number of accessions with allelic variants (41) for two genes (*lin*7, *lin*8) and 8 haplotypes. Oppositely, EU-SOL displays only 19 variants on six genes (*lin*8, *SI*, *Icl*, *Frk*2 and *Susy*3) and 6 haplotypes. Only *lin*7 and *lin*8 show allelic variants in the CBSG set which may illustrate a reduction of the overall genetic diversity. The same two genes show a higher allelic frequency for certain SNPs (SNP 257 and 383) in CBSG (F=0.19) than in EU-SOL (F=0.001). As CBSG is mainly constituted of elite material, we can make an educated guess that those loci are under a positive selection.

-Screening for polymorphism: Non synonymous SNP

In EU-SOL CC2701 accession (*S. pennelli*), an amino acid change occurs in soluble invertase (*tiv*1) protein sequence the original amino acid was polar basic (lysine, K) is replaced by a polar basic amino acid (glutamic acid, E). If the change is predicted as non deleterious by SIFT, it may have an influence on the protein structure as it flanks the main active site. Especially as glutamic acid does not occur at this position in any other invertase family proteins (see figure 8). Like other invertase *tiv*1 cleaves sucrose into hexoses. It was demonstrated that inhibition of *tiv*1 increase sucrose and reduce hexose concentration in the ripe fruit. It was suggested that *tiv*1 (or acid invertase) controls sugar composition in tomato fruit and this differential growth is correlated with the last stage of fruit development (Klann, Chetelat et al. 1993). A recent study showed that acide invertase activity change independently of the other enzymes. As well, for certain accessions such as *S. pennelli*, expression profile of acid invertase is much lower during all fruit developmental stages than for other measured cultivars (Steinhauser, Steinhauser et al. 2010).

-Bridging phenotypic and genotypic data

In EU-SOL set phenotyping issues concerning sugar dosage and metabolite profiling were faced. Experimental design changed and brix set was changed along the project, thus we could not obtain relevant data for the last 2010 greenhouse trial. As well, replicates were not available or were available but in different conditions. Nevertheless, we could identify haplotypes with strong brix values (EUSOL haplotype 6) for which sugar dosage data is missing.

In CBSG set, links between sugar content and allelic variation observed in the targeted genes could not be made as detected polymorphism did not present non synonymous SNP in coding regions.

5. Conclusion

In the current context, Sanger sequencing is declining to the profit of massive throughput "next generation sequencing technologies"(ie: pyrosequencing or the upcoming nanopore technology). If those techniques bring a major advance in knowledge of genomes and characterization, they are designed for large projects with a specific expertise and remain time consuming. Eco-TILLING is an efficient alternative for polymorphism discovery in targeted regions. HRM screening method offers results in a two hours time frame including PCR reaction. This yields a six hours time gain compare to endonucleases digestion screening method (Parry, Madgwick et al. 2009). As a comparison a candidate gene approach 454 run and related DNA libraries to screen 96 accessions for 30 candidate genes ask a three weeks intensive lab work (PCR amplification a step excluded). As well this flexibility of use makes Eco-TILLING a method of choice for genotyping natural collection, as shown on barley for resistance genes introgression (Mejlhede, Kyjovska et al. 2006).

Compare to other screening techniques (endonuclease and LICOR), Eco-TILLING using HRM screening platform improve throughput and allows larger accession set to be screened. It demonstrated an equivalent detection capacity. A careful attention should be paid to the targeted sequence. As HRM and Sanger sequencing techniques together restrict fragment size to 200 to 400 bp. Thus, obtaining corresponding amplicon using a cDNA reference may be difficult and time consuming in some cases. The use of gDNA simplifies the design procedure and the recent release of tomato genome sequence will be extremely helpful to overcome this difficulty. This is particulary true in case of small regions flanked with introns such as the conserved domain regions targeted here. This difficulty was also faced in other crops such as melon (*Cucumis melo*) for highly conserved eIF4E resistance genes to powdery mildew (Nieto, Piron et al. 2007). In case of too low natural polymorphism for a specific gene, a TILLING approach can be implemented to create variability, with a similar protocol as described by Gady *et al* (2009).

As perspective methodological improvement of GenomiPhi DNA screening can be done. Further investigations can be made concerning SNPs. Confirm them with re-sequencing, measurement of induced amino acid changes performing enzyme activity test. 3D protein modeling can be implemented to evaluated impact of non synonymous SNPs and induced changes on protein structure.

6. Figures and tables titles and legends

Figure 1: Schematic representation of the metabolic pathway of Suc-to-starch synthesis. Targeted genes are circled in green and named with * in related table. Dashed lines indicate enzyme inhibition of the sucrose synthase and fructokinase reactions (adapted from Zamski and Schaffer, 1996).

Figure 2: Lightscanner™ software snapshot.
Left part the tested 96-well plate is represented. Right part two graphs display the melting curves patterns related to each well of the plate, including the reference (A1). Melting curves with a different pattern are selected (in black) and may show a polymorphism in the related tested line.

Figure 3: Geographical origins of EU-SOL core collection accessions (source : EU-SOL website)

Figure 4: DNA Plate layouts designed for Eco-TILLING Brix set.
-a) EU-SOL plate 1 ; b) EU-SOL plate and c) EU-SOL plate 3.
For each plate:
-A1: reference 1 control-I (Moneymaker cultivar genomic DNA);
-B1: reference 2 (Heinz 1706 cultivar)
-C1: reference 1 control-II (Moneymaker GenomiPhi DNA),
-D1: negative control (MilliQ water).
-d) CBSG plate:
- A1: a reference (Moneymaker genomic DNA).
-H12: a negative control (MilliQ water)

Figure 5: Annotation pipeline for Eco-TILLING candidate genes.
Annotation procedure used for each candidate gene, including related software

Figure 6: Distribution diagram and scatterplot of EU-SOL accession brix values.
Brix data from EU-SOL Akko 2008 EU-SOL Akko 2009 are compared

Figure 7: Distribution diagram and table of brix values from CBSG and EU-SOL brix set.
Brix data from EU-SOl 2008 and EU-SOL 2009 and CBSG datasets are shown

Figure 8: Invertases (*lin*5, *lin*6, *lin*7, *lin*8 and *tiv*1) protein alignment with consensus sequence. Conserved domain is underligned, position 307 amino acids are framed in green. *tiv*1 amino acid change on position 307 is highlighted in yellow.

Table 1: a) b) EU-SOL Akko 2009 brix list.

Columns left to right: accession number, field code, brix and fruit weight values 1st measurement, brix and fruit value 2nd measurement, phenotyping subset and average values are indicated.

Table 2: Comparison of Eco-TILLING using HRM screening with Eco-TILLING using ENDO1 and LICOR gel from Mosca et al (2009).

Columns from left to right: family Number of fragments, amplicon size, number of individual screened, total sequence screened, total polymorphism on gDNA and cDNA, non synonymous SNPs and related frequencies are listed

Table 3: Candidate gene list for Eco-TILLING.

Columns left to right: Gene, name NCBI accession code, gene full name, chromosome number, *Arabidopsis* ortholog DNA and protein code, conserved domains (CD) position and sequence are indicated.

Table 4: Eco-TILLING primer list.

Columns left to right: Primer code and sequence, reference sequence type used for design (cDNA or gDNA), melting temperature (Tm), GC rate, start and stop position on sequence, primer length, amplicon, observed size from agarose gel and potential introns are indicated.

Table 5: Eco-TILLING screening and sequencing step overview.

Gene names (number of isoforms X number of Conserved Domains), amplification step, HRM screening step, sequencing step and SNP detection step are indicated with related informations.

Table 6: a) b) c) d) e) List of detected polymorphism using Eco-TILLING on candidate genes.

Gene, plate position, accession number, polymorphism type, nucleotide change, coding or non coding region, position on reference sequence, related haplotype, potential amino acid change as well as position are indicated.

Table 7: CBSG haplotypes and related phenotypic values.

Genotype part includes, accessions number (REF), genes haplotype. Phenotype part includes brix and fruit weight values. Metabolites part includes Fructose, Glucose and sucrose dosage. Average values for whole collection (94 individuals) are shown.

Table 8: EU-SOL haplotypes and related phenotypic values.

Accessions number (REF) and taxa are shown in two first columns. Genotype part presents haplotype number for the different genes. Phenotype part includes brix and fruit weight values for 2008 and 2009 as well as averages. Metabolites part includes Fructose, Fructose-6-P, Glucose, Glucose-6-P and Sucrose dosage. Collection average values are shown .

7. References

Bai, Y. and P. Lindhout (2007). "Domestication and breeding of tomatoes: what have we gained and what can we gain in the future?" Ann Bot 100(5): 1085-1094.

Bovy, A., E. Schijlen, et al. (2007). "Metabolic engineering of flavonoids in tomato (<i>Solanum lycopersicum</i>): the potential for metabolomics." Metabolomics 3(3): 399-412.

Cho, M. H., D. Ciulla, et al. (2008). "High-resolution melting curve analysis of genomic and whole-genome amplified DNA." Clin Chem 54(12): 2055-2058.

Claeyssen, E. and J. Rivoal (2007). "Isozymes of plant hexokinase: occurrence, properties and functions." Phytochemistry 68(6): 709-731.

Comai, L., K. Young, et al. (2004). "Efficient discovery of DNA polymorphisms in natural populations by Ecotilling." Plant J 37(5): 778-786.

Finkers, R. (2009). BreeDB: A database suppporting quantitative aspects of plant breeding. Tomato genomics workshop.

Fridman, E., T. Pleban, et al. (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 U S A 97(9): 4718-4723.

Fridman, E. and D. Zamir (2003). "Functional divergence of a syntenic invertase gene family in tomato, potato, and Arabidopsis." Plant Physiol 131(2): 603-609.

Gady, A. L., F. W. Hermans, et al. (2009). "Implementation of two high through-put techniques in a novel application: detecting point mutations in large EMS mutated plant populations." Plant Methods 5: 13.

Gazzani, S., A. R. Gendall, et al. (2003). "Analysis of the molecular basis of flowering time variation in Arabidopsis accessions." Plant Physiol 132(2): 1107-1114.

GE_Healthcare. (2006). "Illustra Genomiphi DNA amplification kit." Product web protocol.

Gur, A. and D. Zamir (2004). "Unused natural variation can lift yield barriers in plant breeding." PLoS Biol 2(10): e245.

Heuvelink, E. (2005). Tomatoes, CABI.

Jenkins, J. (1948). "The origin of cultivated tomato." Economic botany 2(4): 379-392.

Kimura, S. and N. Sinha (2008). "Tomato (Solanum lycopersicum): A Model Fruit-Bearing Crop." Cold Spring Harb Protoc 2008(12): pdb.emo105-.

Klann, E., S. Yelle, et al. (1992). "Tomato fruit Acid invertase complementary DNA : nucleotide and deduced amino Acid sequences." Plant Physiol 99(1): 351-353.

Klann, E. M., R. T. Chetelat, et al. (1993). "Expression of Acid Invertase Gene Controls Sugar Composition in Tomato (Lycopersicon) Fruit." Plant Physiol 103(3): 863-870.

Kortstee, A. J., N. J. Appeldoorn, et al. (2007). "Differences in regulation of carbohydrate metabolism during early fruit development between domesticated tomato and two wild relatives." Planta 226(4): 929-939.

Lecomte, L., P. Duffe, et al. (2004). "Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds." Theor Appl Genet 109(3): 658-668.

Liew, M., R. Pryor, et al. (2004). "Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons." Clin Chem 50(7): 1156-1164.

Luckwill, L. (1943). "The genus Lycopersicon: An historical, biological, and taxonomical survey of the wild and cultivated tomatoes." Aberdeen Univ. Stud 120(1): 1-44.

Mejlhede, N., Z. Kyjovska, et al. (2006). "EcoTILLING for the identification of allelic variation in the powdery mildew resistance genes." Plant Breeding 125(5): 461-467.

Mosca, G., S. Minoia, et al. (2009). SNP discovery in tomato by EcoTILLING technology. EU-SOL 2009 Workshop on tomato genomics: new tools for improvement of yield and quality, Toledo.

Ng, P. C. and S. Henikoff (2003). "SIFT: Predicting amino acid changes that affect protein function." Nucleic Acids Res 31(13): 3812-3814.

Nieto, C., F. Piron, et al. (2007). "EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility." BMC Plant Biol 7: 34.

Parry, M. A., P. J. Madgwick, et al. (2009). "Mutation discovery for crop improvement." J Exp Bot 60(10): 2817-2825.

Parthiban, V., M. M. Gromiha, et al. (2006). "CUPSAT: prediction of protein stability upon point mutations." Nucleic Acids Res 34(Web Server issue): W239-242.

Pego, J. V. and S. C. Smeekens (2000). "Plant fructokinases: a sweet family get-together." Trends Plant Sci 5(12): 531-536.

Pflieger, S., V. Lefebvre, et al. (2000). "The candidate gene approach in plant genetics: a review." Molecular breeding 7.

Rigola, D., J. van Oeveren, et al. (2009). "High-throughput detection of induced mutations and natural variation using KeyPoint technology." PLoS One 4(3): e4761.

Roitsch, T. and M. C. Gonzalez (2004). "Function and regulation of plant invertases: sweet sensations." Trends Plant Sci 9(12): 606-613.

Sato, M., M. Ohtsuka, et al. (2005). "Usefulness of repeated GenomiPhi, a phi29 DNA polymerase-based rolling circle amplification kit, for generation of large amounts of plasmid DNA." Biomol Eng 22(4): 129-132.

Schaffer, A. A. and M. Petreikov (1997). "Sucrose-to-Starch Metabolism in Tomato Fruit Undergoing Transient Starch Accumulation." Plant Physiol 113(3): 739-746.

Steinhauser, M.-C., D. Steinhauser, et al. (2010). "Enzyme Activity Profiles during Fruit Development in Tomato Cultivars and Solanum pennellii." Plant Physiol. 153(1): 80-98.

Stevens, M., A. Kader, et al. (1979). "Potential for increasing tomato flavor via increased sugar and acid content." Journal of the American Society for Horticultural Science 104(1): 40-42.

Tadmor, R. (2009). Core collection phenotyping workshop. EU-SOL Core colelction phnotyping workshop, Akko (Israel).

Tikunov, Y., A. Lommen, et al. (2005). "A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles." Plant Physiol 139(3): 1125-1137.

Till, B. J., T. Zerr, et al. (2006). "A protocol for TILLING and Ecotilling in plants and animals." Nat Protoc 1(5): 2465-2477.

van Berloo, R., S. van Heusden, et al. (2008). "Genetic research in a public–private research consortium: prospects for indirect use of Elite breeding germplasm in academic research." Euphytica 161(1): 293-300.

Voet, D. and J. G. Voet (2004). Biochemistry, Wiley.

Zamir, D. (2001). "Improving plant breeding with exotic genetic libraries." Nat Rev Genet 2(12): 983-989.

Zamski, E. and A. Schaffer (1996). Photoassimilate distribution in plants and crops, Source-sink ralationships, Marcel Dekker Inc.

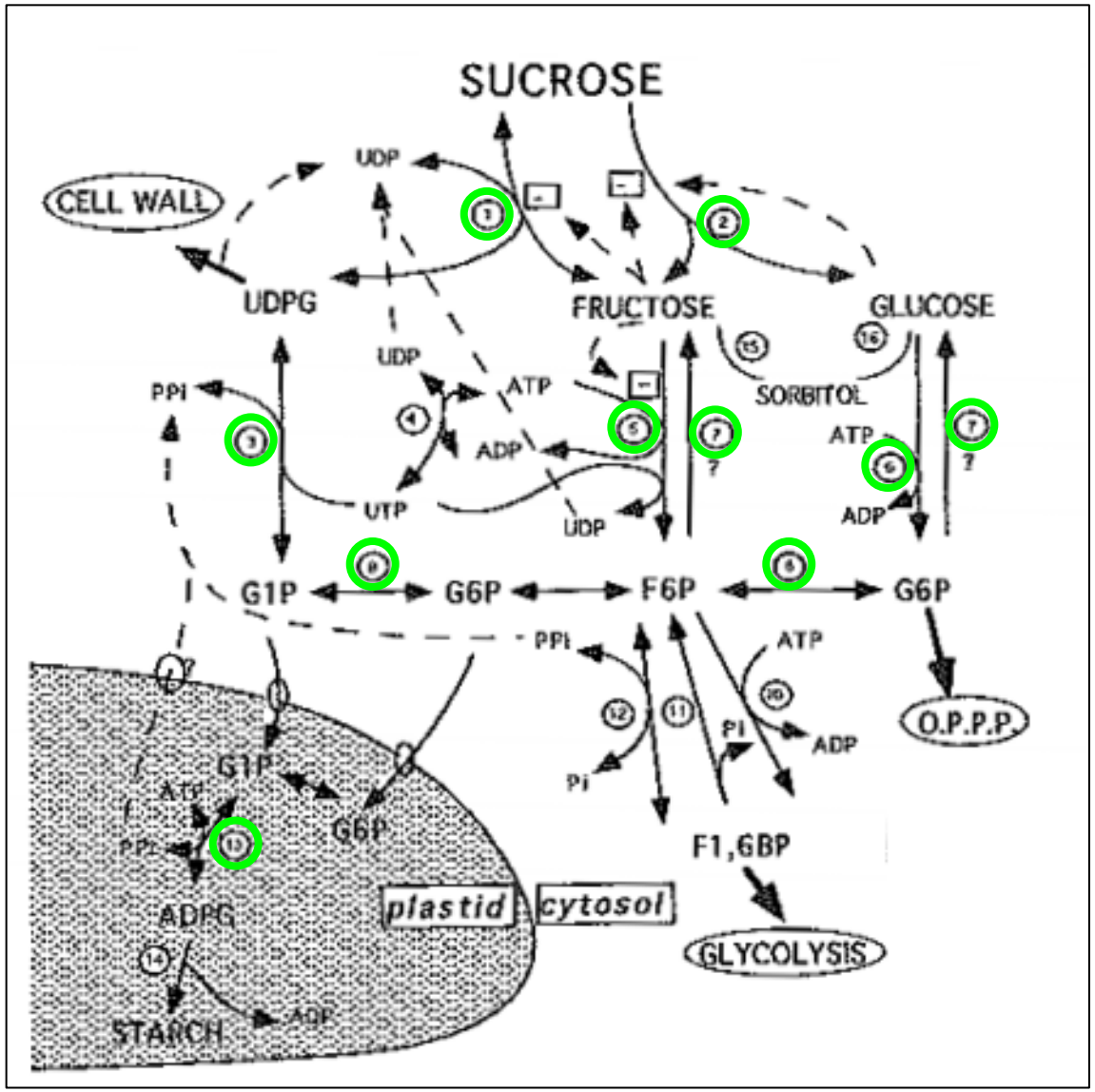
Zanor, M. I., S. Osorio, et al. (2009). "RNA interference of LIN5 in tomato confirms its role in controlling Brix content, uncovers the influence of sugars on the levels of fruit hormones, and demonstrates the importance of sucrose cleavage for normal fruit development and fertility." Plant Physiol 150(3): 1204-1218.

Appendix

Eco-TILLING on tomato core collections for brix: Implementing the methodology and first results

Tables and figures

FIGURES



1- Sucrose synthase *
2- Invertase *
3- UDPGk PPase *
4- NDP kinase
5- Fructokinase *
6- Hexokinase *
7- HexoseP Pase *
8- Phosphoglucoisomerase (GPI) *
9- Phosphoglucomutase (PGM) *
10- ATP-PFK
11- FBPase
12- PPi-PFP
13- ADPGlc PPAse*
14-- Starch synthase
15 - Ketose synthase
16 - Aklose reductase

*: Targeted gene families

Figure 1

Adapted from: **Zamski, E. and A. Schaffer (1996). Photoassimilate distribution in plants and crops, Source-sink ralationships,**

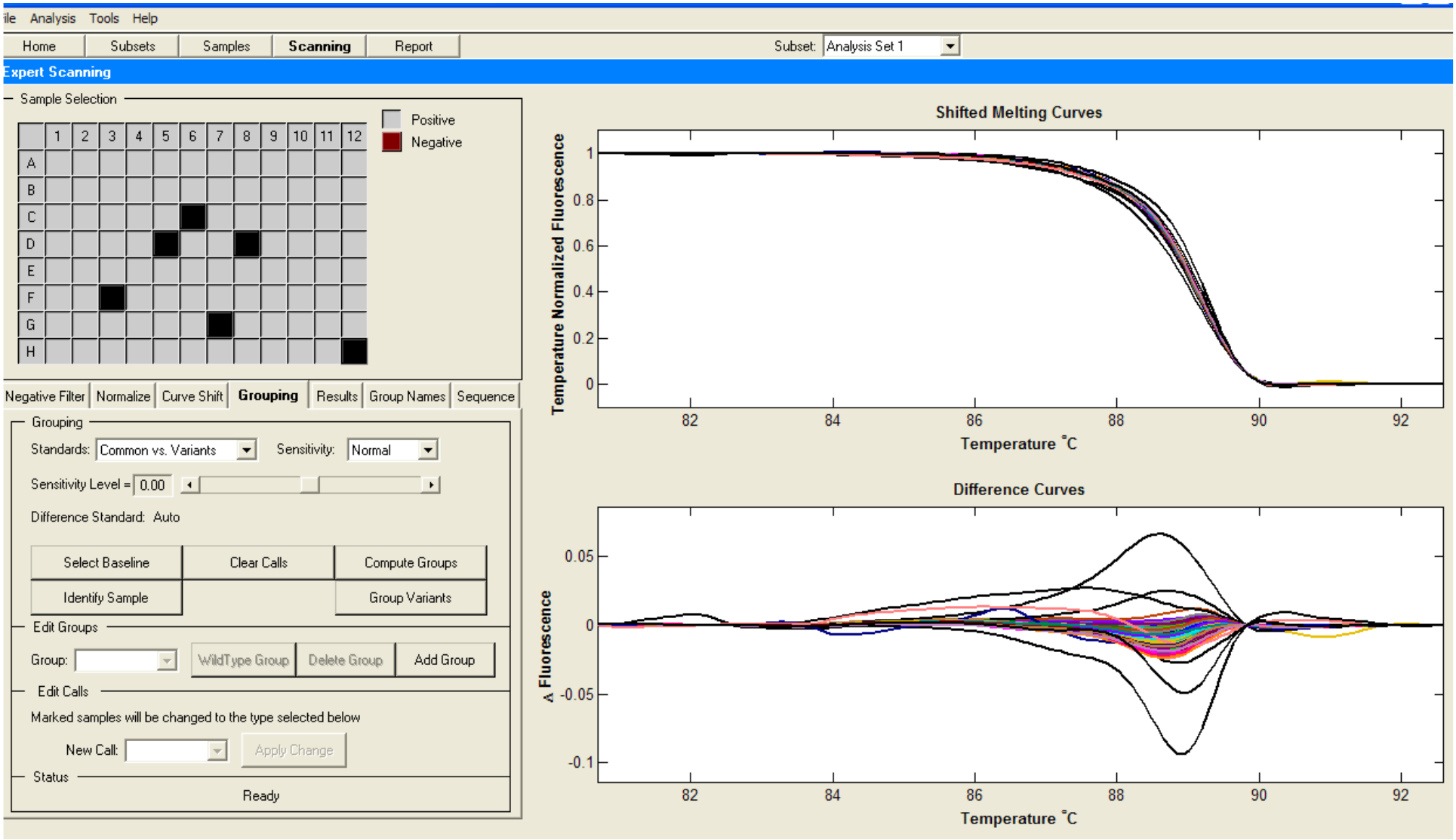

ile Analysis Tools Help
Home
Subsets
Samples
Scanning
Report
Subset: Analysis Set 1
Expert Scanning
Sample Selection
Positive
Negative
1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H
Negative Filter
Normalize
Curve Shift
Grouping
Results
Group Names
Sequence
Grouping
Standards: Common vs. Variants
Sensitivity: Normal
Sensitivity Level = 0.00
Difference Standard: Auto
Select Baseline
Clear Calls
Compute Groups
Identify Sample
Group Variants
Edit Groups
Group:
WildType Group
Delete Group
Add Group
Edit Calls
Marked samples will be changed to the type selected below
New Call:
Apply Change
Status
Ready
Shifted Melting Curves
Temperature Normalized Fluorescence
1
0.8
0.6
0.4
0.2
0
82
84
86
88
90
92
Temperature ˚C
Difference Curves
Δ Fluorescence
0.05
0
-0.05
-0.1
82
84
86
88
90
92
Temperature ˚C

Figure 2

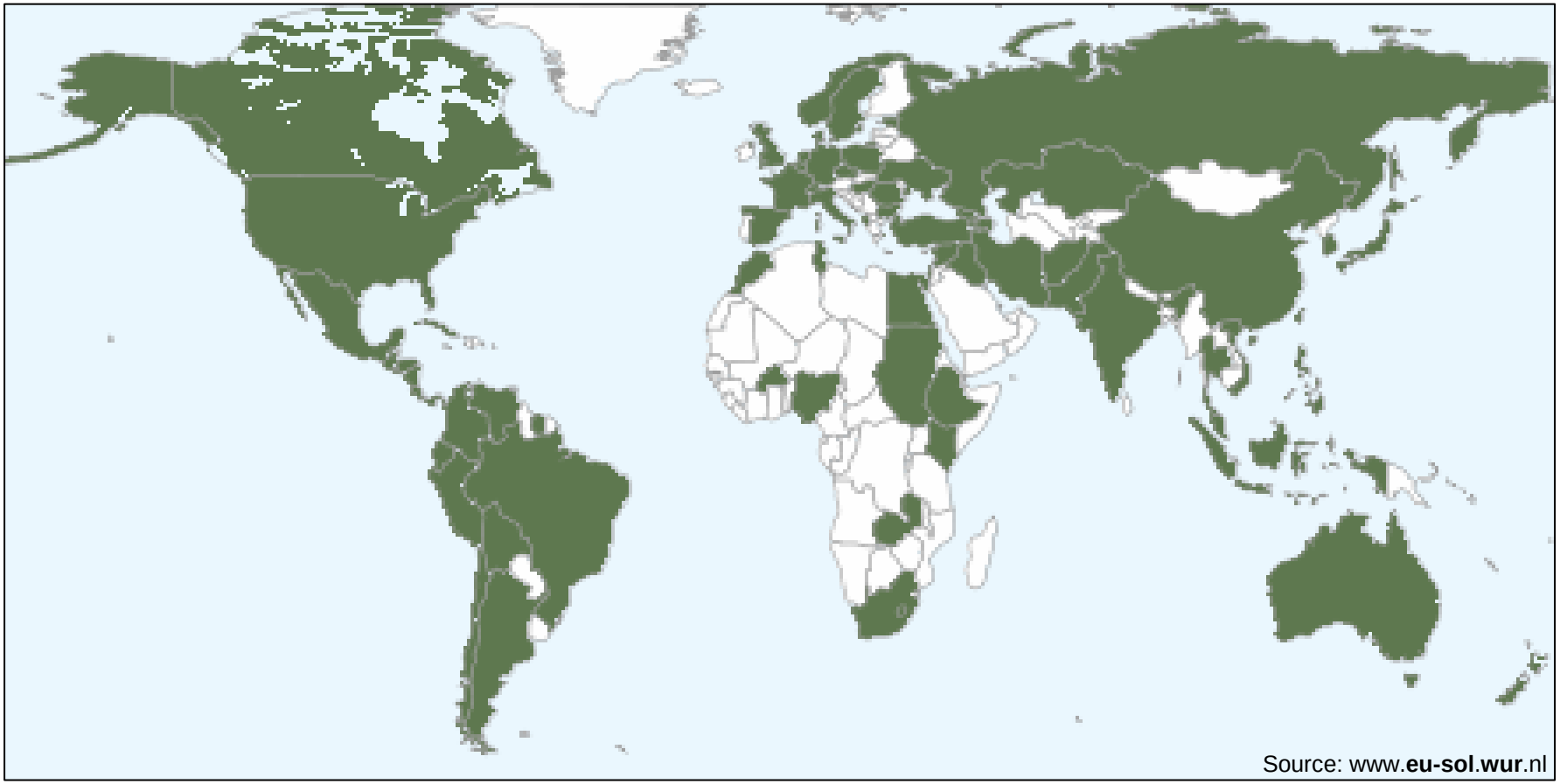


Figure 3

Plate #1	1	2	3	4	5	6	7	8	9	10	11	12
A	Moneymak	CC#03514	CC#05531	CC#01155	CC#03053	CC#02123	CC#02603	CC#02716	CC#02879	CC#2967	CC#2979	CC#03083
	CONTROL	Plate_01	Plate_01	Plate_01	Plate_01	Plate_01	Plate_01	Plate_01	Plate_01	Plate_02	Plate_02	Plate_02
		C12	E05	E08	F09	G06	H08	H10	H11	A03	A04	A08
B	Heinz 170	CC#03527	CC#03574	CC#03686	CC#06408	CC#05772	CC#02860	CC#03136	CC#00369	CC#04074	CC#01617	CC#03099
	Plate_03	Plate_02	Plate_02	Plate_02	Plate_02	Plate_03	Plate_03	Plate_03	Plate_03	Plate_03	Plate_03	Plate_03
	G09	C03	C09	E05	F10	A08	B04	C01	E09	F03	F11	H02
C	Moneymak	CC#00371	CC#00041	CC#01371	CC#01627	CC#00386	CC#01330	CC#00326	CC#03669	CC#00376	CC#00389	CC#00512
	Plate_15	Plate_04	Plate_04	Plate_04	Plate_04	Plate_04	Plate_04	Plate_04	Plate_04	Plate_04	Plate_04	Plate_04
	H05	B07	B08	B09	B10	B12	C07	D12	E02	E03	E04	E10
D		CC#01049	CC#01117	CC#01185	CC#01193	CC#01201	CC#01230	CC#01234	CC#01236	CC#01237	CC#01270	CC#01299
	BLANK	Plate_04	Plate_04	Plate_05	Plate_05	Plate_05	Plate_05	Plate_05	Plate_05	Plate_05	Plate_05	Plate_05
		H07	H12	B03	B06	B08	B12	C01	C02	C03	C09	D02
E	CC#01580	CC#01581	CC#01582	CC#01703	CC#02139	CC#02146	CC#02658	CC#02666	CC#02670	CC#02728	CC#02732	CC#02753
	Plate_05	Plate_05	Plate_05	Plate_05	Plate_05	Plate_05	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06
	D08	D09	D10	F04	G10	G11	B05	B08	B10	C09	C11	D06
F	CC#02754	CC#02895	CC#02898	CC#02990	CC#03001	CC#03002	CC#03006	CC#03040	CC#03051	CC#03023	CC#03028	CC#03150
	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06	Plate_06	Plate_07
	D07	E09	E11	F07	F08	F09	G01	G06	G10	G04	G05	A04
G	CC#03151	CC#03159	CC#03174	CC#03190	CC#03206	CC#03223	CC#03244	CC#03247	CC#03274	CC#03296	CC#03415	CC#03419
	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07
	A05	A08	A10	B01	B07	C04	C12	D02	D06	D07	E07	E08
H	CC#03433	CC#03456	CC#3494	CC#03496	CC#03497	CC#03533	CC#03575	CC#03586	CC#03611	CC#03613	CC#3614	CC#3648
	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07	Plate_07
	E11	E12	F02	F03	F04	F08	G02	G04	G08	G09	G10	H07

Figure 4 a)

Plate #2	1	2	3	4	5	6	7	8	9	10	11	12
A	Moneymak	CC#3650	CC#3657	CC#3659	CC#3673	CC#3967	CC#3981	CC#4073	CC#4089	CC#4109	CC#4186	CC#4199
	CONTROL	Plate_07	Plate_07	Plate_07	Plate_07	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08
		H09	H10	H11	H12	A02	B03	B06	B07	B08	B11	C02
B	Heinz 170	CC#4471	CC#4877	CC#4883	CC#5117	CC#5122	CC#5187	CC#5480	CC#5529	CC#5532	CC#5550	CC#5558
	Plate_03	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08
	G09	C06	C07	C08	C12	D01	D02	D04	D10	D11	E01	E02
C	Moneymak	CC#5612	CC#5651	CC#5653	CC#5694	CC#5747	CC#5774	CC#5775	CC#5783	CC#5784	CC#5788	CC#5789
	Plate_15	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08	Plate_08
	H05	E09	F04	F05	F08	G01	G06	G07	G08	G09	G11	G12
D		CC#5807	CC#5808	CC#1027	CC#1130	CC#1641	CC#2586	CC#6071	CC#6157	CC#6161	CC#6244	CC#6247
	BLANK	Plate_08	Plate_08	Plate_09	Plate_09	Plate_09	Plate_09	Plate_09	Plate_09	Plate_09	Plate_09	Plate_09
		H10	H11	G09	G10	G04	H05	A11	B02	B05	B07	B11
E	CC#6345	CC#6434	CC#6447	CC#6476	CC#6878	CC#1356	CC#258	CC#2705	CC#3321	CC#3439	CC#54	CC#2607
	Plate_09	Plate_09	Plate_09	Plate_09	Plate_09	Plate_10	Plate_10	Plate_10	Plate_10	Plate_10	Plate_10	Plate_11
	C02	C08	C09	D02	F08	E11	B01	F11	G07	B09	D05	B01
F	CC#2685	CC#2723	CC#2764	CC#2924	CC#933	CC#948	CC#3539	CC#4243	CC#4500	CC#5615	CC#2660	CC#2883
	Plate_11	Plate_11	Plate_11	Plate_11	Plate_11	Plate_11	Plate_12	Plate_12	Plate_12	Plate_12	Plate_13	Plate_13
	B02	B06	H07	B11	E08	E09	D01	D05	D09	F03	B11	C10
G	CC#3306	CC#4032	CC#4236	CC#4710	CC#4778	CC#5004	CC#5819	CC#5891	CC#304	CC#422	CC#448	CC#6628
	Plate_13	Plate_13	Plate_14	Plate_14	Plate_14	Plate_14	Plate_14	Plate_14	Plate_15	Plate_15	Plate_15	Plate_15
	F06	H10	B06	C12	D05	E08	H07	H09	F05	F10	F11	B05
H	CC#872	CC#941	CC#983	CC#1002	CC#1020	CC#1040	CC#1062	CC#1118	CC#1198	CC#1235	CC#1311	CC#1640
	Plate_16	Plate_16	Plate_16	Plate_16	Plate_16	Plate_16	Plate_16	Plate_17	Plate_17	Plate_17	Plate_17	Plate_17
	A10	C04	D05	D12	E11	F11	G10	B01	D12	E11	G09	H04

Figure 4 b)

Plate #3	1	2	3	4	5	6	7	8	9	10	11	12
A	Moneymak	CC#1820	CC#1903	CC#1915	CC#1949	CC#1959	CC#1960	CC#1966	CC#1967	CC#2018	CC#2026	CC#2034
	CONTROL	Plate_19	Plate_19	Plate_19	Plate_19	Plate_19	Plate_19	Plate_19	Plate_19	Plate_20	Plate_20	Plate_20
		A01	C11	D10	G4	H2	H3	H9	H10	D05	D12	E07
B	Heinz 170	CC#2038	CC#2340	CC#2348	CC#2431	CC#2435	CC#2460	CC#2472	CC#2490	CC#2516	CC#2523	CC#2534
	Plate_03	Plate_20	Plate_22	Plate_22	Plate_22	Plate_22	Plate_23	Plate_23	Plate_23	Plate_23	Plate_23	Plate_23
	G09	E11	B05	C01	G06	G10	A05	B03	C05	E02	E07	F02
C	Moneymak	CC#2669	CC#2712	CC#2777	CC#2782	CC#2783	CC#2953	CC#2959	CC#2960	CC#2971	CC#2972	CC#2994
	Plate_15	Plate_24	Plate_24	Plate_24	Plate_24	Plate_24	Plate_24	Plate_24	Plate_24	Plate_24	Plate_24	Plate_25
	H05	B08	C03	D06	D07	D08	G05	G10	G11	H05	H06	A02
D		CC#3054	CC#3075	CC#3271	CC#3287	CC#4228	CC#5170	CC#5581	CC#3124	CC#1073	CC#1620	CC#205
	BLANK	Plate_25	Plate_25	Plate_25	Plate_25	Plate_29	Plate_29	Plate_29	Plate_29	Plate_30	Plate_30	Plate_30
		B04	C07	G03	G12	A07	C03	C10	H04	H08	H12	E09
E	CC#3525	CC#3633	CC#4180	CC#4208	CC#882	CC#2701	CC#3979	CC#5057	CC#470	CC#883	CC#951	CC#79
	Plate_30	Plate_30	Plate_30	Plate_30	Plate_30	Plate_31	Plate_31	Plate_31	Plate_31	Plate_31	Plate_31	Plate_31
	C03	B02	B10	B11	H01	A05	A12	B11	D07	D10	E01	E12
F	CC#609	CC#610	CC#669	CC#744	CC#892	CC#915	CC#3026	CC#3107	CC#03567	CC#03426	CC#02890	CC#00891
	Plate_31	Plate_31	Plate_32	Plate_32	Plate_32	Plate_32	Plate_32	Plate_32	Plate_07	Plate_02	Plate_03	Plate_04
	H08	H09	A06	B05	C05	D07	H01	H02	F11	B06	H07	G07
G	CC#4362	CC#5575	CC#5801	CC#2036	CC#2653	CC#2696						
	Plate_08	Plate_08	Plate_08	Plate_20	Plate_24	Plate_25						
	C04	E03	H06	E09	B04	A12						
H												

Figure 4 c)

Plate #4 CBSG	1	2	3	4	5	6	7	8	9	10	11	12
A	MM (gDNA)	2	3	4	5	7	8	9	10	11	12	13
B	14	15	16	17	18	19	20	21	23	24	25	26
C	27	28	30	31	32	33	34	35	36	37	38	39
D	40	41	43	44	45	46	47	48	49	50	53	55
E	56	57	58	59	61	68	70	71	72	73	74	75
F	79	80	81	83	85	86	87	88	89	90	91	92
G	93	94	Mm(95)	96	98	99	100	102	104	110	111	112
H	119	120	121	123	124	125	126	127	131	132	1	Blank

Figure 4 d)

Primer design on candidate genes

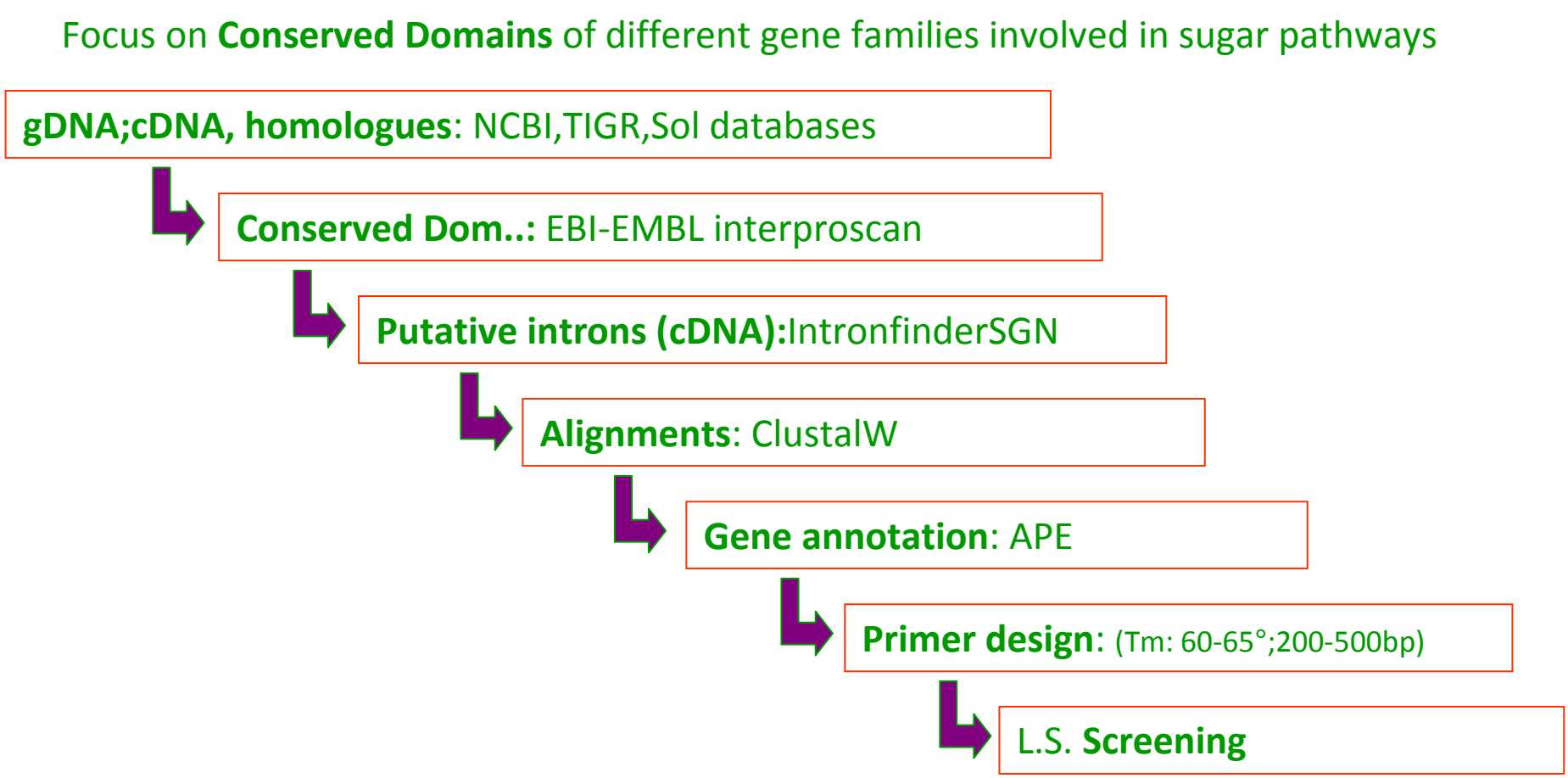
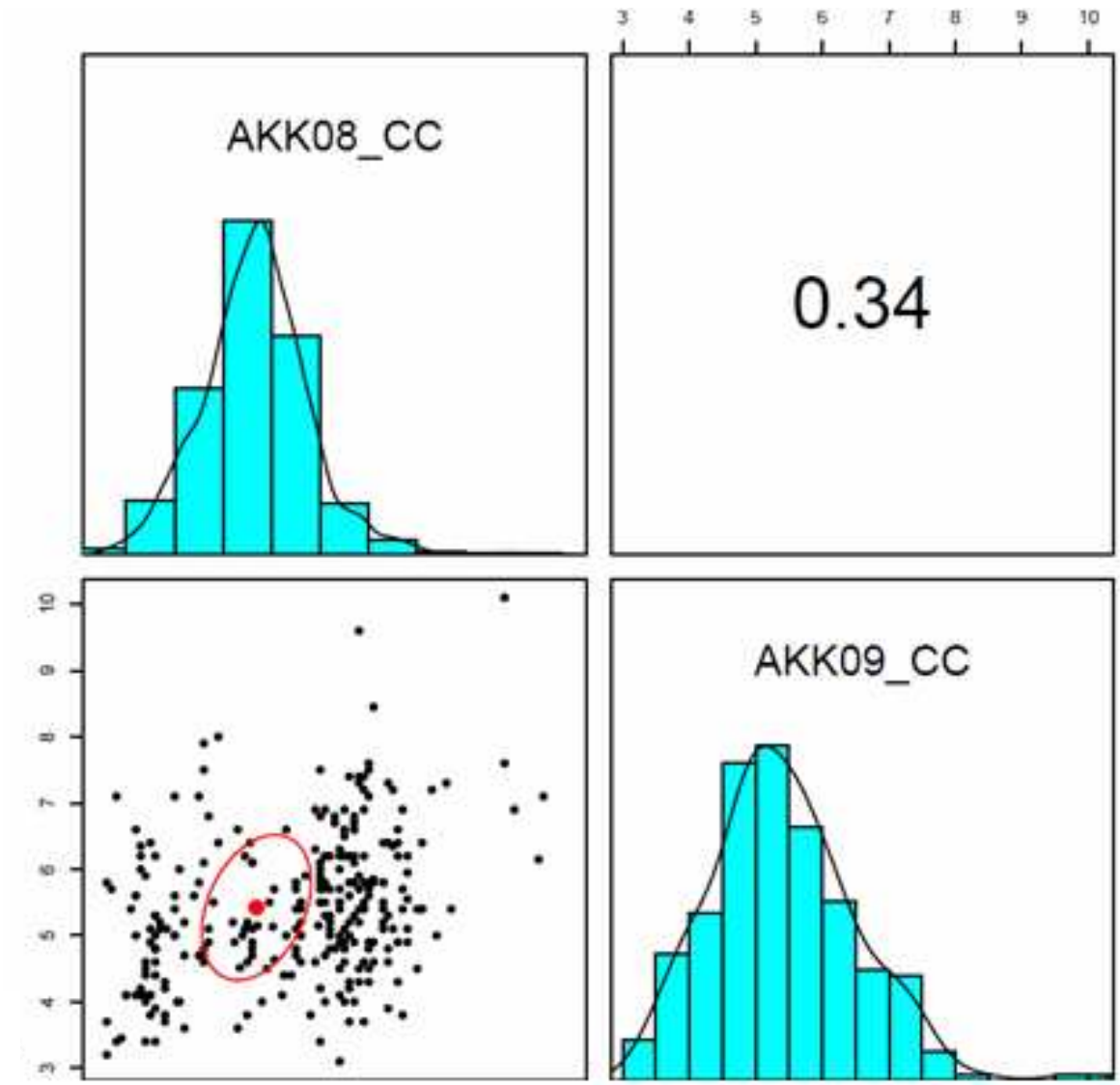
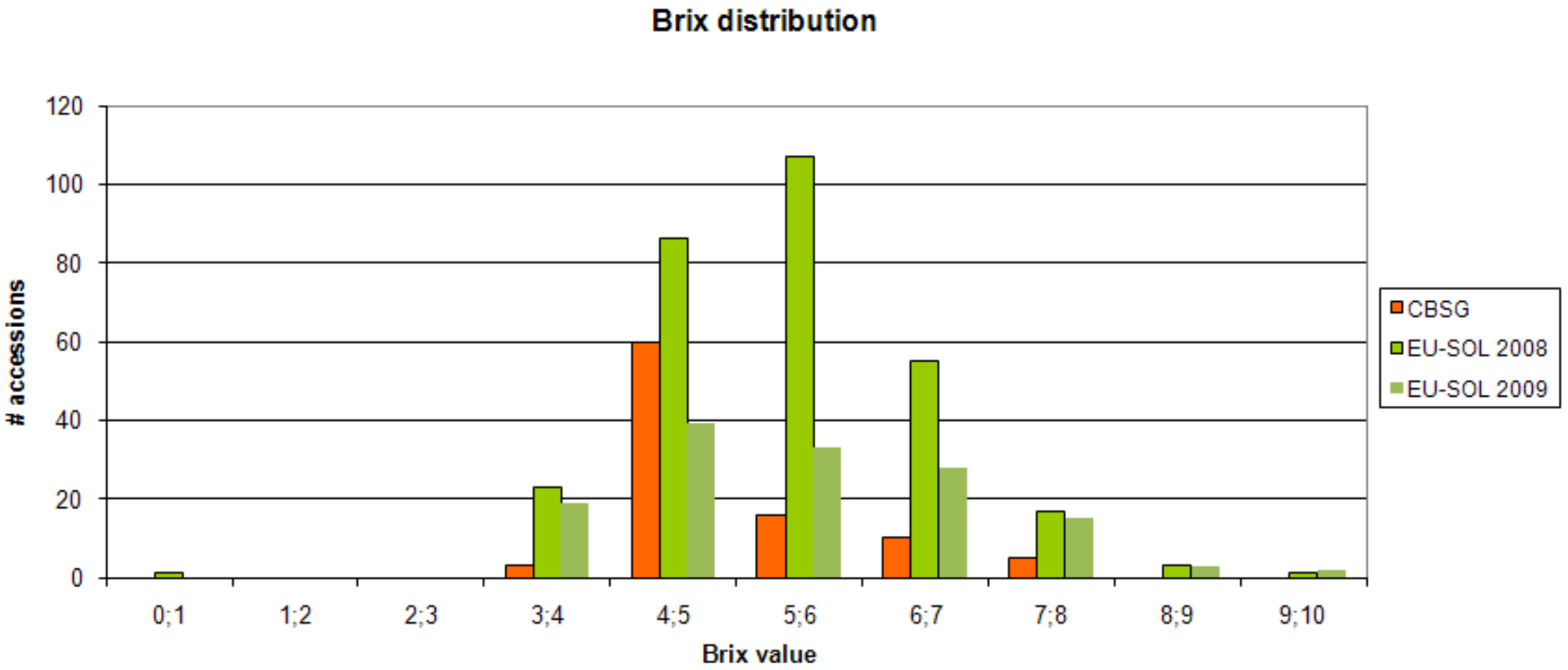


Figure 5



Brix data 2009 vs Brix data 2008

Figure 6



Brix value	CBSG	E-SOL 2008	EU-SOL 2009 (Avg.)
0;1	0	1	0
1;2	0	0	0
2;3	0	0	0
3;4	3	23	19
4;5	60	86	39
5;6	16	107	33
6;7	10	55	28
7;8	5	17	15
8;9	0	3	3
9;10	0	1	2
#individuals	94	293	139

Figure 7

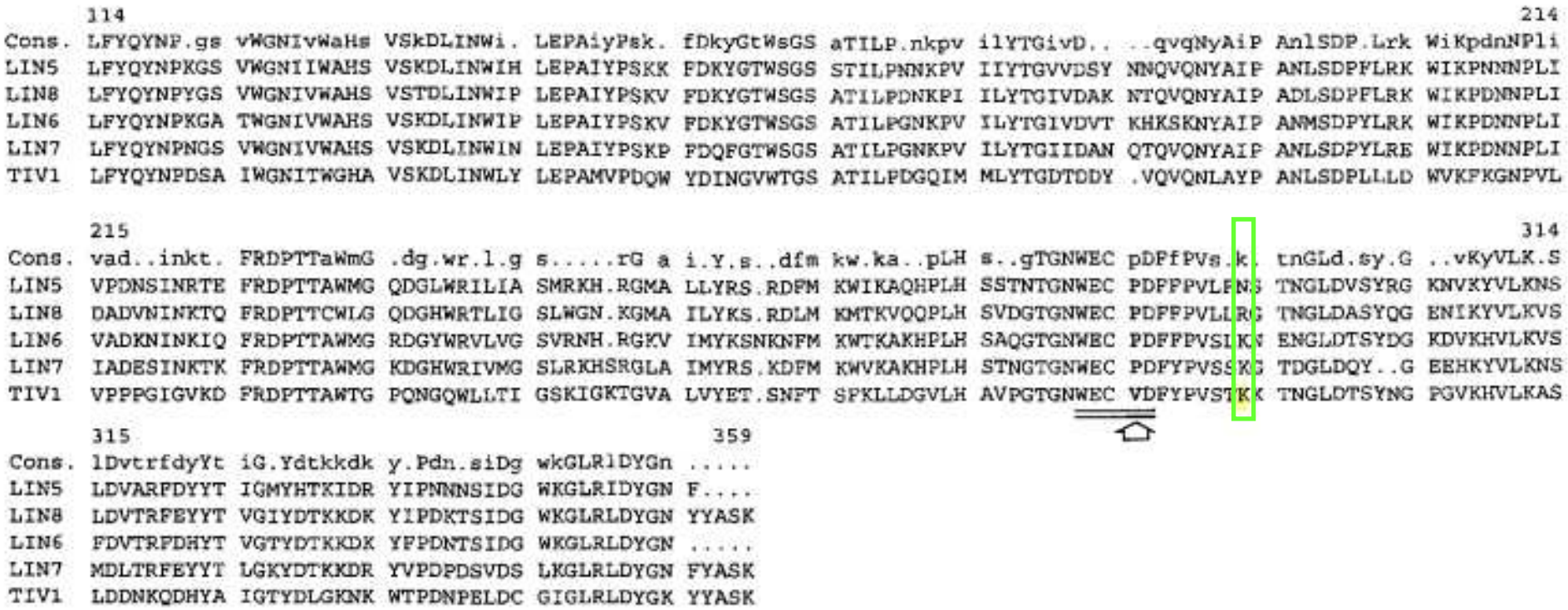

114 214
Cons. LFYQYNP.gs vWGNIvWaHs VSkDLINWi. LEPAiyPsk. fDkyGtWsGS aTILP.nkpv ilYTGivD.. ..qvqNyAiP AnlSDP.Lrk WiKpdnNPli
LIN5 LFYQYNPKGS VWGNIIWAHS VSKDLINWIH LEPAIYPSKK FDKYGTWSGS STILPNNKPV IIYTGVVDSY NNQVQNYAIP ANLSDPFLRK WIKPNNNPLI
LIN8 LFYQYNPYGS VWGNIVWAHS VSTDLINWIP LEPAIYPSKV FDKYGTWSGS ATILPDNKPI ILYTGIVDAK NTQVQNYAIP ADLSDPFLRK WIKPDNNPLI
LIN6 LFYQYNPKGA TWGNIVWAHS VSKDLINWIP LEPAIYPSKV FDKYGTWSGS ATILPGNKPV ILYTGIVDVT KHKSKNYAIP ANMSDPYLRK WIKPDNNPLI
LIN7 LFYQYNPNGS VWGNIVWAHS VSKDLINWIN LEPAIYPSKP FDQFGTWSGS ATILPGNKPV ILYTGIIDAN QTQVQNYAIP ANLSDPYLRE WIKPDNNPLI
TIV1 LFYQYNPDSA IWGNITWGHA VSKDLINWLY LEPAMVPDQW YDINGVWTGS ATILPDGQIM MLYTGDTDDY .VQVQNLAYP ANLSDPLLLD WVKFKGNPVL
215 314
Cons. vad..inkt. FRDPTTaWmG .dg.wr.l.g s.....rG a i.Y.s..dfm kw.ka..pLH s..gTGNWEC pDPfPVs.k. tnGLd.sy.G ..vKyVLK.S
LIN5 VPDNSINRTE FRDPTTAWMG QDGLWRILIA SMRKH.RGMA LLYRS.RDFM KWIKAQHPLH SSTNTGNWEC PDPFPVLPNG TNGLDVSYRG KNVKYVLKNS
LIN8 DADVNINKTQ FRDPTTCWLG QDGHWRTLIG SLWGN.KGMA ILYKS.RDLM KMTKVQQPLH SVDGTGNWEC PDFFPVLLRG TNGLDASYQG ENIKYVLKVS
LIN6 VADKNINKIQ FRDPTTAWMG RDGYWRVLVG SVRNH.RGKV IMYKSNKNFM KWTKAKHPLH SAQGTGNWEC PDFFPVSIKN ENGLDTSYDG KDVKHVLKVS
LIN7 IADESINKTK FRDPTTAWMG KDGHWRIVMG SLRKHSRGLA IMYRS.KDFM KWVKAKHPLH STNGTGNWEC PDFYPVSSKG TDGLDQY..G EEHKYVLKNS
TIV1 VPPPGIGVKD FRDPTTAWTG PQNGQWLLTI GSKIGKTGVA LVYET.SNFT SPKLLDGVLH AVPGTGNWEC VDFYPVSTKK TNGLDTSYNG PGVKHVLKAS
315 359
Cons. lDvtrfdyYt iG.Ydtkkdk y.Pdn.siDg wkGLRlDYGn .....
LIN5 LDVARFDYYT IGMYHTKIDR YIPNNNSIDG WKGLRIDYGN F....
LIN8 LDVTRFEYYT VGIYDTKKDK YIPDKTSIDG WKGLRLDYGN YYASK
LIN6 FDVTRFDHYT VGTYDTKKDK YFPDNTSIDG WKGLRLDYGN .....
LIN7 MDLTRFEYYT LGKYDTKKDR YVPDPDSVDS LKGLRLDYGN FYASK
TIV1 LDDNKQDHYA IGTYDLGKNK WTPDNPELDC GIGLRLDYGK YYASK

Figure 8

TABLES

Table 1 a)

	1st measurement			2nd measurement					Averages (meas. 1+2)	
CC#	09#-1	avg FW-1	Brix 09-1	09 #-2	CC#	Brix 09-2	avg FW-2	Origin	Avg Brix	Tot. Avg Fw (g)
41	2721	0.011	5.5							0.005
79	2722	0.040	4.1	2722	79	3.6	0.048	brix panel subset	3.85	0.044
205	2862	0.361	4.8	2862	205	-		brix panel subset		0.181
304	2864	0.202	5.8	2864	304	4.9	0.167	brix panel subset	5.35	0.185
369	2723	0.208	5.8	2723	369	5.9	0.177	brix panel subset	5.85	0.192
376	2725	0.030	7.5	2725	376	7.9	0.029	brix panel subset	7.7	0.029
389	2726	0.101	6.9	2726	389	8.6	0.103	brix panel subset	7.75	0.102
422	2868	0.064	5.7	2868	422	6.2	0.019	brix panel subset	5.95	0.042
448	2869	0.134	5.5	2869	448	6.4	0.184	brix panel subset	5.95	0.159
512	2871	0.170	3.8	2871	512	3.3	0.144	brix panel subset	3.55	0.157
610	2873	0.016	3.9	2873	610	4.3	0.018	brix panel subset	4.1	0.017
744	2877	0.015	4	2877	744	5	0.012	brix panel subset	4.5	0.013
891	2727	0.014	4.9	2727	891	4.8	0.013	brix panel subset	4.85	0.014
892	2883	0.322	4.7	2883	892	4.9	0.347	brix panel subset	4.8	0.335
915	2884	0.021	9.6	2884	915	8.5	0.009	brix panel subset	9.05	0.015
933	2887	0.006	7.1	2887	933	6.8	0.004	brix panel subset	6.95	0.005
948	2889	0.008	8	2889	948	6.6	0.008	brix panel subset	7.3	0.008
951	2890	0.065	6.7	2890	951	5.6	0.027	brix panel subset	6.15	0.046
983	2891	0.121	5.6	2891	983	5.8	0.108	brix panel subset	5.7	0.114
1002	2892	0.058	5.7	2892	1002	5.4	0.044	brix panel subset	5.55	0.051
1020	2893	0.240	6.3	2893	1020	5.3	0.152	brix panel subset	5.8	0.196
1027	2894	0.048	4.5	2894	1027	5.1	0.091	brix panel subset	4.8	0.070
1040	2896	0.129	6.6	2896	1040	-		brix panel subset	-	0.065
1049	2897	0.287	4.6	2897	1049	4	0.197	brix panel subset	4.3	0.242
1117	2728	0.056	6.2	2728	1117	5.6	0.054	brix panel subset	5.9	0.055
1155	2903	0.362	45	2903	1155	4.2	0.184	brix panel subset	24.6	0.273
1159	2729	0.058	4.2	2729	1159	6.1	0.083	brix panel subset	5.15	0.071
1185	2905	0.038	6.5	2905	1185	5.8	0.029	brix panel subset	6.15	0.033
1193	2906	0.100	3.1	2906	1193	-		brix panel subset	-	0.050
1198	2907	0.154	5.9	2907	1198	5.2	0.151	brix panel subset	5.55	0.152
1230	2909	0.059	6.2	2909	1230	6.1	0.026	brix panel subset	6.15	0.042
1234	2730	0.063	4.4	2730	1234	5.1	0.053	brix panel subset	4.75	0.058
1237	2732	0.108	3.4	2732	1237	3.9	0.120	brix panel subset	3.65	0.114
1270	2911	0.055	5.7	2911	1270	5.7	0.036	brix panel subset	5.7	0.045
1356	2914	0.007	7.2	2914	1356	7.4	0.007	brix panel subset	7.3	0.007
1371	2915	0.237	5.5	2915	1371	5.7	0.184	brix panel subset	5.6	0.210
1582	2736	0.060	6.8	2736	1582	5.8	0.074	brix panel subset	6.3	0.067
1684	2921	0.130	4.3	2921	1684	3.6	0.118	brix panel subset	3.95	0.124
1712	2923	0.171	4.8	2923	1712	4.6	0.110	brix panel subset	4.7	0.141
1756	2928	0.034	3.9	2928	1756	4	0.045	brix panel subset	3.95	0.039
1784	2930	0.048	4	2930	1784	4.2	0.046	brix panel subset	4.1	0.047
1802	2931	0.014	3.7	2931	1802	4.9	0.015	brix panel subset	4.3	0.015
1804	2738	0.033	4.5	2738	1804	4.4	0.030	brix panel subset	4.45	0.032
1820	2739	0.031	4.5	2739	1820	5.1	0.040	brix panel subset	4.8	0.035
1903	2933	0.060	3.7	2933	1903	3.3	0.093	brix panel subset	3.5	0.076
1915	2934	0.078	3.2	2934	1915	3	0.066	brix panel subset	3.1	0.072
1953	2936	0.006	3.8	2936	1953	6.4	0.006	brix panel subset	5.1	0.006
1960	2938	0.079	3.7	2938	1960	3.7	0.061	brix panel subset	3.7	0.070
1966	2940	0.086	5.7	2940	1966	-		brix panel subset		
1982	2944	0.073	3.5	2944	1982	3.4	0.063	brix panel subset	3.45	0.068
1989	2946	0.106	4.1	2946	1989	4.6	0.091	brix panel subset	4.35	0.098
2139	2744	0.061	6.1	2744	2139	6.3	0.073	brix panel subset	6.2	0.067
2146	2745	0.038	6.6	2745	2146	6	0.038	brix panel subset	6.3	0.038
2304	2952	0.038	4.8	2952	2304	4.6	0.028	brix panel subset	4.7	0.033
2431	2956	0.003	4.6	2956	2431	7.5	0.003	brix panel subset	6.05	0.003
2435	2957	0.032	4.1	2957	2435	4.7	0.030	brix panel subset	4.4	0.031
2586	2968	0.156	5.8	2968	2586	4.7	0.049	brix panel subset	5.25	0.103
2660	2971	0.013	6.8	2971	2660	7.7	0.014	brix panel subset	7.25	0.013
2669	2972	0.033	3.4	2972	2669	3.7	0.037	brix panel subset	3.55	0.035
2696	2973			2978	2761	6.5	0.028	brix panel subset	3.25	0.014
2723	2752	0.031	7.1	2752	2723	7	0.025	brix panel subset	7.05	0.028
2728	2753	0.057	4.8	2753	2728	5.1	0.050	brix panel subset	4.95	0.054
2732	2754	0.076	6.6	2754	2732	5.5	0.064	brix panel subset	6.05	0.070
2753	2755	0.032	3.6	2755	2753	3.6	0.035	brix panel subset	3.6	0.034
2764	2756	0.066	6.4	2756	2764	5.4	0.033	brix panel subset	5.9	0.050
2879	2985	0.075	3.8	2985	2879	3.8	0.058	brix panel subset	3.8	0.067
2883	2986	0.003	4.1	2986	2883	4.2	0.003	brix panel subset	4.15	0.003
2890	2760	0.028	4.9	2760	2890	-	-	brix panel subset		
2895	2761	0.039	6.4	2761	2895	5.4	0.036	brix panel subset	5.9	0.037
2898	2762	0.036	6.9	2762	2898	5.9	0.023	brix panel subset	6.4	0.029
2953	2988	0.008	7.5	2988	2953	6.2	0.004	brix panel subset	6.85	0.006
2959	2989	0.012	6.7	2989	2959	6.8	0.013	brix panel subset	6.75	0.012
2979	2765	0.006	7.3	2765	2979	7.8	0.006	brix panel subset	7.55	0.006
2994	2994	0.001	6.9	2994	2994	9.1	0.001	brix panel subset	8	0.001
3002	2767	0.052	4.5	2767	3002	4.4	0.031	brix panel subset	4.45	0.042
3026	2770	0.052	4	2770	3026	4.8	0.039	brix panel subset	4.4	0.046
3028	2771	0.032	4.3	2771	3028	6.1	0.041	brix panel subset	5.2	0.036
3051	2996	0.005	6.9	2996	3051	5	0.068	brix panel subset	5.95	0.037

Table 1 b)

	1st measurement			2nd measurement					Averages (meas. 1+2)	
CC#	09#-1	avg FW-1	Brix 09-1	09 #-2	CC#	Brix 09-2	avg FW-2	Origin	Avg Brix	Tot. Avg Fw (g)
3053	2773	0.042	4.4	2773	3053	4.6	0.040	brix panel subset	4.5	0.041
3083	2775	0.060	6.4	2775	3083	6.4	0.039	brix panel subset	6.4	0.050
3099	2776	0.066	6.3	2776	3099	5.9	0.063	brix panel subset	6.1	0.065
3126	2777	0.051	6.6	2777	3126	5.9	0.062	brix panel subset	6.25	0.056
3150	2780	0.123	5.7	2780	3150	6.4	0.103	brix panel subset	6.05	0.113
3159	2782	0.081	7.1	2782	3159	6.2	0.050	brix panel subset	6.65	0.065
3174	2783	0.076	6.2	2783	3174	6.2	0.084	brix panel subset	6.2	0.080
3190	2784	0.055	4.3	2784	3190	4.6	0.059	brix panel subset	4.45	0.057
3206	2786	0.099	6.2	2786	3206	5.8	0.061	brix panel subset	6	0.080
3223	2788	0.060	3.7	2788	3223	3.7	0.055	brix panel subset	3.7	0.057
3244	2789	0.073	4.4	2789	3244	3.9	0.060	brix panel subset	4.15	0.067
3271	3001	0.003	8	3001	3271	8.7	0.002	brix panel subset	8.35	0.003
3274	2791	0.194	5.5	2791	3274	4.3	0.064	brix panel subset	4.9	0.129
3306	3004	0.003	6.9	3004	3306	9.9	0.002	brix panel subset	8.4	0.002
3327	3006	0.006	6.6	3006	3327	-		brix panel subset		
3426	2794	0.145	3.6	2794	3426	4.1	0.047	brix panel subset	3.85	0.096
3439	2795	0.069	6.6	2795	3439	6.3	0.069	brix panel subset	6.45	0.069
3456	2796	0.067	4.8	2796	3456	4.1	0.070	brix panel subset	4.45	0.068
3463	3014	0.108	3.8	3014	3463	3.6	0.108	brix panel subset	3.7	0.108
3496	3015	0.034	4.8	3015	3496	-		brix panel subset		
3525	2800	0.013	5.4	2800	3525	5	0.016	brix panel subset	5.2	0.014
3527	2801	0.058	4	2801	3527	3.8	0.048	brix panel subset	3.9	0.053
3533	2802	0.044	7.4	2802	3533	6	0.048	brix panel subset	6.7	0.046
3539	3016	0.003	8.3	3016	3539	9.3	0.003	brix panel subset	8.8	0.003
3567	2803	0.033	3.6	2803	3567	-		brix panel subset		0.016
3577	3018	0.104	4	3018	3577	5.5	0.171	brix panel subset	4.75	0.137
3586	2807	0.049	4.8	2807	3586	3	0.049	brix panel subset	3.9	0.049
3611	2808	0.027	7.4	2808	3611	7.2	0.028	brix panel subset	7.3	0.028
3613	2809	0.050	4	2809	3613	4.6	0.038	brix panel subset	4.3	0.044
3633	2811	0.079	4.3	2811	3633	5.1	0.065	brix panel subset	4.7	0.072
3648	3020	0.147	3.9	3020	3648	4.3	0.174	brix panel subset	4.1	0.160
3650	3021	0.132	5.4	3021	3650	4.2	0.053	brix panel subset	4.8	0.092
3657	2812	0.042	6.4	2812	3657	6.3	0.033	brix panel subset	6.35	0.038
3659	2813	0.055	6	2813	3659	5.8	0.040	brix panel subset	5.9	0.047
3669	2814	0.067	5.6	2814	3669	5.3	0.075	brix panel subset	5.45	0.071
3673	2815	0.062	3.7	2815	3673	8.2	0.020	brix panel subset	5.95	0.041
3981	2820	0.059	3.6	2820	3981	4.4	0.044	brix panel subset	4	0.052
4001	3023	0.159	4,7	3023	4001	4.7	0.071	brix panel subset	4.7	0.115
4032	3024	0.025	4.2	3024	4032	-		brix panel subset		0.013
4186	2822	0.081	4.9	2822	4186	5.5	0.056	brix panel subset	5.2	0.068
4228	3034	0.084	5.7	3034	4228	7.7	0.007	brix panel subset	6.7	0.045
4236	2823	0.055	7.9	2823	4236	7.2	0.024	brix panel subset	7.55	0.040
4243	3035	0.048	6.6	3035	4243	5.8	0.021	brix panel subset	6.2	0.034
4877	3040	0.046	6.4	3040	4877	7.7	0.002	brix panel subset	7.05	0.024
5170	3046	0.002	4.5	3046	5170	6.8	0.002	brix panel subset	5.65	0.002
5228	3049	0.027	4.6	3049	5228	6	0.026	brix panel subset	5.3	0.027
5480	3050	0.050	4	3050	5480	5.8	0.042	brix panel subset	4.9	0.046
5531	2827	0.119	4.2	2827	5531	-		brix panel subset	-	
5550	2829	0.039	6.9	2829	5550	7.5	0.031	brix panel subset	7.2	0.035
5575	2830	0.045	4.6	2830	5575	5.1	0.026	brix panel subset	4.85	0.035
5578	2831	0.158	5.1	2831	5578	4.6	0.176	brix panel subset	4.85	0.167
5581	3053	0.103	6.1	3053	5581	-		brix panel subset		
5612	3054	0.054	7.2	3054	5612	5.6	0.093	brix panel subset	6.4	0.073
5653	2833	0.037	7.2	2833	5653	6.7	0.029	brix panel subset	6.95	0.033
5703	3057	0.006	3.9	3057	5703	7.5	0.002	brix panel subset	5.7	0.004
5747	3059	0.107	5.2	3059	5747	5.2	0.106	brix panel subset	5.2	0.107
5774	3060	0.035	4.6	3060	5774	4.4	0.035	brix panel subset	4.5	0.035
5775	2835	0.094	4.5	2835	5775	4.8	0.117	brix panel subset	4.65	0.105
5801	2840	0.056	5.6	2840	5801	5.7	0.068	brix panel subset	5.65	0.062
5808	2842	0.052	7.6	2842	5808	7.5	0.037	brix panel subset	7.55	0.045
6161	2845	0.071	4.7	2845	6161	4.7	0.058	brix panel subset	4.7	0.064
6244	2846	0.020	7.5	2846	6244	7.4	0.015	brix panel subset	7.45	0.018
6345	2848	0.128	5.3	2848	6345	6.4	0.043	brix panel subset	5.85	0.085
6408	2849	0.115	6	2849	6408	-		brix panel subset		0.057
6434	3064	0.031	4	3064	6434	4.4	0.051	brix panel subset	4.2	0.041
6485	3066	0.165	5.7	3066	6485	5.3	0.061	brix panel subset	5.5	0.113
6660	3069	0.005	7.3	3069	6660	7.5	0.004	brix panel subset	7.4	0.005
6689	3074	0.132	5.4	3074	6689	4.8	0.086	brix panel subset	5.1	0.109
6690	3075	0.052	6.4	3075	6690	5.1	0.047	brix panel subset	5.75	0.050
6728	3079	0.004	7.4	3079	6728	8.2	0.005	brix panel subset	7.8	0.005
6756	2852	0.113	4.7	2852	6756	4.8	0.158	brix panel subset	4.75	0.136
6824	2856	0.059	5.9	2856	6824	6	0.052	brix panel subset	5.95	0.056
6843	2857	0.170	6.3	2857	6843	5.8	0.132	brix panel subset	6.05	0.151
6878	2858	0.193	6.1	2858	6878	6.3	0.174	brix panel subset	6.2	0.184
6902	3081	0.048	3.4	3081	6902	3.4	0.040	brix panel subset	3.4	0.044

	# Fragments screened	Amplicon size (bp)	# Individuals	Total screened sequence (bp)	Polymorphism (Exons and Introns)	Polymorphism (Exons only)	Non synonymous	Total screened sequence/Total polymorphism frequency (bp)	Total screened sequence/SNP in exon (bp)
EcoTILLING with Licor on 89 individuals (*Mosca et al*, 2009)	6 fragments	1000 (min) 1500 (max)	89	534000 801000	144	30	8	3708* 5562*	66750* 100125*
	Average*			667500*				4635*	**22250***
EcoTILLING with Light Scanner on 370 individuals	4 fragments Invertases (lin6, lin7, lin8, SI)	400 508 526 378	352	637824	188	26	1	3393	**24532**

*: inferred values

Table 2

Gene code	NCBI accession	Name	Chromosome location	Fragment size (AA)	Fragment size (BP)	ConsDom start (AA)	ConsDom stop (AA)	size of ConsDom (AA)	ConsDom start (bp) on cDNA	ConsDom stop (bp) on cDNA	size of ConsDom (bp) on cDNA	Arabidopsis Homologues CODE	PROSITE code	CD name	Pattern	Sequence
APL1	U81033	ADP-glucose pyrophosphorylase large subunit (AGP-S1)	(1-4)	516	1548	91	110	19	273	330	57	AT4G39210	PS00808	ADP_GLC_PYROPHOSPH_	[AG]-G-G-x-G-[STKA]-x-L-x(2)-L-[TA]-x(3)-[AST]-x-P-[AS]-[LV].	GGGeGTkLfpLTsrtAtPAV
						181	189	8	543	567	24		PS00809	ADP_GLC_PYROPHOSPH_	W-[FY]-x-G-[ST]-[AS]-[DNSH]-[AS]-[LIVMFYW].	WFqGTADAV
						295	305	10	885	915	30		PS00810	ADP_GLC_PYROPHOSPH_	APV]-[GS]-M-G-[LIVMN]-Y-[IVC]-[LIVMFY]-x(2)-[DENPHKRQS].	ASMGVYVFktD
APL2	U81034	ADP-glucose pyrophosphorylase large subunit (AGP-S2)	(7-4)	518	1554	91	110	19	273	330	57	AT1G27680	PS00808	ADP_GLC_PYROPHOSPH_	[AG]-G-G-x-G-[STKA]-x-L-x(2)-L-[TA]-x(3)-[AST]-x-P-[AS]-[LV].	GGGvGTrLfpLTsrrAkPAV
						178	189	11	534	567	33		PS00809	ADP_GLC_PYROPHOSPH_	W-[FY]-x-G-[ST]-[AS]-[DNSH]-[AS]-[LIVMFYW].	WFqGTADAV
						293	305	12	879	915	36		PS00810	ADP_GLC_PYROPHOSPH_	APV]-[GS]-M-G-[LIVMN]-Y-[IVC]-[LIVMFY]-x(2)-[DENPHKRQS].	ASMGVYVFktD
APL3	U85497	ADP-glucose pyrophosphorylase large subunit (agpL3)	1-1/1-2	516	1548	91	110	19	273	330	57	AT5G19220.1	PS00808	ADP_GLC_PYROPHOSPH_	[AG]-G-G-x-G-[STKA]-x-L-x(2)-L-[TA]-x(3)-[AST]-x-P-[AS]-[LV].	GGGgGTrLfpLTkrrAkPAV
						181	189	8	543	567	24		PS00809	ADP_GLC_PYROPHOSPH_	W-[FY]-x-G-[ST]-[AS]-[DNSH]-[AS]-[LIVMFYW].	WFqGTADAV
						295	305	10	885	915	30		PS00810	ADP_GLC_PYROPHOSPH_	APV]-[GS]-M-G-[LIVMN]-Y-[IVC]-[LIVMFY]-x(2)-[DENPHKRQS].	ASMGVYVFkkD
SS	L41126	ADP-glucose pyrophosphorylase small subunit	7-2/7-3	521	1563	96	110	14	288	330	42	AT5G48300.1	PS00808	ADP_GLC_PYROPHOSPH_	[AG]-G-G-x-G-[STKA]-x-L-x(2)-L-[TA]-x(3)-[AST]-x-P-[AS]-[LV].	GGGaGTrLypLTkkrAkPAV
						185	189	4	555	567	12		PS00809	ADP_GLC_PYROPHOSPH_	W-[FY]-x-G-[ST]-[AS]-[DNSH]-[AS]-[LIVMFYW].	WFqGTADAV
						297	305	8	891	915	24		PS00810	ADP_GLC_PYROPHOSPH_	APV]-[GS]-M-G-[LIVMN]-Y-[IVC]-[LIVMFY]-x(2)-[DENPHKRQS].	ASMGIYVIskD
Fsk 1	U64817 LEU64817	Fructose kinase 1	(4-4)			58	82	24	174	246	72		PS00583	PFKB_KINASES_1	[AG]-G-x(0,1)-[GAP]-x-N-[AGLS]-[STA]-x(2)-[AI]-x-[G]-[GNKA]-[GS]-x(9)-G	GGaPaNVAvcIsKLGgssafigkvG
						279	292	13	837	876	39		PS00584	PFKB_KINASES_2	[DNSK]-[PSTV]-x-[SAG](2)-[GD]-D-x(3)-[SAGV]-[AG]-[LIVMFYA]-[LIVMS	DTtGAGDaftGGVL
Fsk 2	U64818 LEU64818	Fructose kinase 2	6	328	984	42	66	24	126	198	72	AT3G59480.1	PS00583	PFKB_KINASES_1	[AG]-G-x(0,1)-[GAP]-x-N-[AGLS]-[STA]-x(2)-[AI]-x-[G]-[GNKA]-[GS]-x(9)-G	GGaPaNVAiaVtRLGgksafvgkIG
						261	274	13	783	822	39		PS00584	PFKB_KINASES_2	[DNSK]-[PSTV]-x-[SAG](2)-[GD]-D-x(3)-[SAGV]-[AG]-[LIVMFYA]-[LIVMS	DTtGAGDsfvGALL
Fxk3	AY323226.1	Fructose kinase 3		386	1158	98	122	24	294	366	72					
						318	331	13	954	993	39					
Fsk4	AY099454	Fructose kinase 4	(4-4)	375	1125	85	109	24	255	327	72	AT5G51830.1				
						306	319	13	918	957	39					
Hxk 1	AJ401153	Hexokinase 1		498	1494	172	197	25	516	591	75	AT4G29130.1	PS00378	HEXOKINASES	[LIVM]-G-F-[TN]-F-S-[FY]-P-x(5)-[LIVM]-[DNST]-x(3)-[LIVM]-x(2)-W-T-K-x-	LGFTFSFPvmqtsINsgnImrWTKgF
Hxk 2	AF208543	Hexokinase 2		496	1488	172	197	25	516	591	75	AT2G19860.1	PS00378	HEXOKINASES	[LIVM]-G-F-[TN]-F-S-[FY]-P-x(5)-[LIVM]-[DNST]-x(3)-[LIVM]-x(2)-W-T-K-x-	LGFTFSFPimqtsINsgtLirWTKgF
Hxk3	DQ05686*1*	Hexokinase 3	12	499	1497							AT2G19860		HEXOKINASES	[LIVM]-G-F-[TN]-F-S-[FY]-P-x(5)-[LIVM]-[DNST]-x(3)-[LIVM]-x(2)-W-T-K-x-	[LF].
Hxk 4	DQ056862	Hexokinase 4	(4-4)	498	1494							AT1G47840.1	PS00378	HEXOKINASES	[LIVM]-G-F-[TN]-F-S-[FY]-P-x(5)-[LIVM]-[DNST]-x(3)-[LIVM]-x(2)-W-T-K-x-	[LF].
Icl	U18678	Isocitrate lyase		575	1725	211	216	5	633	648	15	AT3G21720.1		ISOCITRATE_LYASE	K-[KR]-C-G-H-[LMQR].	KKCGHM
Lin5	AJ272304	Invertase	9;2	584	1752	57	70	13	171	210	39	AT2G36190	PS00609		H-x(2)-[PY]-x(4)-[LIVMA]-N-D-P-N-[GA].	HfqPpkhwfNDPNA
Lin6	AF506005	Invertase	10	582	1746	56	69	13	168	207	39	AT3G52600.1	PS00609	GLYCOSYL_HYDROL_F32	H-x(2)-[PY]-x(4)-[LIVMA]-N-D-P-N-[GA].	HfqPpknwfNDPNG
Lin7	AF506006	Invertase	9	584	1752	57	70	13	171	210	39					
Lin8	AF506007	Invertase	10	607	1821	7	20	13	21	60	39	AT3G52600.1				
Mls	AT5G03860.1	Malate synthase		497	1491	215	230	15	645	690	45	AT5G03860.1	PS00510	MALATE_SYNTHASE	[KR]-[DENQ]-[HN]-x(2)-G-L-N-x-G-x-W-D-Y-[LIVM]-F	RDHsvGLNcGrWDYIF
Susy2	AJ011535	Sucrose synthase 2	7-5 ; 7-4	805	2415							AT3G43190.1			NO conserved domain displayed	
Susy3	AJ011319	Sucrose synthase 3	12	805	2415							AT5G20830.1			NO conserved domain displayed	
Tiv1/SI	M81081	Sucrose invertase	3	636	1908	108	121	13	324	363	39	AT1G62660.1	PS00609	GLYCOSYL_HYDROL_F32	H-x(2)-[PY]-x(4)-[LIVMA]-N-D-P-N-[GA].	HfqPqknwMNDPNG
UGPase	AW217311 EST	UDP-glucose pyrophosphorylase														
PGI =GPI	DQ451687	GPI glucose-6-phosphate isomerase	12? (penn-neo	568	1704	269	282	13	807	846	39	AT5G42740.1	PS00174	P_GLUCOSE_ISOMERASE_	[GSA]-x-[LIVMCAYQS]-[LIVMFYWN]-x(4)-[FY]-[DNTH]-Q-x-[GA]-[IV]-[EQ	DwVGGRFSVcSAVG

Table 3

Primer name	sequence (in 5'----> 3' order)	Reverse complement	DNA type	Tm	Gc	Start	Stop	Length	Exp. Size (bp)	Obs. size (from gel 1Kb+ ladder)	Introns
Apl1_cd1_2_bp_R	GTGCTGACACAGTACAATTCTGCTG	CAGCAAAGAATTGTACTGTGTCAGCAC	cDNA	60	48	406	430	25		500	
Apl1_cd1_2_bp_R BIS	AGGATGCTACAGGCTGATAGACATC	GATGTCTATCAGCCTGTAGCATCCT	cDNA	60	48	339	363	25	105		
Apl1_cd2_1_ bp_F	GCTGCAACTCAGACACCTGG		cDNA	59	60	505	524	20			
Apl1_cd2_1_bp_R	TAGGATGGATTATATGGAGTTGGTGC	GCACCAACTCCATATAATCCATCCTA	cDNA	58	42	648	573	26	68	300	1 intron
Apl1_cd2_2_bp_F			cDNA								
Apl1_cd2_2_bp_R	CTTTCATGTGCACCAGCTGAGG	CCTCAGCTGGTGCACATGAAAG	cDNA	60	55	706	727	22	727	250	
Apl1_cd3_1_bp_F	GGATTATCTCCACAAGATGCGAAG		cDNA	57	44	844	867	24			
Apl1_cd3_1_bp_R	ATCTTTCTATAATGCTAGCTTGGCG	CGCCAAGCTAGCATTATAGAAAGAT	cDNA	57	40	1053	1077	25	233	350	
Apl1_cd3_2_bp_F	GATTGACAGCAGAGGCAGAGTTGT		cDNA	61	50	759	782	24			
Apl1_cd3_2_bp_R	TTTGGCTCTGAAATTATACCAGCAGA	GCTGCTGGTATAATTTCAGAGCCAAA	cDNA	60	42	961	986	26	227	350	
Apl2_cd2_1_ bp_R	TTGTCTGGTGACCATCTTTACCG	CGGTAAAGATGGTCACCAGACAA	cDNA	59	46	630	653	23		500	2 introns
Apl2_cd3_1_ bp_F	GACACATCTATCCTCGGACTATCTGA		cDNA	59	46	835	860	26			
Apl2_cd3_1_bp_R	ATTGGAACAGTGAAGTCTTTCTTTGATGC	GCATCAAAGAAAGACTTCACTGTTCCAAT	cDNA	60	38	1045	1073	29	238	350	0 intron
Apl2_cd3_2_ bp_F	GCCCTGCTTTGAAGGCA		cDNA	57	59	809	825	17			
Apl2_cd3_2_ bp_R	TCTGAAATTATCCCTTCTGCTGTG	CACAGCAGAAGGGATAATTTCAGA	cDNA	57	42	973	996	24	187	350	1 intron
Apl3_cd1_1_ bp_F	GAACAGTAGTAGCAATCATTCTAGGA		cDNA	56	38	239	264	26			
Fk2_cd2_3_ bp_F	CCGTTGGAGGATTCCATGTGAAGA		cDNA	60	50	752	775	24			
Fk2_cd2_3_ bp_R	CCTCACTTTGCTCAAGGGAGG	CCTCCCTTGAGCAAAGTGAGG	cDNA	59	57	960	980	21	228	250	
Icl_cd1_1BIS_bp_F	GCAAGAATGAGCATGTGCAGAGA		cDNA	59	48	497	519	23			
Icl_cd1_1bp_R	ATCAGTCCTCTGTTACTAAGAAATGTGG	CCACATTTCTTAGTAACAGAGGACTGAT	cDNA	58	39	660	687	28	190	200	in combination with Se11_1_1 bisF
Icl_cd1_2_bp_F	CTGTTAAGCTGTGTAAGCTTTTCGTG		cDNA	58				26			0 introns
Icl_cd1_3_bp_R	CCAATGTGGATACAAGGGATCATCAG	CTGATGATCCCTTGTATCCACATTGG	cDNA	59	46	828	853	26		350	0 introns
Mls_cd1_3_bp_F	GGAACAATGTGTTTGACAGGGCA		cDNA	60	48	527	549	23			
Mls_cd1_3_bp_R	TTGCTTATCCATACCTGTCATAAGCG	CGCTTATGACAGGTATGGATAAGCAA	cDNA	59	42	793	818	26	291	200	check!
Lin5-CD1-1_F	GAATTATTTATGAAAAACTCTTCTCT		gDNA	50	23	9815	9840	26			
Lin5-CD1-1_R	ACATTTATTGACTTGGTATATATATCAGTACG	CGTACTGATATATATACCAAGTCAATAAATGT	gDNA	55	28	10162	10193	32	378	400	
Lin5-CD1-2_F	CAAACATTAATAGGGCATTTGCTTC		gDNA	55	25	9885	9909	25			
Lin5-CD1-2_R	TATATAGAGAAGTTCAAATGTTGAGG	CCTCAACATTTGAACTTCTCTATATA	gDNA	52	31	10272	10297	26	412	450	
Lin5-CD1-3_R	TATATAGAGAAGTTCAAATGTTGAGGGTTC	GAACCCTCAACATTTGAACTTCTCTATATA	gDNA	57	32	10272	10302	31	417	450	
Lin6-CD1-1_F	GTTTACTCCTATGTACTTTTTGCTCAACA		gDNA	58	34	11905	11933	29			
Lin6-CD1-1_R	ATGATTTAGACGTTCTGAGTCGAGC	GCTCGACTCAGAACGTCTAAATCAT	gDNA	58	44	12308	12332	25	427	450	
Lin6-CD1-2_F	GCCAATTCTTGTGTTGTGTTTCTTTATCA		gDNA	59	34	12115	12143	29			

Table 4 a)

Primer name	sequence (in 5'----> 3' order)	Reverse complement	DNA type	Tm	Gc	Start	Stop	Length	Exp. Size (bp)	Obs. size (from gel 1Kb+ ladder)	Introns
Lin6-CD1-1_F	GTTTACTCCTATGTACTTTTTGCTCAACA		gDNA	58	34	11905	11933	29			
Lin6-CD1-1_R	ATGATTTAGACGTTCTGAGTCGAGC	GCTCGACTCAGAACGTCTAAATCAT	gDNA	58	44	12308	12332	25	427	450	
Lin6-CD1-2_F	GCCAATTCTTGTGTTGTGTTTCTTTATCA		gDNA	59	34	12115	12143	29			
Lin6-CD1-2_R	ATGCAACAGTTTTAATTTGATTAGTGCATATG	CATATGCACTAATCAAATTAAAACTGTTGCAT	gDNA	58	28	12483	12514	32	399	450	
Lin7-CD1-1_F	CATAATTTTCATTTGTTCTTCATTCCCATG		gDNA	56	30	14819	14848	30			
Lin7-CD1-1_R	GTTGTGTTTAGTCTAGATTTAATACTAGTGA	TCACTAGTATTAAATCTAGACTAAACACAAC	gDNA	55	29	15206	15236	31	417	450	
Lin7-CD1-2_F	GTTTTTATTCACTTGCAATCTCAAAATGCTGTA		gDNA	59	30	15054	15086	33			
Lin7-CD1-2_R	TCTGTTATACTTTAGAGACATTTTTGCCGTC	GACGGCAAAAATGTCTCTAAAGTATAACAGA	gDNA	59	35	15529	15559	31	505	500	
Lin8-CD1-1_F	GTGAACTAGTATTCAGCGTAATGTAACC		gDNA	57	39	532	559	28			
Lin8-CD1-1_R	AATGACCTAGTTATAATTTTGTGGTGC	GCACCACAAAATTATAACTAGGTCATT	gDNA	56	33	1031	1057	27	525	550	
Lin8-CD1-2_F	GCTTCACACAGAGTTTTTCCAGG		gDNA	58	48	797	819	23			
Lin8-CD1-2_R	ATCATTGCATTCTATAATTATGACCTCATG	CATGAGGTCATAATTATAGAATGCAATGAT	gDNA	56	30	1083	1112	30	315	500	
Sl-1-1_F	ACTGGGTCAAGTTCAAAGGCA		cDNA	58	48	656	676	21			
Sl-1-1_R	AGATCATTATGCTATTGGTACGTATGACTTGG	CCAAGTCATACGTACCAATAGCATAATGATCT	cDNA	60	38	1002	1033	32	377	500	
Hk1-CD1-1_F	GAGCTTGCAAAATTTGTTGCTGC		cDNA	58	43	448	470	23			
Hk1-CD1-1_R	TGGACAAAAGGCTTCTCTATTGATGATG	CATCATCAATAGAGAAGCCTTTTGTCCA	cDNA	59	39	577	604	28	156	200	
Hk2-CD1-1_F	CTATATTGCGGCAGAACTTGCAAAATTCGTAGA		cDNA	63	39	435	467	33			
Hk2-CD1-1_R	AGGTGGACGAAAGGTTTCTCCATTGATGAC	GTCATCAATGGAGAAACCTTTCGTCCACCT	cDNA	64	47	574	603	30	168	200	
Hk3-CD1-1_F	CACGGCACTTGCAGAATTTGTAGCTA		cDNA	61	46	486	511	26			
Hk3-CD1-1_R	TAAATGGACAAAGGGCTTCTCCATAGAAG	CTTCTATGGAGAAGCCCTTTGTCCATTTA	cDNA	60	41	615	643	29	157	200	
Hk4-CD1-1_F	CACAAAAGGAAGGTGGTAATTTTGAGTTGCA		cDNA	62	39	476	506	31			
Hk4-CD1-1_R	ATGGACAAAAGGTTTTGCTGTCTC	GAGACAGCAAAACCTTTTGTCCAT	cDNA	58	42	588	611	24	135	200	
Fk1cd2-1F	TTGAAGAGGCTTTTCCATCCTAAT		cDNA	56	38	717	740	24			
Fk1cd2-1R	TGATGCTAGTCTTTATCAGGATGAAAA	TTTTCATCCTGATAAAGACTAGCATCA	cDNA	56	33	892	918	27	201	500	
Fk1cd2-2F	CAAGGGAAGAGTAAATTCGATCAAGGTAA		cDNA	59	38	800	828	29			
Fk1cd2-2R	TAGCTTCTGATGCTAGTCTTTATCAGGATG	CATCCTGATAAAGACTAGCATCAGAAGCTA	cDNA	60	40	886	915	30	115	200	
Fk2cd2-1Rnew	CTTCTAACCAAGATTGTTGATGATCAAACCATTCT										
UGP F	GCACAAAGTGCATGTTGCGGATGTTGA										
UGP R	CATGTTTTGTCTAATCACCTCTCCCCGGTA										
SUS3 F	GACTCACAATTTGAATGGCCAGTTCCG										
SUS3 R	CAGCAACAGGAGATTAGAATCACTTTGAATTAATG										
SUS2 F	CTCCACGGGTGTCGTAACGTTGGA										
SUS2 R	CAGACCACAAGTCATTGCTTCGACAACA										

Table 4 b)

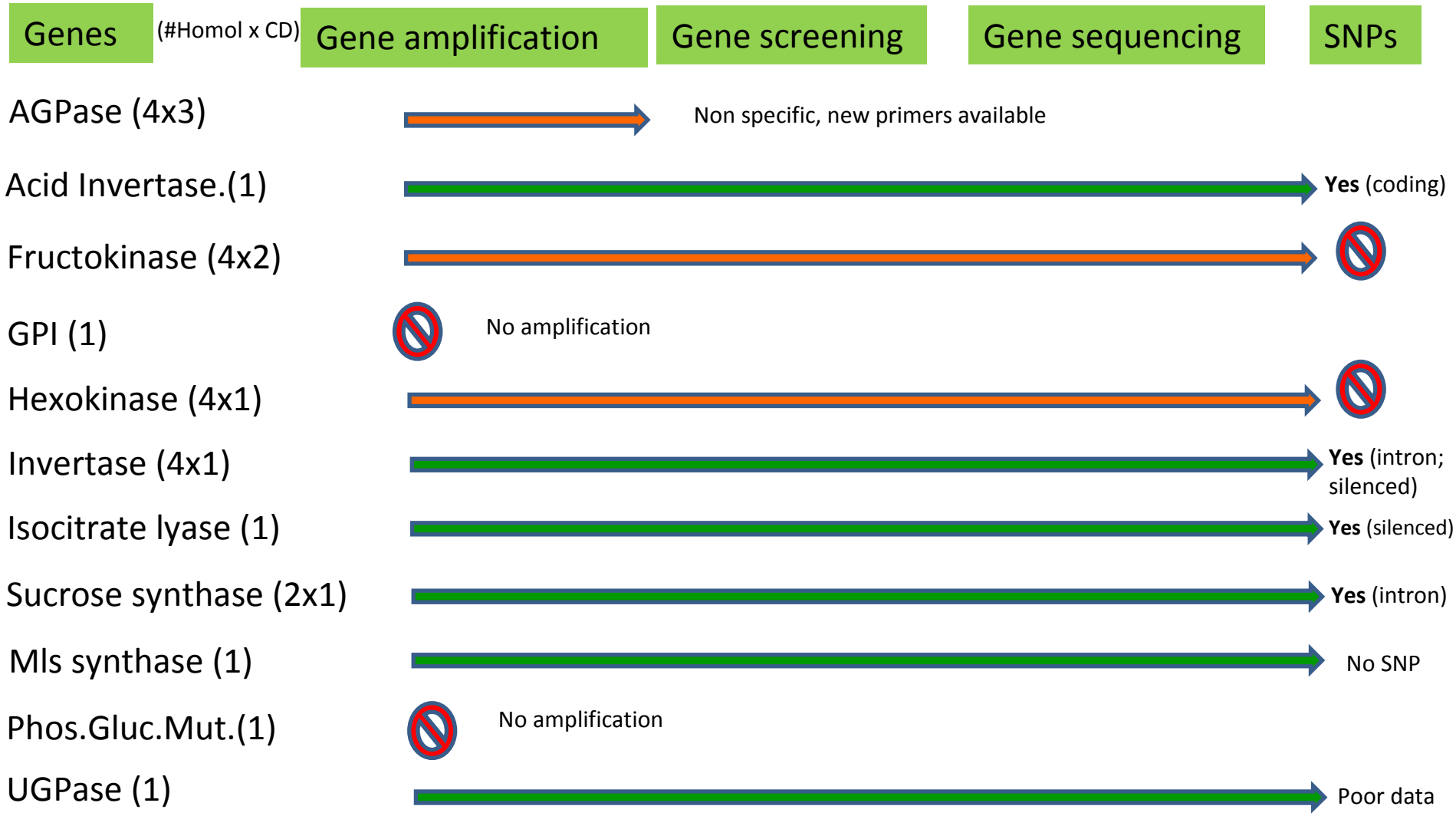

Genes
(#Homol x CD)
Gene amplification
Gene screening
Gene sequencing
SNPs
AGPase (4x3)
Non specific, new primers available
Acid Invertase.(1)
Yes (coding)
Fructokinase (4x2)
GPI (1)
No amplification
Hexokinase (4x1)
Invertase (4x1)
Yes (intron; silenced)
Isocitrate lyase (1)
Yes (silenced)
Sucrose synthase (2x1)
Yes (intron)
Mls synthase (1)
No SNP
Phos.Gluc.Mut.(1)
No amplification
UGPase (1)
Poor data

Table 5

Gene	Plate position (+box# for EUSOL acc.)	Accession		Polymorphism	Nucleotide change		gDNA or cDNA	Exon/Intron	Position on ref seq	General haplotype	Position AA	Predicted effect
		EUSOL	CBSG		Reference	variant						
Lin8	B9		23	SNP	AA	GA	gDNA	Intron	257	7	_	_
Lin8	B12		26	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	C6		33	SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	C7		34	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	C8		35	SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	D3		43	SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	D4		44	SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	D8		48	SNP	AA	GA	gDNA	Intron	257	7	_	_
Lin8	E3		58	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	E4		59	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	F1		79	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	F4		83	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	G10		110	SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	G11		111	SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	G2		94	SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	G4		96	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	H5		124	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	A7		8	SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	H10		132	SNP	AA	GA	gDNA	Intron	257	7	_	_
Lin8	B1/G1	3151		SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	B1/D8	1234		SNP	AA	GG	gDNA	Intron	257	9	_	_
Lin8	B1/F5	3001		SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	B3/C9	2960		SNP	AA	GA	gDNA	Intron	257	8	_	_
Lin8	B1/E11	2732		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	B1/E6	2146		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	B1/B4	3686		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	B1/D4	1185		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	B3/B11	2523		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	B3/E11	951		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	B3/A10	2018		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	B3/F12	891		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	B3/F5	892		SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	A3		3	SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	F10		90	SNP	TT	AA	gDNA	Intron	445	10	_	_
Lin8	A11		12	SNP	TT	TA	gDNA	Intron	445	11	_	_

Table 6 a)

Gene	Plate position (+box# for EUSOL acc.)	Accession		Polymorphism	Nucleotide change		gDNA or cDNA	Exon/Intron	Position on ref seq	General haplotype	Position AA	Predicted effect
		EUSOL	CBSG		Reference	variant						
Lin8	G5		98	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	E6		68	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	B5		18	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	H4		123	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	G8		102	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	F8		88	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	B6		19	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	E1		56	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	E5		61	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	D11		53	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	H7		126	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	D10		50	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	H1		119	SNP	TT	TA	gDNA	Intron	445	11	_	_
Lin8	B9		23	SNP	TT	TA	gDNA	Intron	445	7	_	_
Lin8	D8		48	SNP	TT	TA	gDNA	Intron	445	7	_	_
Lin8	H10		132	SNP	TT	TA	gDNA	Intron	445	7	_	_
Lin8	B9		23	SNP	CC	TC	gDNA	exon	383	7	not defined	silenced
Lin8	B12		26	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	C6		33	SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	C7		34	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	C8		35	SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	D3		43	SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	D4		44	SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	D8		48	SNP	CC	TC	gDNA	exon	383	7	not defined	silenced
Lin8	E3		58	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	E4		59	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	F1		79	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	F4		83	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	G10		110	SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	G11		111	SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	G2		94	SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	G4		96	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	H5		124	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	A7		8	SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	H10		132	SNP	CC	TC	gDNA	exon	383	7	not defined	silenced
Lin8	B1/G1	3151		SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	B1/D8	1234		SNP	CC	TT	gDNA	exon	383	9	not defined	silenced

Table 6 b)

Gene	Plate position (+box# for EUSOL acc.)	Accession		Polymorphism	Nucleotide change		gDNA or cDNA	Exon/Intron	Position on ref seq	General haplotype	Position AA	Predicted effect
		EUSOL	CBSG		Reference	variant						
Lin8	B1/D8	1234		SNP	CC	TT	gDNA	exon	383	9	not defined	silenced
Lin8	B1/F5	3001		SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin8	B3/C9	2960		SNP	CC	TC	gDNA	exon	383	8	not defined	silenced
Lin7	C6		33	SNP	T	G	gDNA	intron	123	4	_	_
Lin7	C6		33	SNP	C	T	gDNA	intron	309	4	_	_
Lin7	C6		33	SNP	A	G	gDNA	intron	425	4	_	_
Lin7	E11		74	SNP	T	G	gDNA	intron	123	4	_	_
Lin7	E11		74	SNP	C	T	gDNA	intron	309	4	_	_
Lin7	E11		74	SNP	A	G	gDNA	intron	425	4	_	_
Lin7	E4		59	SNP	T	G	gDNA	intron	123	4	_	_
Lin7	E4		59	SNP	C	T	gDNA	intron	309	4	_	_
Lin7	E4		59	SNP	A	G	gDNA	intron	425	4	_	_
Lin7	G11		111	SNP	T	G	gDNA	intron	123	4	_	_
Lin7	G11		111	SNP	C	T	gDNA	intron	309	4	_	_
Lin7	G11		111	SNP	A	G	gDNA	intron	425	4	_	_
Lin7	G4		96	SNP	T	G	gDNA	intron	123	5	_	_
Lin7	G4		96	SNP	C	G	gDNA	intron	324	5	_	_
Lin7	G4		96	SNP	T	A	gDNA	intron	325	5	_	_
Lin7	G4		96	SNP	T	A	gDNA	intron	326	5	_	_
Lin7	G4		96	SNP	A	T	gDNA	intron	332	5	_	_
Lin7	G4		96	SNP	A	C	gDNA	intron	333	5	_	_
Lin7	G4		96	SNP	A	T	gDNA	intron	334	5	_	_
Lin7	G4		96	SNP	A	C	gDNA	intron	343	5	_	_
Lin7	G4		96	SNP	A	G	gDNA	intron	425	5	_	_
Lin7	C1		27	SNP	T	G	gDNA	intron	123	6	_	_
Lin7	C6		33	Indel	-	T	gDNA	intron	308	4	_	_
Lin7	C6		33	Indel	G	-	gDNA	intron	380	4	_	_
Lin7	C6		33	Indel	T	-	gDNA	intron	381	4	_	_
Lin7	C6		33	Indel	A	-	gDNA	intron	382	4	_	_
Lin7	C6		33	Indel	T	-	gDNA	intron	383	4	_	_
Lin7	C6		33	Indel	A	-	gDNA	intron	384	4	_	_
Lin7	C6		33	Indel	T	-	gDNA	intron	385	4	_	_
Lin7	C6		33	Indel	T	-	gDNA	intron	386	4	_	_
Lin7	C6		33	Indel	T	-	gDNA	intron	387	4	_	_
Lin7	C6		33	Indel	T	-	gDNA	intron	388	4	_	_
Lin7	C6		33	Indel	T	-	gDNA	intron	389	4	_	_
Lin7	C6		33	Indel	T	-	gDNA	intron	390	4	_	_

Table 6 c)

Gene	Plate position (+box# for EUSOL acc.)	Accession		Polymorphism	Nucleotide change		gDNA or cDNA	Exon/Intron	Position on ref seq	General haplotype	Position AA	Predicted effect
		EUSOL	CBSG		Reference	variant						
Lin7	C6		33	Indel	T	-	gDNA	intron	390	4	_	_
Lin7	E11		74	Indel	-	T	gDNA	intron	308	4	_	_
Lin7	E11		74	Indel	G	-	gDNA	intron	380	4	_	_
Lin7	E11		74	Indel	T	-	gDNA	intron	381	4	_	_
Lin7	E11		74	Indel	A	-	gDNA	intron	382	4	_	_
Lin7	E11		74	Indel	T	-	gDNA	intron	383	4	_	_
Lin7	E11		74	Indel	A	-	gDNA	intron	384	4	_	_
Lin7	E11		74	Indel	T	-	gDNA	intron	385	4	_	_
Lin7	E11		74	Indel	T	-	gDNA	intron	386	4	_	_
Lin7	E11		74	Indel	T	-	gDNA	intron	387	4	_	_
Lin7	E11		74	Indel	T	-	gDNA	intron	388	4	_	_
Lin7	E11		74	Indel	T	-	gDNA	intron	389	4	_	_
Lin7	E11		74	Indel	T	-	gDNA	intron	390	4	_	_
Lin7	E4		59	Indel	-	T	gDNA	intron	308	4	_	_
Lin7	E4		59	Indel	G	-	gDNA	intron	380	4	_	_
Lin7	E4		59	Indel	T	-	gDNA	intron	381	4	_	_
Lin7	E4		59	Indel	A	-	gDNA	intron	382	4	_	_
Lin7	E4		59	Indel	T	-	gDNA	intron	383	4	_	_
Lin7	E4		59	Indel	A	-	gDNA	intron	384	4	_	_
Lin7	E4		59	Indel	T	-	gDNA	intron	385	4	_	_
Lin7	E4		59	Indel	T	-	gDNA	intron	386	4	_	_
Lin7	E4		59	Indel	T	-	gDNA	intron	387	4	_	_
Lin7	E4		59	Indel	T	-	gDNA	intron	388	4	_	_
Lin7	E4		59	Indel	T	-	gDNA	intron	389	4	_	_
Lin7	E4		59	Indel	T	-	gDNA	intron	390	4	_	_
Lin7	G11		111	Indel	-	T	gDNA	intron	308	4	_	_
Lin7	G11		111	Indel	G	-	gDNA	intron	380	4	_	_
Lin7	G11		111	Indel	T	-	gDNA	intron	381	4	_	_
Lin7	G11		111	Indel	A	-	gDNA	intron	382	4	_	_
Lin7	G11		111	Indel	T	-	gDNA	intron	383	4	_	_
Lin7	G11		111	Indel	A	-	gDNA	intron	384	4	_	_
Lin7	G11		111	Indel	T	-	gDNA	intron	385	4	_	_
Lin7	G11		111	Indel	T	-	gDNA	intron	386	4	_	_
Lin7	G11		111	Indel	T	-	gDNA	intron	387	4	_	_
Lin7	G11		111	Indel	T	-	gDNA	intron	388	4	_	_
Lin7	G11		111	Indel	T	-	gDNA	intron	389	4	_	_
Lin7	G11		111	Indel	T	-	gDNA	intron	390	4	_	_

Table 6 d) and e) (next page)

Gene	Plate position (+box# for EUSOL acc.)	Accession		Polymorphism	Nucleotide change		gDNA or cDNA	Exon/Intron	Position on ref seq	General haplotype	Position AA	Predicted effect
		EUSOL	CBSG		Reference	variant						
Lin7	G11		111	Indel	T	-	gDNA	intron	390	4	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	294	5	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	295	5	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	296	5	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	297	5	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	298	5	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	299	5	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	300	5	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	301	5	_	_
Lin7	G4		96	Indel	A	-	gDNA	intron	302	5	_	_
Lin7	G4		96	Indel	T	-	gDNA	intron	303	5	_	_
Lin7	G4		96	Indel	-	G	gDNA	intron	331	5	_	_
Lin7	G4		96	Indel	-	G	gDNA	intron	339	5	_	_
Fxk2	B3/E6	2701		SNP	C	G	cDNA	intron	125	2	_	_
Fxk2	B3/E6	2701		SNP	C	A	cDNA	intron	136	2	_	_
Fxk2	B3/E6	2701		SNP	C	A	cDNA	intron	152	2	_	_
Fxk2	B3/E6	2701		SNP	C	G	cDNA	intron	200	2	_	_
Fxk2	B3/E6	2701		SNP	G	A	cDNA	intron	210	2	_	_
Fxk2	B3/E6	2701		SNP	G	A	cDNA	intron	243	2	_	_
Fxk2	B3/C12	2994		SNP	C	G	cDNA	intron	92	3	_	_
Fxk2	B3/E6	2701		Indel	-	A	cDNA	intron	116	2	_	_
Fxk2	B3/E6	2701		Indel	-	T	cDNA	intron	117	2	_	_
Fxk2	B3/E6	2701		Indel	-	A	cDNA	intron	119	2	_	_
Fxk2	B3/E6	2701		Indel	-	G	cDNA	intron	121	2	_	_
Fxk2	B3/E6	2701		Indel	-	G	cDNA	intron	122	2	_	_
Fxk2	B3/E6	2701		Indel	-	C	cDNA	intron	124	2	_	_
Fxk2	B3/E6	2701		Indel	-	T	cDNA	intron	126	2	_	_
Fxk2	B3/E6	2701		Indel	-	A	cDNA	intron	129	2	_	_
Fxk2	B3/E6	2701		Indel	-	T	cDNA	intron	265	2	_	_
Fxk2	B3/E6	2701		Indel	-	A	cDNA	intron	267	2	_	_
Fxk2	B3/E6	2701		Indel	-	G	cDNA	intron	268	2	_	_
Fxk2	B3/E6	2701		Indel	-	T	cDNA	intron	269	2	_	_
Fxk2	B3/E6	2701		Indel	-	A	cDNA	intron	270	2	_	_
Fxk2	B3/E6	2701		Indel	-	C	cDNA	intron	271	2	_	_
Fxk2	B3/E6	2701		Indel	A	T	cDNA	intron	272	2	_	_
Fxk2	B3/C12	2994		Indel	A	G	cDNA	intron	257	3	_	_
Fxk2	B3/C12	2994		Indel	A	-	cDNA	intron	266	3	_	_
Fxk2	B3/C12	2994		Indel	A	-	cDNA	intron	272	3		
Tiv1	B3/E6	2701		SNP	T	G	cDNA	exon	708	12	not defined	silenced
Tiv1	B3/E6	2701		SNP	G	A	cDNA	exon	802	12	not defined	silenced
Tiv1	B3/E6	2701		SNP	AA	AG	cDNA	exon	807	12	not defined	silenced
Tiv1	B3/E6	2701		SNP	AA	AG	cDNA	exon	919	12	307	K to E
Susy3	B2/G8	5891		SNP	G	A	cDNA	intron	74 (to check)	13	_	_
Icl	B3/E6	2701		SNP	T	C	cDNA	exon	691	14	not defined	silenced
Icl	B3/E6	2701		SNP	TT	TG	cDNA	exon	700	14	not defined	silenced

Table 7

CBSG set	Genotype			Phenotype		Metabolites		
General Haplotyp		Haplotype Lin7	Haplotype Lin8	Brix	Fruit weight	Fructose	Glucose	Sucrose
	REF	1	1					
1	12	1	11	4.31	4.09	2.9	4.44	2.69
1	18	1	11	5	4.05	3	4.54	3.18
1	19	1	11	4.77	4.11	3.01	4.56	3.16
1	50	1	11	4.25	4.07	2.98	4.53	3.14
1	53	1	11	4.86	4.09	2.87	4.46	2.96
1	56	1	11	4.71	4.08	2.89	4.42	3.26
1	61	1	11	5.51	4.11	3.06	4.58	3.41
1	68	1	11	5.19	4.07	3.01	4.55	3.28
1	88	1	11	4.32	4.09	2.93	4.52	3.18
1	98	1	11	4.47	3.96	3	4.58	3.29
1	102	1	11	4.66	4.09	3	4.49	3.09
1	119	1	11	4.83	3.97	3.01	4.53	3.17
1	123	1	11	4.6	4.08	3.06	4.59	3.08
1	126	1	11	4.72	4.01	3.04	4.59	3.13
Average values				**4.73**	**4.06**	**2.98**	**4.53**	**3.14**
2	8	1	8	5.22	4.08	3.02	4.58	3.19
2	26	1	8	4.68	4.1	3.07	4.65	3.18
2	34	1	8	4.7	4.09	3.03	4.59	3.08
2	58	1	8	6.78	3.86	3.18	4.71	3.73
2	59	1	8	7.35	3.99	3.21	4.8	3.94
2	79	1	8	6.5	3.91	3.05	4.59	3.77
2	83	1	8	6.3	3.98	3.14	4.71	3.76
2	96	1	8	5.71	4.01	3.09	4.7	3.53
2	124	1	8	5.02	4.09	3.01	4.58	3.02
Average values				**5.81**	**4.01**	**3.09**	**4.66**	**3.47**
3	33	1	9	7.91	3.89	3.23	4.77	3.54
3	35	1	9	6.86	3.95	3.01	4.63	3.56
3	43	1	9	4.39	3.98	3.01	4.57	3.02
3	44	1	9	4.87	4.04	3.03	4.53	3.19
3	94	1	9	4.66	4.06	2.93	4.71	3.13
3	110	1	9	6.79	3.96	3.22	4.75	3.9
3	111	1	9	7.2	4.02	3.19	4.74	3.74
Average values				**6.10**	**3.99**	**3.09**	**4.67**	**3.44**
4	23	1	7	4.75	4.1	2.93	4.46	3.17
4	48	1	7	5.47	4.08	3.11	4.68	3.45
4	132	1	7	5.89	4.04	3.16	4.72	3.18
Average values				**5.37**	**4.07**	**3.07**	**4.62**	**3.27**
5	3	1	10	6.42	3.96	3.19	4.75	3.4
5	90	1	10	4.72	3.92	2.97	4.58	3.13
Average values				**5.57**	**3.94**	**3.08**	**4.67**	**3.27**
6	33	4	1	7.91	3.89	3.23	4.77	3.54
6	59	4	1	7.35	3.99	3.21	4.8	3.94
6	74	4	1	7.69	3.96	3.19	4.83	3.81
6	111	4	1	7.2	4.02	3.19	4.74	3.74
Average values				**7.54**	**3.97**	**3.21**	**4.79**	**3.76**
7	96	5	1	**5.71**	**4.01**	**3.09**	**4.70**	**3.53**
8	27	6	1	**4.44**	**3.99**	**2.91**	**4.50**	**3.15**
Average values on 94 CBSG accessions	Mean value			5.03	4.03	3.01	4.57	2.69
	Minimum value			3.94	3.75	2.79	4.42	4.01
	Maximum value			7.91	4.20	3.24	4.83	0.26
	Standard deviation			0.89	0.08	0.09	0.09	0.26
	# Individuals			n=94	n=94	n=94	n=94	n=94

EU-SOL Core Collection		Genotype					Phenotype						Metabolites				
General Haplotype	Taxa/Country of origin	Haplo. Lin8	Haplo. Fruc.K2	Haplo. Sol.Inv	Haplo. Iso.Lyase	Haplo. Susy3	Brix 2008	Brix 2009	Brix 2009-2	Brix gh 2010	Haplotype mean	FW (gr.)*	Fructose	Fructose-6-P	Glucose	Glucose-6-P	Sucrose
	REF	1	1	1	1	1											
1	891 na/Guatemala	10	1	1	1	1	4.7	4.9	4.8	na	_	14.0	0.831	0.874	0.809	0.922	0.916
1	892 na/Guatemala	10	1	1	1	1	4.5	4.7	na	4.57	_	322.0	na	na	na	na	na
1	951 na/Unknown	10	1	1	1	1	7.3	6.7	na	4.77	_	65.2	na	na	na	na	na
1	1185 na/Unknown	10	1	1	1	1	7.5	6.5	na	6.13	_	37.6	na	na	na	na	na
1	2018 *Lycopersicon*/Unknown	10	1	1	1	1	4	5.4	na	na	_	53.0	na	na	na	na	na
1	2146 *Solanum lycopersicum var. esculentum Mill. Convar. infiniens Lehm. Var. commune Bail.*/Poland	10	1	1	1	1	7.5	6.6	6	na	_	38.0	1.084	0.697	1.077	0.729	2.427
1	2523 *Solanum lycopersicum var. esculentum Mill.* /Bulgaria	10	1	1	1	1	8.5	6.6	na	na	_	6.0	na	na	na	na	na
1	2732 *Solanum lycopersicum var. esculentum Mill. Convar. Fruticosum. Lehm. Var. finiens. Lehm.* /Costa Rica	10	1	1	1	1	7.7	6.6	5.5	na	_	75.6	na	na	na	na	na
1	3686 *Solanum lycopersicum var. esculentum Mill.*/Peru	10	1	1	1	1	6.6	5.5	na	na	**6.0**	57.0	na	na	na	na	na
2	1234 na/ Russian Federation	9	1	1	1	1	7.1	4.4	5.1	na	_	62.8	0.749	0.886	0.719	0.886	0.580
2	3151 *Solanum lycopersicum var. esculentum. Mill. Convar. fruticosum Lehm. Var. finiens Lehm* /Argentina	9	1	1	1	1	8.1	6	na	na	**6.1**	46.4	na	na	na	na	na
3	2960 *Solanum pimpinellifolium (Jusl.) Mill.*/CzechRepublic	8	1	1	1	1	8.6	6.4	na	na	_	7.2	na	na	na	na	na
3	3001 *Solanum lycopersicum var. esculentum. Mill. Convar. fruticosum Lehm. Var. finiens lehm.*/Russian	8	1	1	1	1	8.6	4.9	na	na	**7.1**	58.8	na	na	na	na	na
4	5891 *Solanum lycopersicum var. cerasiforme/* United States	1	1	1	1	13	9.7	5.4	na	na	**7.6**	3.6	na	na	na	na	na
5	2994 *Solanum pimpinellifolium (Jusl.) Mill. Var. pimpinellifolium./* Russian Federation	1	3	1	1	1	8.7	6.9	na	na	**7.8**	*0.666	1.107	0.000	1.086	0.000	4.981
6	2701 *Solanum pennellii* / Peru	1	2	12	14	1	8.5	10.2	na	na	**9.35**	na	na	na	na	na	na
Average values on EU-SOL core collections		Mean value					na	5.47	5.56	na	_	65.17	0.97	0.98	0.96	0.94	1.63
		Minimum value					na	0.59	3.00	na	_	0.77	0.16	0.47	0.14	0.56	0.03
		Maximum value					na	10.20	9.90	na	_	334.50	2.49	2.57	2.71	1.50	8.51
		Standard deviation					na	2.82	1.41	na	_	56.97	0.34	0.31	0.36	0.20	1.47
Haplotype 1= identical to ref. sequence *: weighted on 5 fruits **: weighted on 4 fruits ***: weighted on 15 fruits		# Individuals					na	293	152	na	_	138	171	171	171	171	171

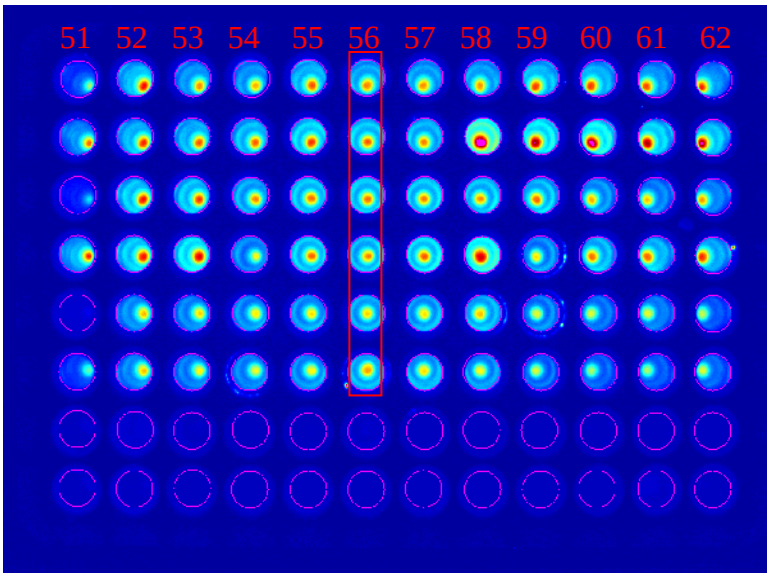
Table 8

Compare Keygen Keypoint™ experiment with Eco-TILLING using High Res-Melting screening technique

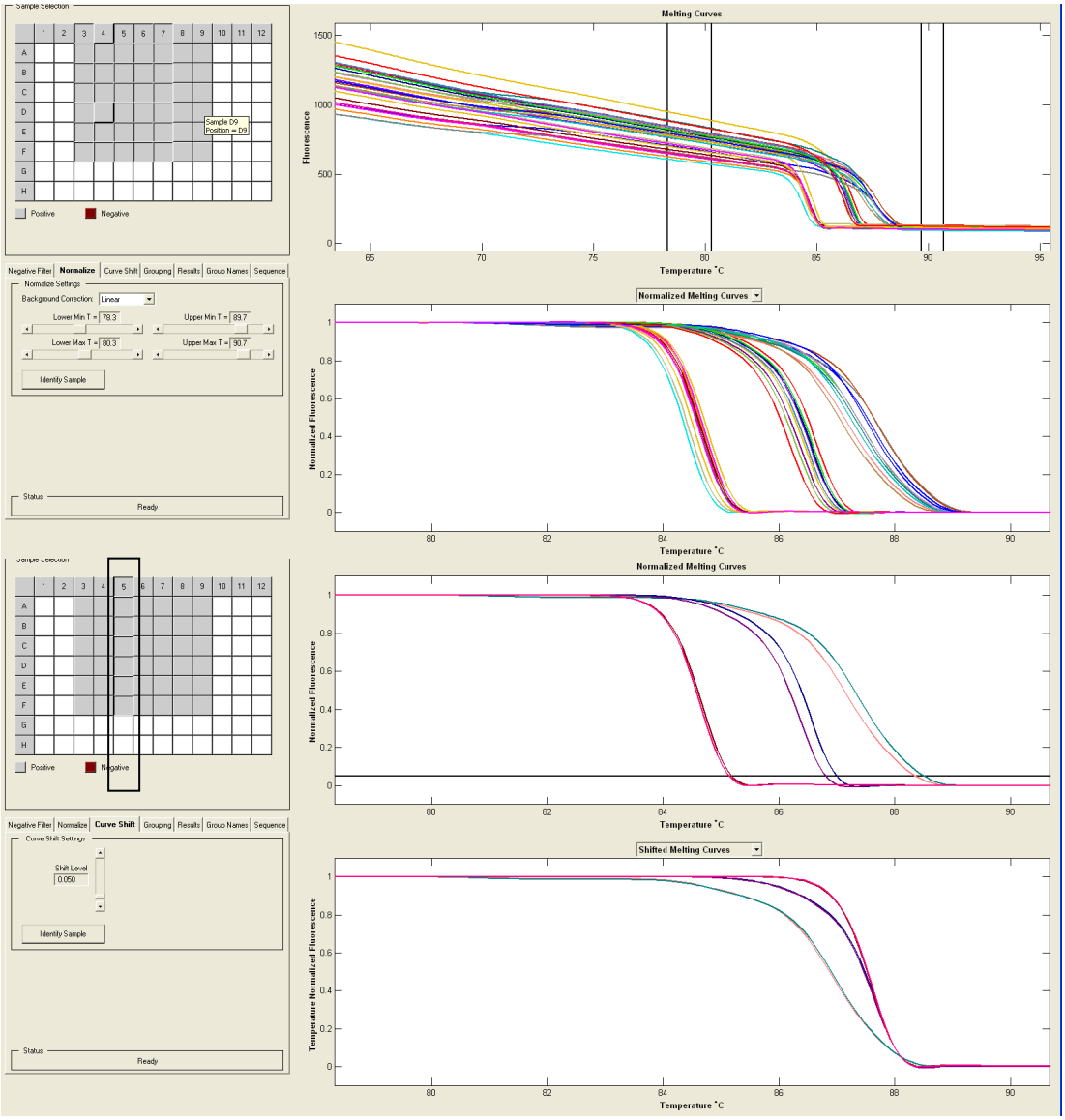
- **Primer ordering**

name primer	**sequence** (5'----> 3')	**size (bp)**
Exon1amp_1-287 bp_F	ATGGCAGCAGCTGAAATGG	19
Exon1amp_1-287 bp_R	CCCCAAAAAATTTTCAACAGTG	21
Exon2 amp_2-402 bp_F	TGCTTACAATAATATCCATCAC	22
Exon2 amp_2-402 bp_R	CCTGAGCTGTTTCATTTGC	19
Exon5 amp_3-200 bp_F	TTAGCATTGGTAAGCAATGG	20
Exon5 amp_3-200 bp_R	CTATACGGTGTAACGATTC	19

- **Gradient PCR with new primer set from KG**



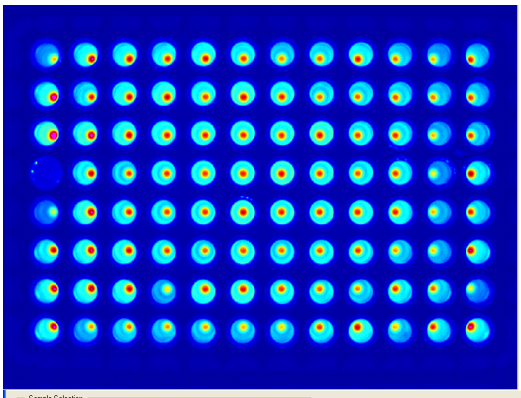
- **Melting curves for each primer pairs**

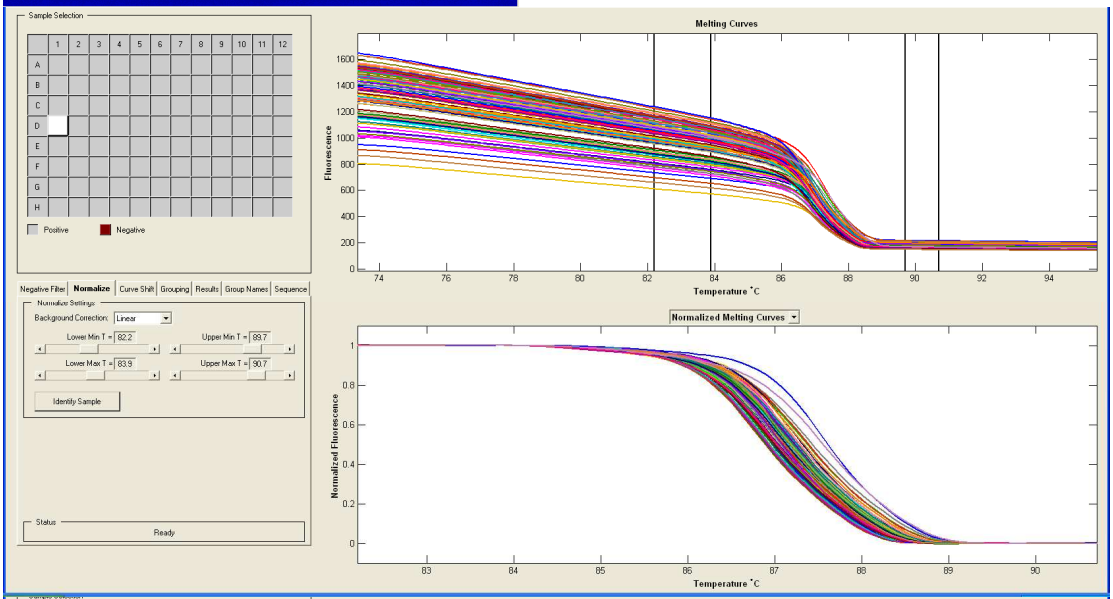


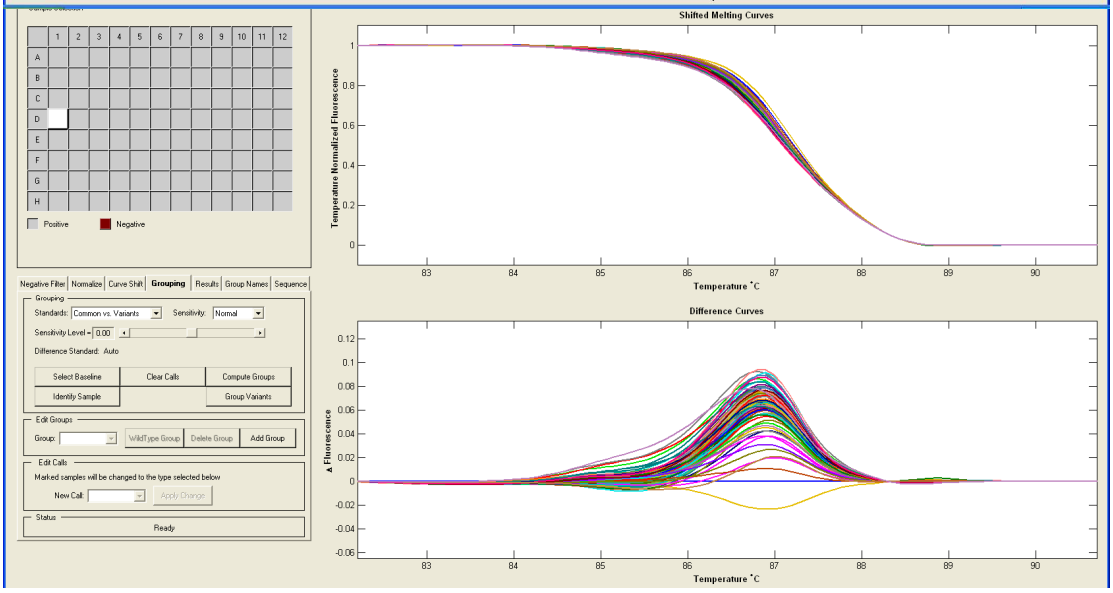
From Keypoint paper, 56 degree nevertheless 55 degree seems the best for lightscanner experiment

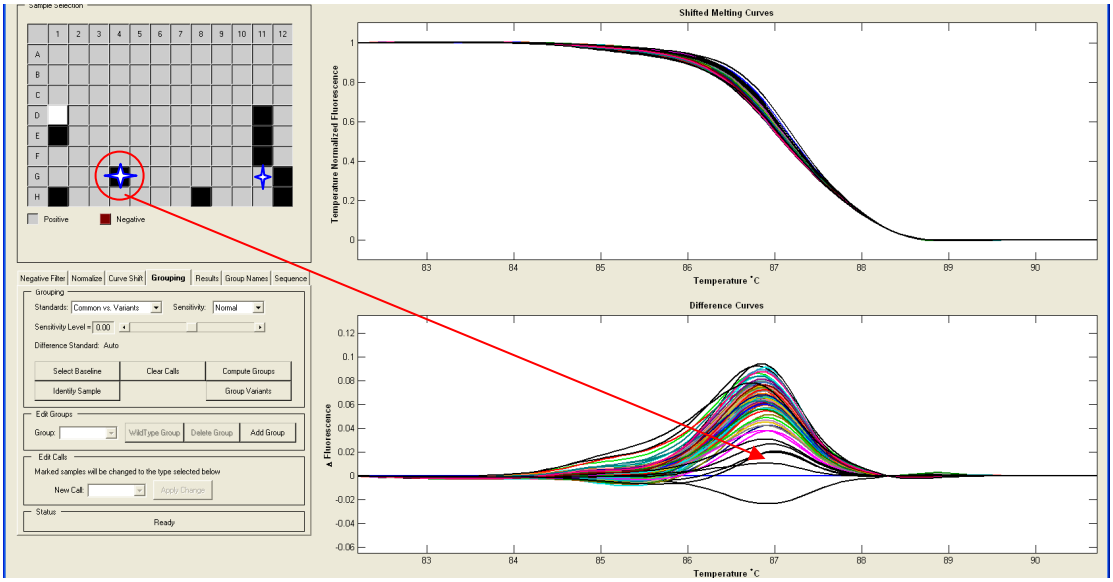
- **Results**

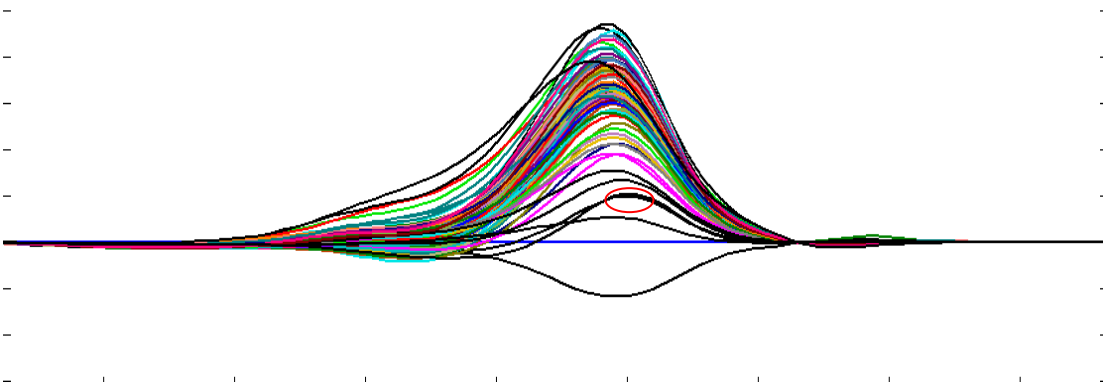
Amplicon 1: E1A1



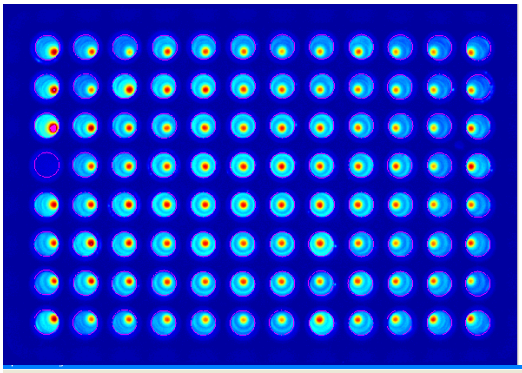


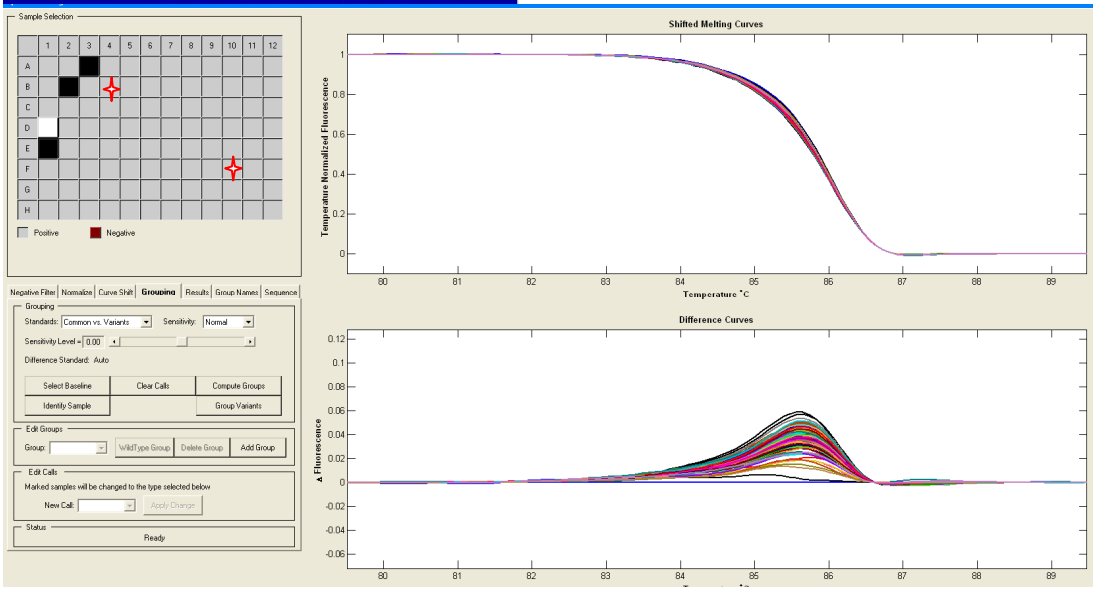



Sample Selection
1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
H
Positive
Negative
Negative Filter | Normalize | Curve Shift | Grouping | Results | Group Names | Sequence
Grouping
Standards: Common vs. Variants
Sensitivity: Normal
Sensitivity Level = 0.00
Difference Standard: Auto
Select Baseline
Clear Calls
Compute Groups
Identity Sample
Group Variants
Edit Groups
Group:
WildType Group
Delete Group
Add Group
Edit Calls
Marked samples will be changed to the type selected below
New Call:
Apply Change
Status
Ready
Shifted Melting Curves
Temperature Normalized Fluorescence
Temperature °C
Difference Curves
Δ Fluorescence
Temperature °C

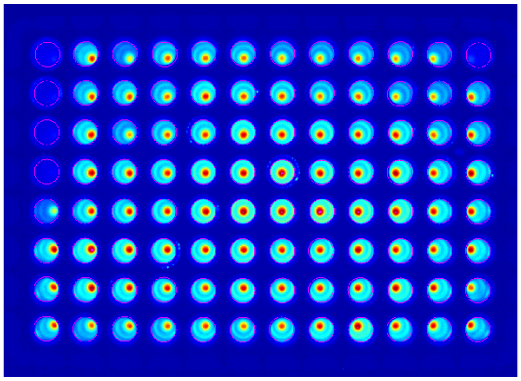


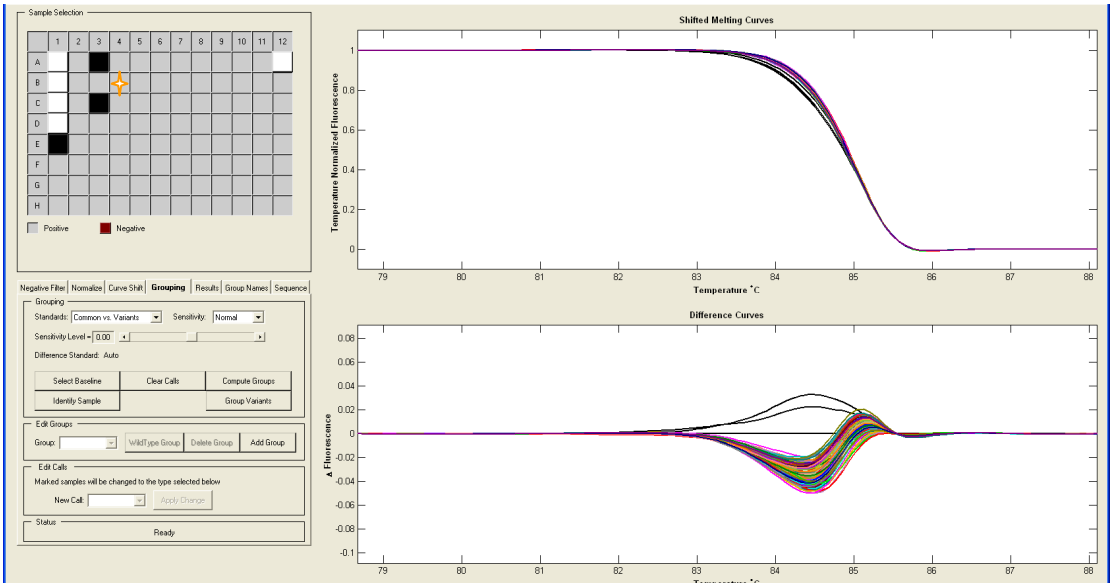
Amplicon 2: E2A2





Amplicon 3: E3A5





E3A5 has to be done again (controls amplification failed)

- **Results compiled from KG study:**

E1A1	mut position (BP)	WT	SNP	plate position	Observed from LS	E2A2	mut position	WT	SNP	plate position	Obeserved from LS	E3A5	mut position	WT	SNP	plate position	Observed from LS
	160	G	A	G11			**47**	G	A	A7,F7			**41**	T	C	B4	
	161	T	A	G11						A11,F11			**84**	C	T	B4	
	172	T	G	D2,D4,D8			**171**	T	G	B4			**85**	A	G	A2,A4,A5,A7,A8,A9,A10,A11	
				E2,E4,E8			**193**	C	T	B4						B2,B4,B5,B7,B8,B9,B10,B11	
				F2,F4,F8			**203**	A	C	F10						D2,D4,D5,D7,D8,D9,D10,D11	
				H2,H4,H8			**209**	T	C	B4						E2,E4,E5,E7,E8,E9,E10,E11	
	210	C	T	G4	X		**245**	C	A	B4						F2,F4,F5,F7,F8,F9,F10,F11	
							266	A	G	B4						H2,H4,H5,H7,H8,H9,H10,H11	
							269	T	A	E2,E4,E8							
										H2,H4,H8							