Confirmation of QTLs for Seed and Seedling Quality Traits in Tomato Using a HIF approach

MSc Minor thesis report (PPH80424)

Habtamnesh Habtemariam

Supervisors: Dr Wilco Ligterink

Leo Willems

Examiners: Dr Wilco Ligterink

Dr Henk Hilhorst

Confirmation of QTLs for Seed and Seedling Quality Traits in Tomato Using a HIF approach

Habtamnesh Habtemariam

(Reg. number: 861103296040)

MSc Plant Biotechnology (Specialisation: Molecular Plant Breeding and Pathology)

Period: October 2012 - January 2013

Wageningen University

Laboratory of plant physiology,

Wageningen seed lab

Acknowledgement

I would like to thank Dr Henk Hilhorst and Dr Wilco Ligterink for allowing me to work on this project and for their guidance throughout the project.

My deepest gratitude goes to Leo Willems for his unreserved guidance and supervision during the experiments and report writing. Leo, you helped me to understand the subject in short time and taught me to be organised and handle experiments efficiently.

I am very thankful to Noorullah Khan for his guidance during seedling experiment.

I would like to thank all members of Wageningen Seedlab for their valuable feedback on my project and for the nice social events we had. It made my stay wonderful in the group.

Finally my deepest gratitude goes to my husband, Adane, and my son Japheth for their love, support and encouragement all the time.

Abstract

Seed quality is a broad term which can be described as the ability of seeds to germinate, tolerate adverse environmental conditions and give rise to a healthy seedling. Seed quality is a complex trait which is determined by multiple factors. Genetic and environmental factors can influence the ultimate seed quality through interactions during seed developmental processes in the mother plant. Seed quality is a quantitatively inherited trait and genetic variation for seed quality in wild and cultivated species is demonstrated. Several tomato seed weight, seed size and seedling traits were studied and quantitative trait loci (QTLs) were identified in our lab. Fine-mapping is a follow up step for a detailed mapping and characterisation of individual loci. Verification and confirmation of the identified QTLs is crucial prior to fine-mapping and applying in a practical breeding program. The aim of this study was to confirm some of the seed and seedling QTLs previously detected on chromosome 9 and 6 of tomato following a heterozygous inbred family (HIF). Eight seed quality traits and 3 plant phenotypes were scored for the HIFs of chromosome 9, and 5 seedling growth traits were scored for the HIF on chromosome 6. The progenies of the HIFs were characterized per segregating flanking markers of the QTLs for hypocotyl length, seed weight and seed size and plant phenotypes. Significant differences for the traits were observed confirming the presence of the QTLs for the traits and it was in agreement with the original QTL data.

Contents

1. Introduction

1.1. Seed quality

Seed is a structure which helps a fully developed plant embryo to disperse and to survive in harsh environments during the period between seed maturation and seedling establishment, thus ensuring the beginning of a next generation (Koornneef, Bentsink and Hilhorst, 2002). Seed takes up a critical position in the life cycle of higher plants as a merely mean of propagation with which new individuals are established and multiplied, with an exception for vegetative propagated plants (Bewley and Black, 1994). A good quality seed contributes to better germination, maximal vigour and ensures successful establishment of plants during the early stages of the plants life-cycle (Hilhorst and Toorop, 1997). Seed quality is a broad term which can be defined as the ability of seeds to germinate, tolerate adverse environmental conditions and give rise to a healthy seedling (Kazmi et al., 2012). Vigour, size, weight and germination are among seed quality traits. Seed vigour determines the performance of the seed lot with regard to rate and uniformity of seed germination, seedling emergence and growth in the field under unfavourable environmental conditions (Ellis, 1992). Seed size is shown to influence seedling growth. A large seed result in faster seedling growth and superior seedlings due to higher amounts of reserve food, but does not guarantee final yield of a plant since this is a complex trait influenced by several genetic and subsequent environmental conditions (Nieuwhof et al., 1989; Pet and Garretsen, 1983). Seed quality is a complex trait which is determined by multiple factors. Genetic and environmental factors can influence the ultimate seed quality through interactions during seed developmental processes in the mother plant (Nieuwhof et al., 1989). Seeds are composed of three genetically distinct structures: the embryo, the endosperm and the seed coat. The embryo contains parental genetic information, the endosperm serves as reserve food, and the seed coat provides protection for the embryo and endosperm. These structures contribute to the variation in seed weight (Doganlar et al., 2000).

1.2. Genetics of seed quality

Marker assisted selection (MAS) is widely being used in agricultural research to assist phenotypic selection by using presence/absence of a marker. MAS leads to more efficient, effective, reliable and cost-effective selection compared to the more classical plant breeding methodology. Many of agriculturally essential traits (e.g. yield, quality and some kinds of

disease resistance) are known as quantitative traits and governed by many genes. Regions on the genome, containing genes associated with a quantitative trait are referred to as quantitative trait loci (QTL) (Collard et al., 2005). Tomato seed weight and seed size are genetically complex, quantitatively inherited traits controlled by additive effects of genes (Nieuwhof et al., 1989; Kazmi et al., 2012). Quantitatively inherited traits are becoming more amenable for genetic analysis because the position of individual QTL and the relative contribution of these loci can be identified (Koornneef, Bentsink and Hilhorst, 2002; Collard et al., 2005). QTL analysis of recombinant inbred lines (RIL) is an important tool that has been used to exploit natural variation (Ligterink et al., 2012). Studies in a RIL population of tomato and Arabidopsis resulted in the discovery of several QTLs responsible for seed, germination and seedling traits (Joosen et al., 2012; Kazmi et al., 2012 and Khan et al., 2012). A RIL population generated from a cross between *Solanum lycopersicum* (cv. Moneymaker) x *Solanum pimpinellifolium* (Pimp) was used in a QTLs analysis to unravel the complex genetics of seed and seedling traits in tomato (Khan et al., 2012; Kazmi et al., 2012). The cultivated tomato cultivar Moneymaker is commercially important and it gives bigger fruits and larger seeds, but it is sensitive to cold, salt and drought stress conditions. Whereas its closest wild species tomato, pimp, produces smaller fruits and seeds, but has a faster germination rate. Although *S. pimpinellifolium* has horticulturally undesirable traits, it is the only species that can natural introgress with *S. lycopesicum*. The RIL population of *S. lycopersicum* x *S. pimpinellifolium* allowed identification of QTLs influencing various seed and seedling phenotypes across the 12 chromosomes of tomato (Kazmi et al., 2012 and Khan et al., 2012). Forty two seed quality traits were analysed and 120 QTLs were identified under normal and stress conditions (Kazmi et al., 2012). Furthermore, 20 QTLs for seedling traits under starvation and nutrient supplied conditions were detected (Khan et al., 2012). Co-location of different seed and seedling QTLs was identified on the bottom of chromosome 1, 4, 6, 9 and 11 (Figure 1).

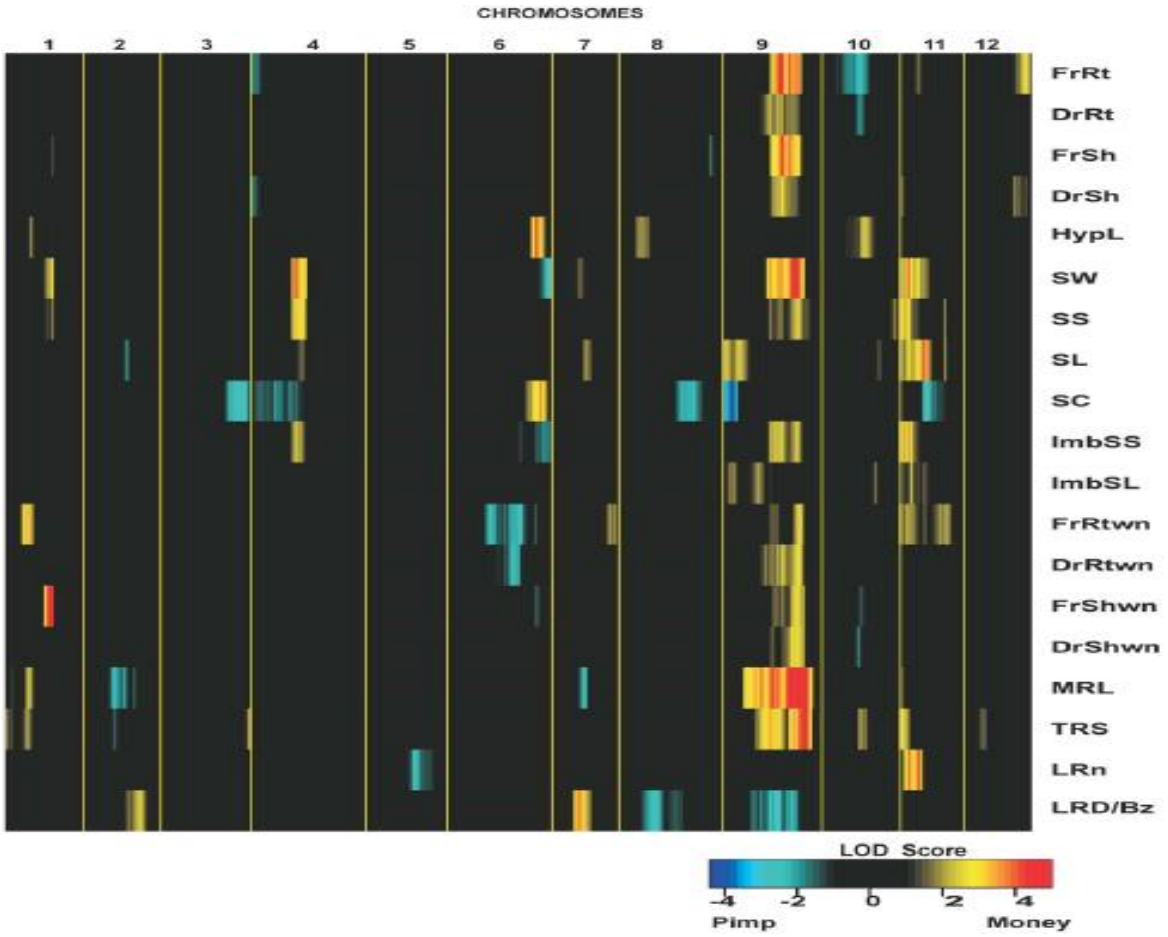


Figure 1. Heatmap of QTLs identified for seed and seedling quality traits across 12 chromosomes of tomato from Khan et al. (2012). Traits are abbreviated as: SW = Seed Weight, SS = Seed Size SL = Seed Length. FrRt = Fresh Root weight, DrRt = Dry Root weight, FrSh = Fresh Shoot weight, DrSh Dry Shoot weight. Similar traits measured in nutrientless condition are abbreviated as FrShwn, DrShWn, FrRtwn, DrRtwn. MRL = Main Root Length, TRS = Total Root Size, LRn = Lateral Root number per main root, LRD/Bz = Lateral Root Density per branched zone. Colors indicate QTLs significant at P = 0.002 in multiple QTL mapping models (1-LOD intervals). Blue and light blue colors indicate larger effect of the trait in *S. pimpinellifolium*, and yellow and red in *S. lycopersicum*.

Verification of QTLs identified with DNA markers is vital before selection and fine-mapping of the QTLs in a practical breeding program (Jakkula et al., 2001; Joosen et al., 2012). QTL confirmation study is recommended as a possibility to study both parental alleles at the locus of interest in an elsewhere homozygous genetic background (Figure 2., Tuinstra et al., 1997). Unlike to classical near isogenic lines, the genetic background of heterozygous inbred family (HIF) comprises a mixture of the two parental genomes.

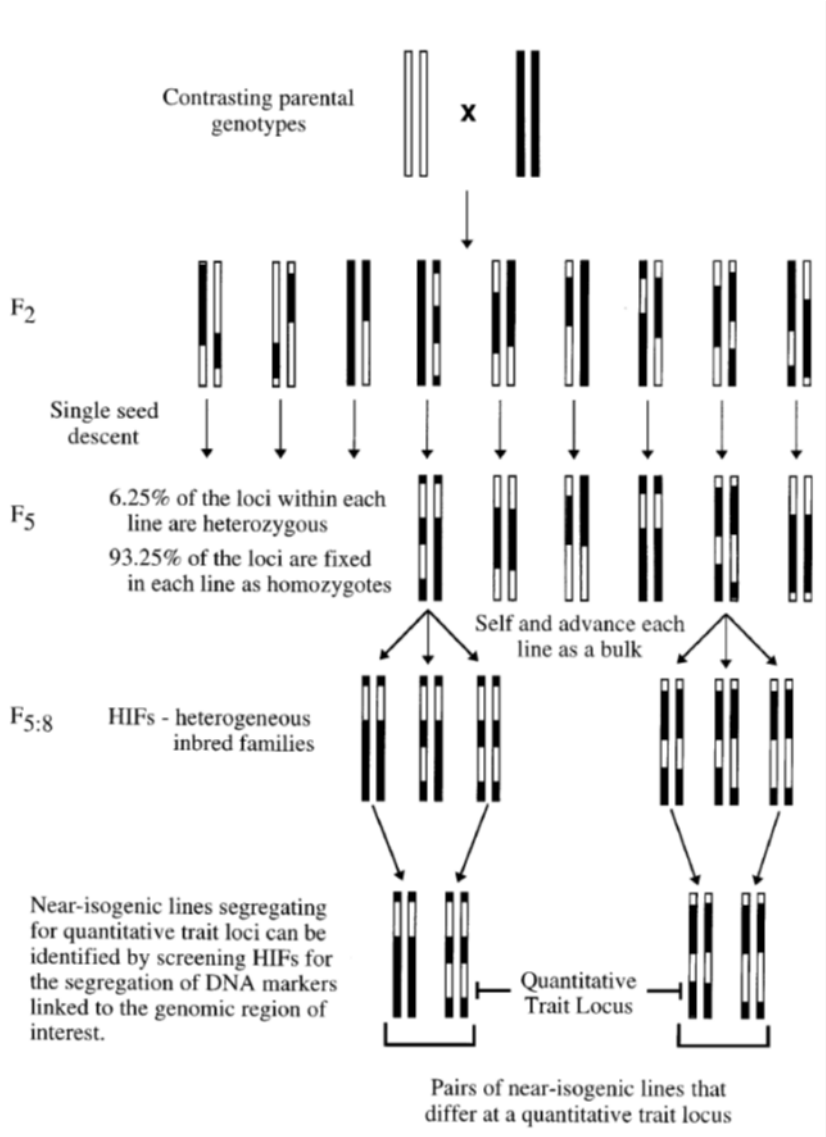


Figure 2. Generating heterogeneous inbred families from recombinant inbred lines and deriving NILs by Tuinstra et al. (1997). Selfing the F1 obtained from two heterozygous parents with single seed decent approach for five generations results in 93.25% of the loci fixed homozygous and 6.25% loci heterozygous with in each line. After continues selfing and advancing each line as a bulk provides heterozygous inbred family. Near-isogenic lines (NILs) that are heterozygous at the regions of interest (QTL) are selected from HIFs by using segregating markers for phenotypic data analysis.

During this minor thesis project in Wageningen Seed Lab, I worked on confirmation of various previously detected QTLs by the researchers in Wageningen Seed Lab (Khan and Kazmi et al., 2012) for seedling traits on chromosome 6 and seed quality traits on chromosome 9 of tomato following the HIF approach.

1.3. Molecular marker

The use of DNA markers that are tightly linked to a trait of interest in plant breeding opened an era of molecular breeding and is being widely used in crop improvement (Collard et al., 2005). Polymorphic DNA markers that reveal differences between individuals of the same or different species are principally useful. A low level of polymorphism at the DNA level among tomato cultivars can be alleviated by using a population derived from an interspecific cross between a cultivar and wild species (Monforte and Tanksley, 2000). Differences in seed phenotype occurred among wild and domesticated/cultivated cultivars of tomato (Khan et al., 2012). The variation is mostly quantitatively controlled by molecular variation at multiple loci in the genome (QTLs) and thereby through multiple genes (Alonso-Blanco et al., 2009). Variation in the DNA sequence of closely related species can be detected by application of molecular markers. The variation can be due to insertion/deletions of a small part of a sequence in genome. The insertion/deletion of ~100 bp long can be detected on agarose gel after PCR amplification of the sequence. Random mutations resulting in base pair substitutions are a source of variation which is called single nucleotide polymorphisms (SNPs). SNPs are an abundant form of DNA variation in the genome and are frequently used in genotyping studies. SNPs can be detected by a suitable restriction endonuclease whose recognition sequence has been changed or introduced by the SNP (Thiel et al., 2004). With application of a PCR assay, SNPs can be identified as Cleaved Amplified polymorphic sequences (CAPs). CAPs are co-dominant, locus specific markers, easily scored and interpreted for use in genotyping studies (Figure 3).

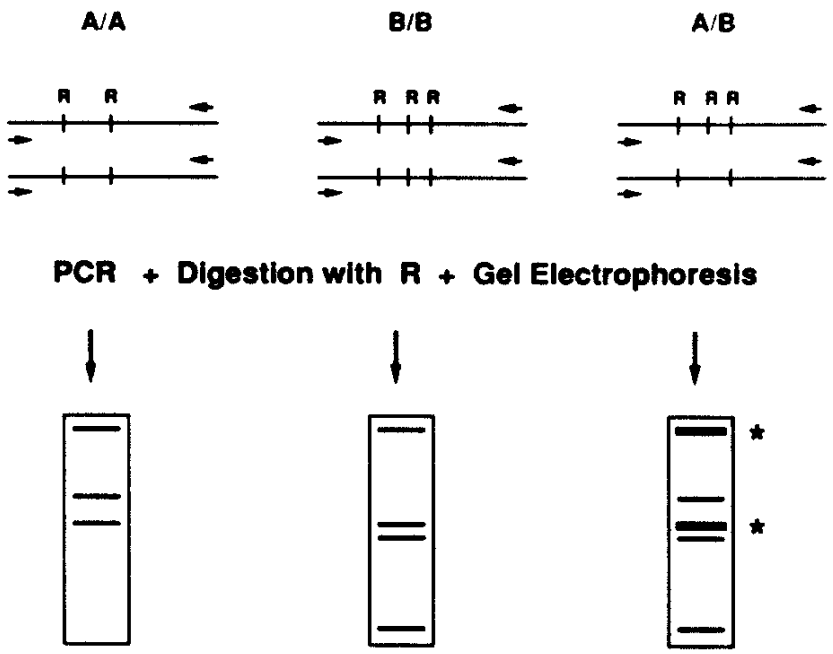


Figure 3. Generation and detection of ecotype specific CAPs markers adopted from Konieczny and Ausubel (1993). Unique sequence primers design at specific locus are used to amplify DNA sequence from related species of homozygous (A/A or B/B) or heterozygous (A/B) background. A/A contains two restriction enzyme (RE) recognition sites whereas, B/B has three RE sites. In case of A/B, two PCR products will be obtained, one which is cleaved three times and one which is cleaved twice. The PCR products digested by RE will give readily distinguishable patterns when fractionated by agarose or acrylamide gel electrophoresis. *: some bands will appear as doublets.

1.4. Objective

- To confirm QTLs for seedling traits which were previously detected by Khan and Kazmi et al. (2012) on chromosome 6
- To confirm QTLs for seed quality traits which were previously detected on chromosomes 9 by Khan and Kazmi et al. (2012)

2. Methodology

2.1. Plant material

Heterozygous inbred family's (HIFs) were generated from specific heterozygous loci on chromosome 6 and 9 from an $F_{7:8}$ Recombinant Inbred Lines (RIL) used by Kazmi and Khan et al. (2012) and Khan and Kazmi et al. (2012). The RIL population was initially developed from a cross between *Solanum lycopersicum* (cv. Moneymaker) and *Solanum pimpinellifolium* (G1.1554) and various seed quality and seedling growth QTLs were detected across the 12 chromosomes of tomato. Seeds of HIF on chromosome 6 (HIF233) containing 20 progenies were used for seedling traits QTL confirmation. The confidence intervals of seed quality QTLs on chromosome 9 ranges from 54.142 to 105.399 cM for which six HIFs were selected based on segregating molecular markers that are linked to the locus of interest at this specific position (Table 1). Parents were grown alongside 120 lines in total of the six different HIFs. Plants were grown in the greenhouse (Wageningen UR, The Netherlands) under long day conditions (16hr light/ 8hr dark) and approximately 25^0C day and 15^0C night temperature.

Table 1. Position of heterozygous part of the HIFs in the genome, numbers are given in cM. Right columns indicate which markers are used for genotyping

HIFs	Heterozygous		Homozygous		Markers	
	low	high	low	high		
233	98,748	115,208	98,225	116,545	STW1058/1059	STW1070/1071
239	54,933	70,853	54,142	71,113	STW998/999	STW1000/1001
241	78,009	92,585	78,009	94,2	STW1054/1055	STW1056/1057
259	94,46	112,29	94,2	113,627	STW1058/1059	STW1070/1071
266	90,969	116,805	87,271	116,805	STW1056/1057	STW1058/1059
288	78,009	94,46	78,009	94,984	STW1054/1055	STW1056/1057

2.2. Phenotyping

Part I: Seedling traits (HIF233 Chromosome 6)

HIF233 (Chromosome 6) seeds were germinated for subsequent experiments on seedling traits such as, hypocotyl length, fresh shoot weight, fresh root weight, dry shoot weight and dry root weight. HIF233 consists of 20 lines and the parents were included as a control. Hundred randomly picked seeds from each seed lot were imbibed using 50 ml demi water in a plastic tray with double blue filter paper. The trays were closed and covered with a plastic bag to avoid evaporation and stratified at 4^0C for 3 days to break dormancy. Subsequently, the trays were kept in an incubator at 25^0C in the dark for 48 hours to initiate germination.

Figure 4. An example image showing seed trays covered in a plastic bag and kept in an incubator (25^0C) for germination

The first 40 germinated seedlings were transferred to a Copenhagen table and grown under water supply (nutrient less) conditions for 14 consecutive days after time to fifty percent seeds germinated (t_{50}). The seedlings were placed on a wet blue filter paper in randomised block design with four replications per line with 10 seedlings per replication (Figure 5).

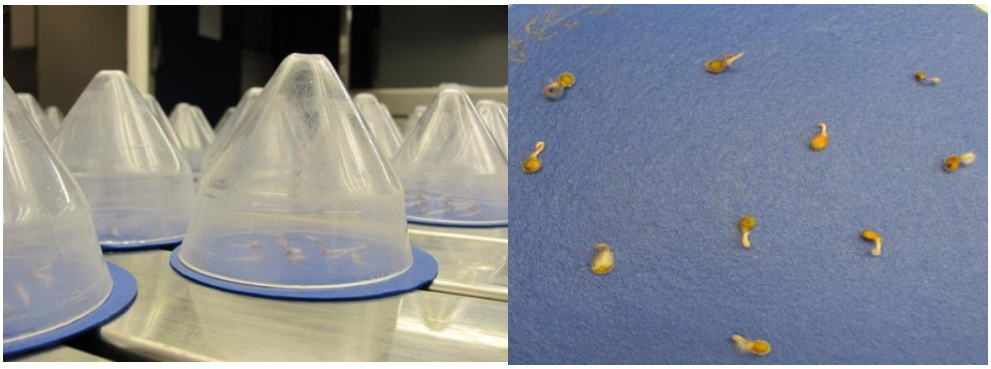


Figure 5. Seedlings grown on a Copenhagen table on a moist blue filter paper with 10 seedlings per replication

Measurements for hypocotyl length, fresh and dry shoot and root weight were taken from HIF233 (Chromosome 6) seedlings grown on a Copenhagen table for 14 days. Hypocotyl length of all 10 seedlings per replication was measured using a ruler. Fresh weight of shoot including hypocotyl and cotyledon was taken. The roots of all 10 seedlings were dissected from shoot and weighed. To measure dry weight, the shoots and roots were dried for 3 days at 80^0C.

Part II: Plant growth and seed quality phenotypes (HIFs chromosome 9)

Plant growth phenotypes

Six HIFs on chromosome 9 and parental lines were scored for plant height, number of trusses and number of fruits per trusses (Figure 6.). The HIFs consist of 120 lines in total and 20 parental lines (10 from each) were grown alongside. Measurements were taken for one month on a weekly basis and the measurement at 90 days after sowing was used for further analysis.

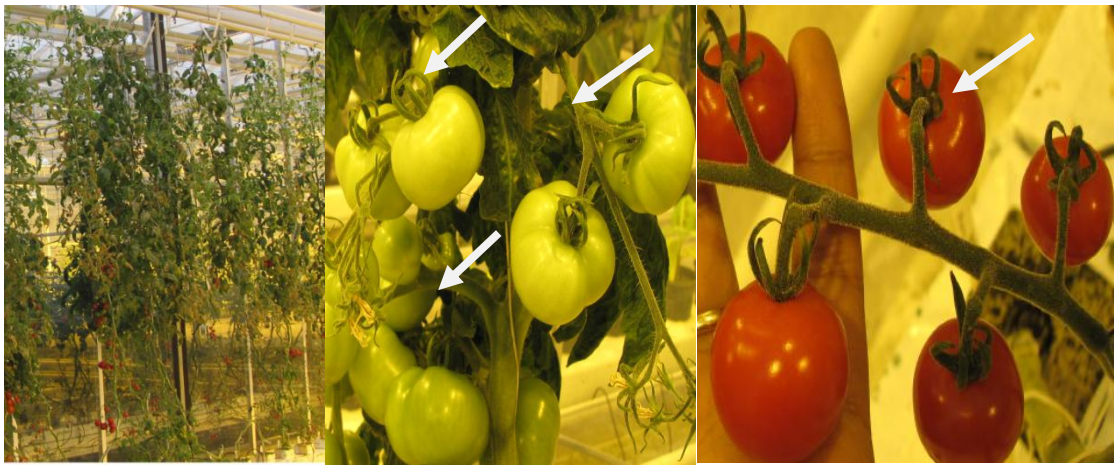


Figure 6. Plant growth traits measured. Plant height: image in the left Number of truss per plant: was scored as image in the middle indicated by arrows. Number of fruits per truss: each fruit per truss was counted as shown by arrow

Seed traits

All the HIFs were evaluated for seed quality traits (seed weight, dry seed area, dry seed length, imbibed seed area and imbibed seed length). For seed traits measurements, on average 120 seeds were randomly taken from the seed lots and weighed on a sensitive balance. Average weight per seed was determined by dividing the total weight by the number of seeds. The weighed seeds were transferred to a tray with a white filter paper. Seeds were separated not to touch each other and images were taken using a camera with fixed settings for seed experiments. Fourteen mL demi water was added to the seed tray to imbibe the seeds. The trays of imbibed seeds were stored in a plastic bag (Figure 4) and kept in an incubator at 25^0C for 18 hours. After imbibition, images were taken to measure for imbibed seed traits. Imbibed seeds were kept in an incubator (25^0C) to initiate germination and number of germinated seeds was scored after 5 days of imbibition.

2.3. Genotyping

DNA extraction

DNA was extracted from the 120 HIF lines for the locus on chromosome 9 and 20 parental lines (10 from each parent). A part of young leaf or undifferentiated floral bulb was harvested and DNA extraction was done following a quick DNA extraction protocol (Cheung et al., 1993) with adjustment for 96 wells plates. Detailed procedures of the protocol can be found in appendix 1.

Marker analysis

Cleaved/cut amplified polymorphic sequence (CAPS) markers were used for genotyping. CAPS that create or eliminate a restriction endonuclease/restriction enzyme at specific positions on Money and Pimp background were developed in the lab prior to this project. Each set of HIFs was genotyped with two marker combinations. The parental lines (Money and pimp) were included in the analysis as a control.

Polymerase chain reaction (PCR) amplification with primers flanking the SNPs was done in the reaction mixture described in table 2.

Table 2. Master mix for PCR reactions

Components	Volume per reaction (µl)
Primer (Forward/Reverse)	0.4
dNTPS	0.3
MgCl	1.5
Buffer B	1.5
Taq. Polymerase	0.3
MQ water	10
DNA	1
Total volume	15

A prepared PCR mix was amplified following the PCR program described in Table 3.

Table 3. PCR program used for amplification

Step	Temperature	Time	
Initial denaturation	95^0C	5 minutes	
Denaturation	95^0C	20 seconds	38 cycles
Annealing	55^0C	30 seconds	
Elongation	72^0C	1.30 minutes	
Elongation	72^0C	7 minutes	
Hold	4^0C	∞	

Five µl of amplified product was electrophoresed on 1.5% agarose gel at 100 volt for 40 minutes to verify the amplification of DNA with the respective primers. Digestion of the

amplified PCR product with proper restriction enzymes for each marker was carried out using the reaction mixture described in Table 4.

Table 4. Restriction enzyme reaction mix

Component	Volume per reaction (µl)
PCR product	5
Buffer 4	2
BSA (optional)	2
Buffer B	2
Restriction enzyme	0.2*
MQ water	x
Total volume	20

*In case of Hind III 1 µl was used. X= amount of MQ water was adjusted to make a final volume of 20 µl

A final volume of 20µl digestion reaction was incubated at 37^0C for six hours and the reaction was stopped by incubation at $65/80^0$C (depending on the restriction enzyme) for 20 minutes in the PCR machine. In order to detect fragment length polymorphisms, the digested reaction mix (10µl) was electrophoresed on a 1.5% agarose gel with gel-red for staining at 80 volt for 80 minutes. Digestion with restriction enzymes was not needed for marker STW1000/1001 thus; genotyping for this marker was scored from the resulting differences in base pairs directly after amplification reaction. The amplified product (4 µl) was loaded on a 2% agarose gel and run at 80 volt for 80 minutes.

Genotyping results were analysed from gel photos based on differences in the expected fragment size after digestion (Table 5) of the amplified PCR product with the respective restriction enzyme. No restriction enzyme digestion was applied for amplification with marker STW1000/1001 because it discriminates genotypes based on a size difference without digestion with restriction enzymes.

Table 5. Marker-restriction enzyme combination and expected fragment size after digestion

Markers	Restriction enzyme	Back-ground	Fragment size (bp)	Fragments (bp/kb) (Money, pimp)	Pimp size (bp)	Money maker size (bp)
STW998/ 999	Nhe1	money	745	240, 500		
STW1000/1001	none	none	-	-	145	177
STW1054/1055	HindIII	Pimp	844	312, 532		
STW1056/1057	xba1	pimp	406	304, 102		
STW1058/1059	HindIII	pimp	691	261, 430		
STW1070/1071	Nhe1	pimp	451	265, 186		

2.4. Data analysis

Scoring and analysis of seed phenotypes such as number of seeds, dry and imbibed seed area and length was carried out in an automated procedure which assists the very laborious and time consuming scoring of seed traits and provides a fast and reliable automated high-throughput scoring. Scoring was done from the images taken from the germination trays by a Nikon D80 camera with Nikon camera control pro software version 2.0. Images were analysed using the open-source image analysis software ImageJ (http://rwbweb.nih.gov/ij/) by using a colour threshold combined with particle analysis. The program automatically counts the amount of pixels per seed and calculates number of objects (seeds), area (mm^2) and length (mm).

Statistical analysis including average, t-test and standard error were conducted to be able to detect QTL association according to single-marker analysis (Collard et al., 2005). Progenies in each HIF were grouped per each segregating marker according to their background (Money, Pimp and Heterozygous). A 2-tailed student's t-test was done to analyse differences in mean values between groups carrying parental allele (Money and pimp) at 5% probability.

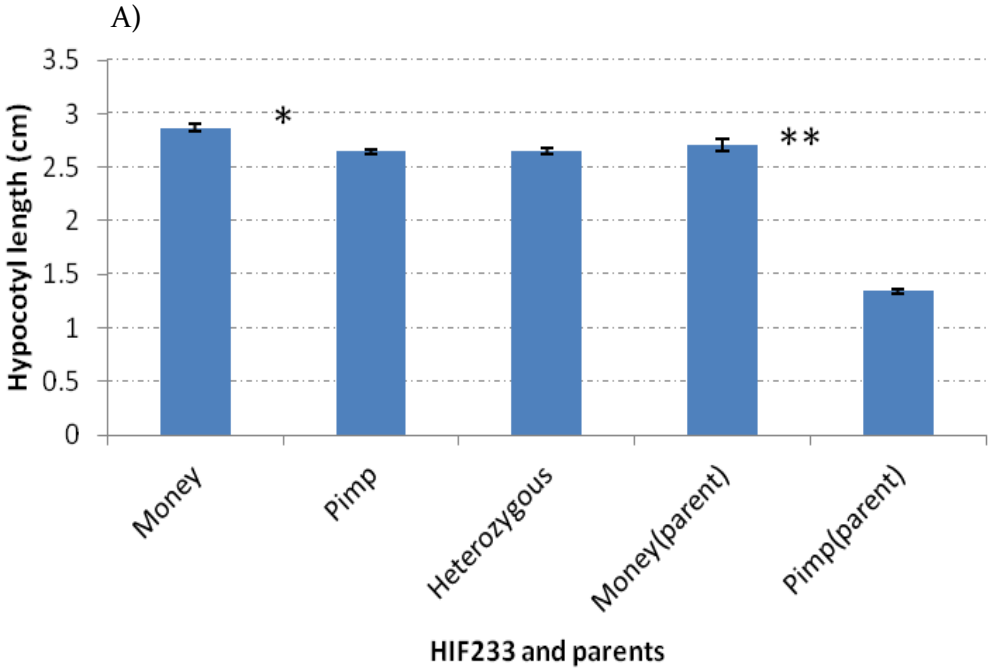
3. Result

Eight traits were scored in this study for the HIFs of chromosome 9, including 5 seed quality traits, such as seed weight (mg per seed), seed area (mm^2), seed length (mm), imbibed seed area (mm^2) and imbibed seed length (mm) and 3 plant growth phenotypes such as plant height, number of trusses per plant and number of fruits per truss (HIF233, HIF239, HIF241, HIF259, HIF266 and HIF288). Furthermore, 5 seedling growth traits such as hypocotyl length, fresh shoot weight, fresh root weight, dry shoot weight and dry root weight were measured from HIFs on chromosome 6 (HIF233). Parental lines (Money and Pimp) were included in all measurements.

3.1. Part one: Seedling QTL on Chromosome 6

HIF 233 lines were grouped into Money, Pimp and Heterozygous background according to the segregating marker (STW1070/1071) at the locus of interest. Significant difference ($p<0.05$) for hypocotyl length was observed between lines of Money and pimp background (Figure 7A). Thus, a statistical evidence for the QTL controlling hypocotyl length is shown. Hypocotyl length of around 2.9 cm on average was measured from Money lines whereas 2.6 cm from lines of pimp and heterozygous background (Figure 7A and Appendix 3) indicating that Pimp allele is dominant at this locus.

A)



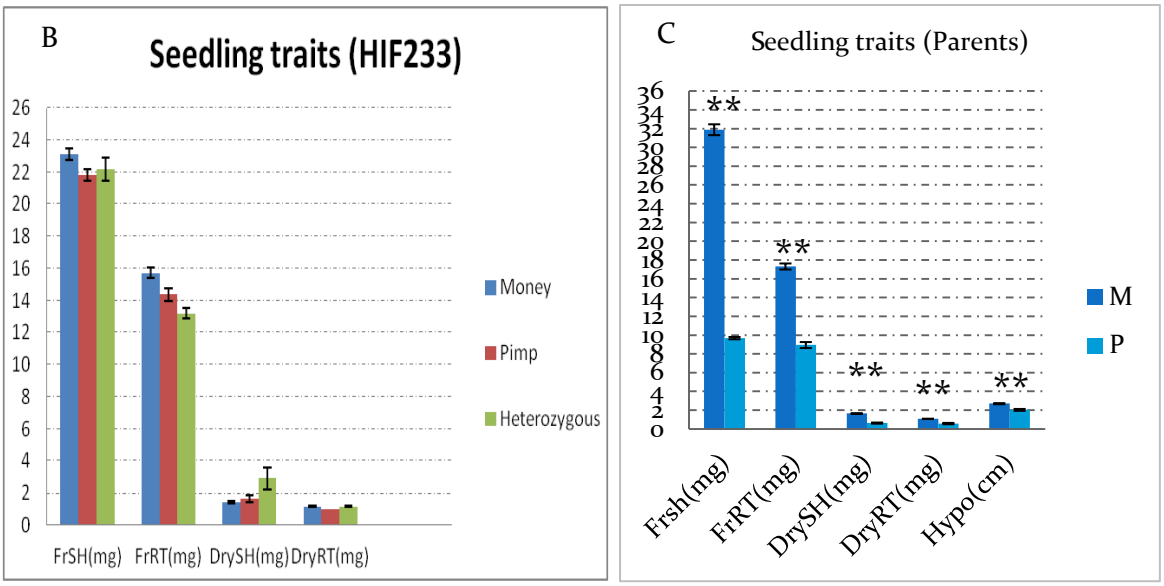


Figure 7. Mean values for hypocotyls length in HIF233 (A) Mean values of seedling traits in HIF233 B) All seedling traits in parental lines C) * = significant difference ($p<0.05$) **= significant difference ($p<0.01$)

HIF233 lines with Money background showed higher values for the tested traits except for dry shoot weight (Figure 7B). The difference for the seedling traits (FrSH, FrRT, DrySH and DryRT) was not statistically significant. Differences between the parents in all the studied seedling traits were highly significant ($p<0.01$). The Money parent was superior in all the measured traits comparing to that of the Pimp parent (Figure 7C). Money seedlings were heavier and longer compared to pimp seedlings.

3.2. Part two: Plant and Seed quality QTLs on Chromosome 9

3.2.1. Plant phenotyping

Plant height

Six HIFs on chromosome 9 of tomato grown in controlled conditions in the greenhouse were evaluated for three plant growth phenotypes such as plant height, number of truss per plant and number of fruits per truss during growing season. The two parents were grown alongside and scored for these phenotypes.

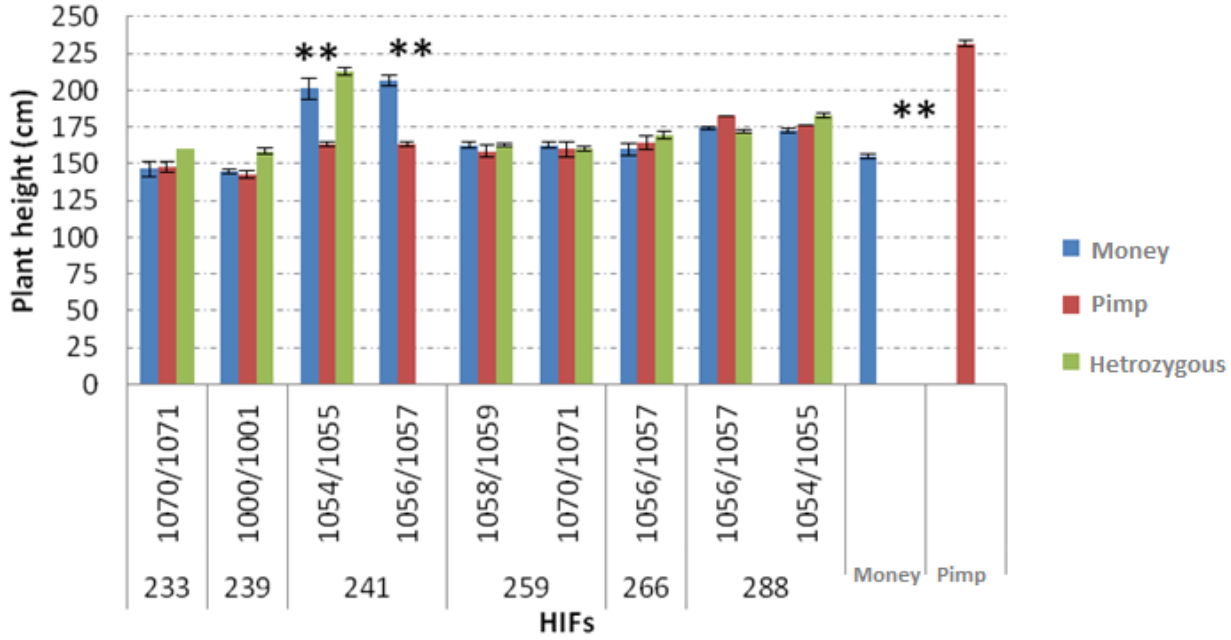


Figure 8. Average plant height (cm) of HIFs on chromosome 9 when clustered per flanking markers at the locus of interest. The data was analysed from measurements taken on 90 days after sowing.The HIFs, parents and markers placed on the x-axis. The error bars represent the standard error of a minimum of 2 biological replicates. **= significant difference ($p<0.01$)

Variation was observed for plant phenotype among the parental lines and the HIFs. Significant difference ($p<0.01$) in plant height was observed between parents and HIF241 for both markers (STW1054/1055 and 1056/1057) (Figure 8). The Pimp parental lines were taller than the Money lines (Figure 8; Appendix 4).

3.2.2. Fruits and Truss

Pimp parental lines produced 35 fruits per truss whereas Money produced on average five fruits per truss (Figure 9; Appendix 5). However, Money lines produced bigger fruits compared to pimp lines.

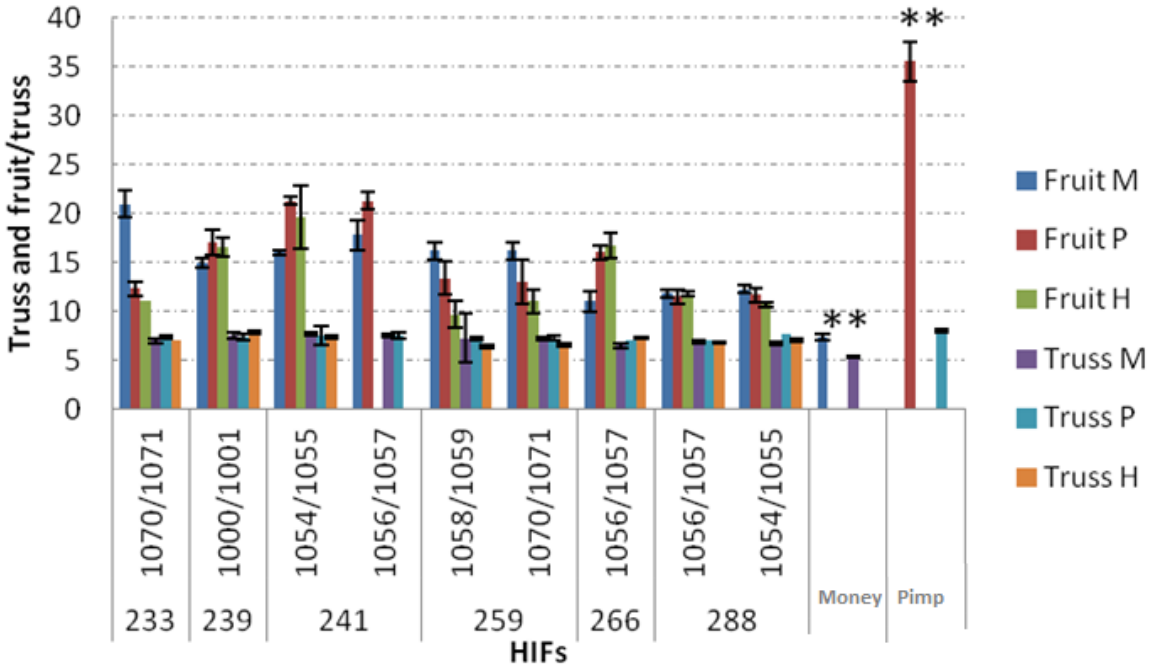


Figure 9. Average number of fruits per truss and number of trusses per plant in HIFs (chromosome 9). The error bars represent the standard error of a minimum of 2 biological replicates. The HIFs, parents and markers placed on the x-axis. **= significant difference ($p<0.01$)

3.2.3. Seed quality QTLs

In total, five seed quality traits such as seed weight (mg per seed), seed area (mm^2), seed length (mm), imbibed seed area (mm^2) and imbibed seed length (mm) were studied from HIFs of chromosome 9. Parents were included in the analysis as a control.

Seed weight

Figure 10 shows seed weight measurement results of the six HIFs and parents. Highly significant differences in seed weight were observed between the parents. The Money parent showed higher seed weight values than pimp parent. Significant difference ($0.01<p<0.05$) in HIF241 and HIF 259 for both markers provide statistical evidence to the presence of QTL. Variation was observed in the remaining HIFs but not at a statistically significant level. HIFs with Money background showed higher values for seed weight when compared with pimp alleles except at the marker position STW1000/1001 in HIF239.

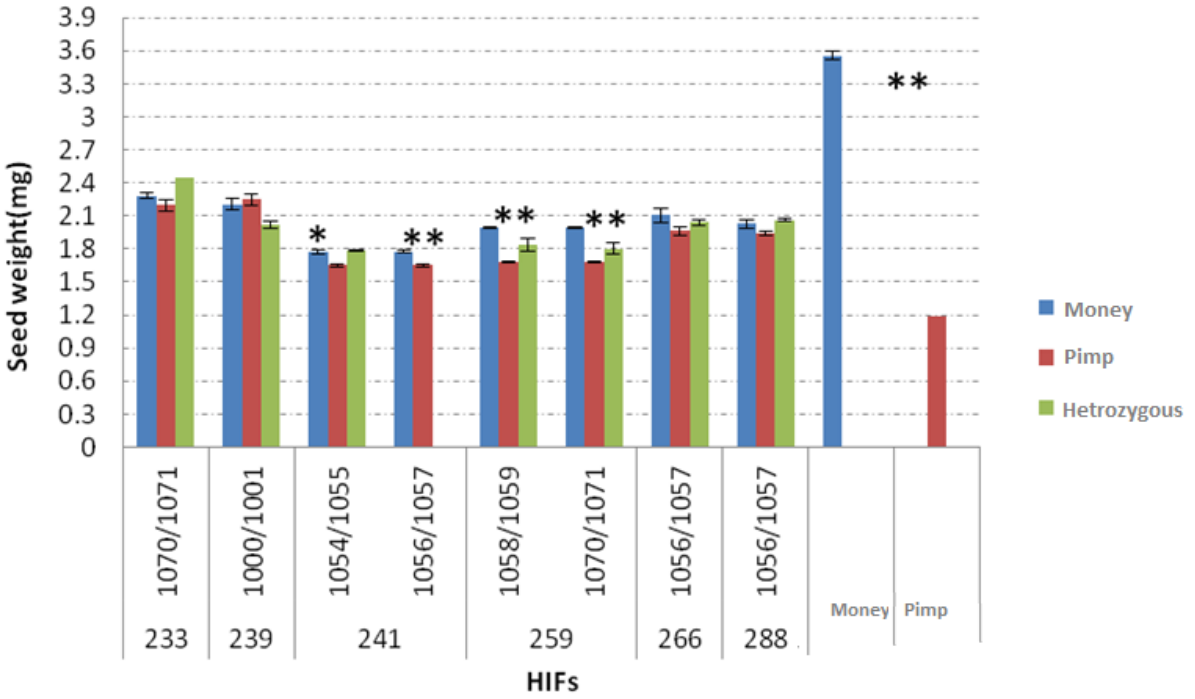


Figure 10. Seed weight (mg per seed) measurement from HIFs on chromosome 9. The error bars represent the standard error of a minimum of 2 biological replicates. The HIFs, parents and markers placed on the x-axis. *= significant difference ($p<0.05$) **= significant difference ($p<0.01$)

Seed area

The Money parental line showed significantly larger dry seed area than pimp seeds ($p<0.01$). Significant difference was observed for seed area in HIF259 for both marker loci and for HIF266 at a single marker locus. In HIFs233, 239, 259 and 266, larger seed area was shown in lines having Money background at the marker positions (Figure 11).

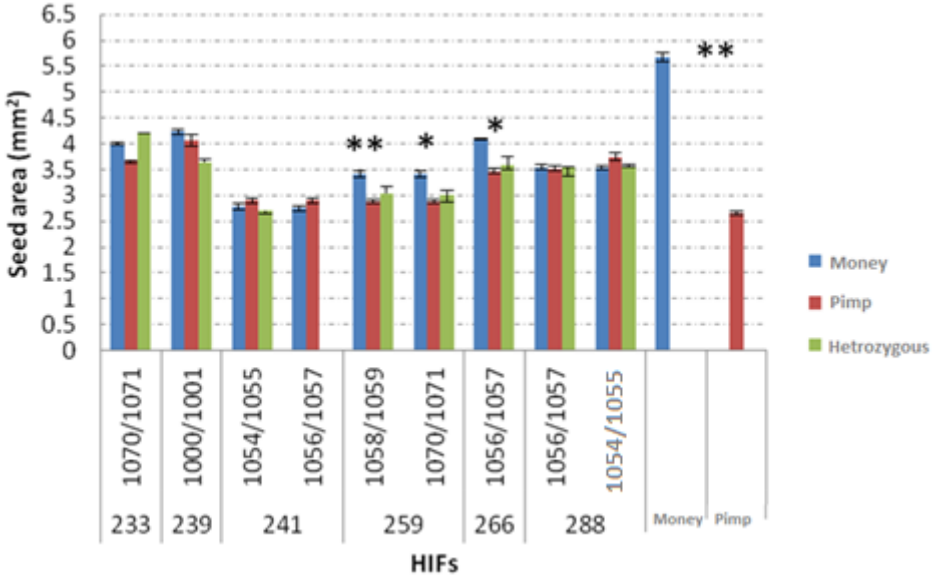


Figure 11. Average seed are (mm^2). The error bars represent the standard error of a minimum of 2 biological replicates. The HIFs, parents and markers placed on the x-axis. * = significant difference ($p<0.05$) ** = significant difference ($p<0.01$). HIFs were grouped per segrigating markers only.

Seed length

Difference in seed length was significant among the parents with higher value for Money than Pimp (Figure 12) ($p<0.01$). The variation in seed length was significant between HIF259 lines of different backgrounds for both markers loci. In the remaining HIFs the difference was not statistically significant.

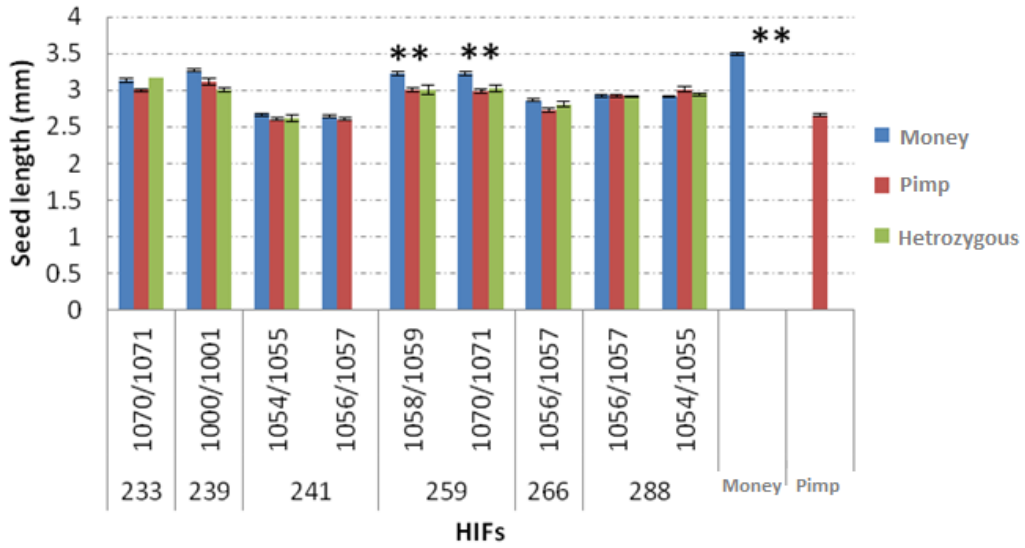


Figure 12. Average dry seed length (mm). The error bars represent the standard error of a minimum of 2 biological replicates. The HIFs, parents and markers placed on the x-axis. ** = significant difference ($p<0.01$)

Imbibed seed area

In addition to dry seed area measurements, seed area from imbibed seeds was taken because in imbibed condition seeds get bigger due to water uptake and it makes the visualisation easier. Imbibed seed area was scored from images taken after 18 hours of imbibition. Highly significant difference in imbibed seed area was observed ($p<0.01$) in HIFs 241 and 259 for the two marker loci (Figure 13). Lines with Moneymaker background showed larger seed area than lines with pimp and heterozygous background with the exception of HIF233 at the marker locus STW1070/1071.

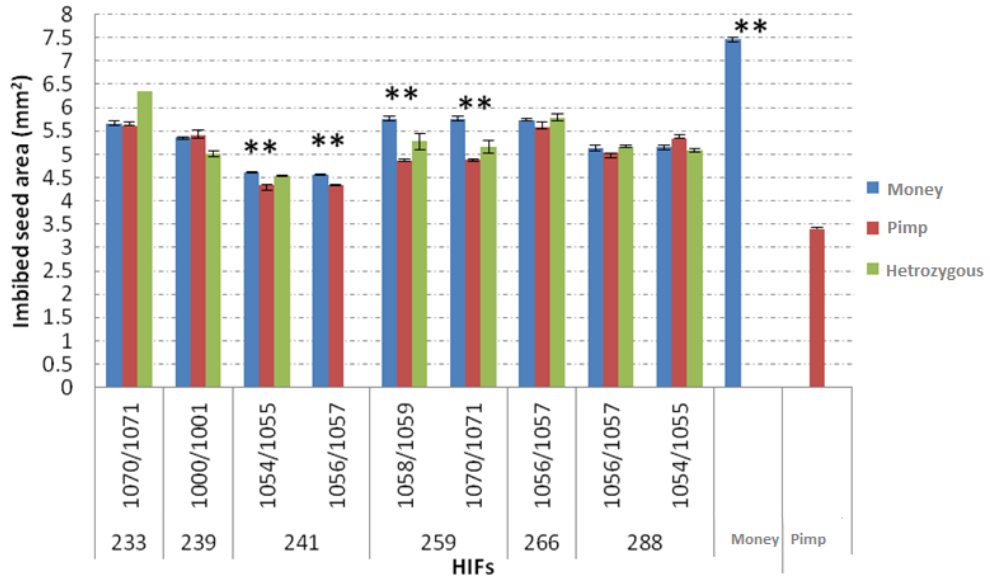


Figure 13. Average imbibed seed area (mm^2). The error bars represent the standard error of a minimum of 2 biological replicates. The HIFs, parents and markers placed on the x-axis. **= significant difference ($p<0.01$)

Imbibed seed length

Figure 14 shows the result of imbibed seed length measurements analysed from images taken after 18 hours of imbibition. Significant difference for imbibed seed length was observed between parents ($p<0.01$). HIFs 241 and 259 showed significant differences for all marker loci for which the HIFs were selected.

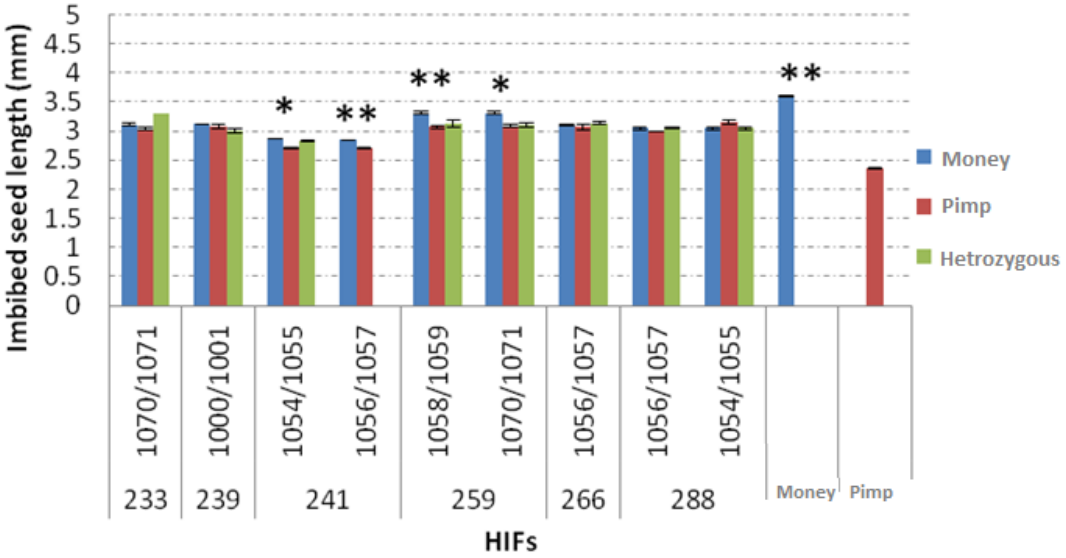


Figure 14. Average imbibed seed length (cm). The error bars represent the standard error of minimum of 2 biological replicates. The HIFs, parents and markers placed on the x-axis. *= significant difference ($p<0.05$)
**= significant difference ($p<0.01$)

Germination

Number of seeds germinated per line was scored 5 days after imbibition to check if there is a variation on germination rate in the HIFs. Higher germination percentages were observed for the pimp compared to the Money parent. In the HIFs, there was no significant difference for rate of germination when lines were clustered per segregating marker locus (Figure 15).

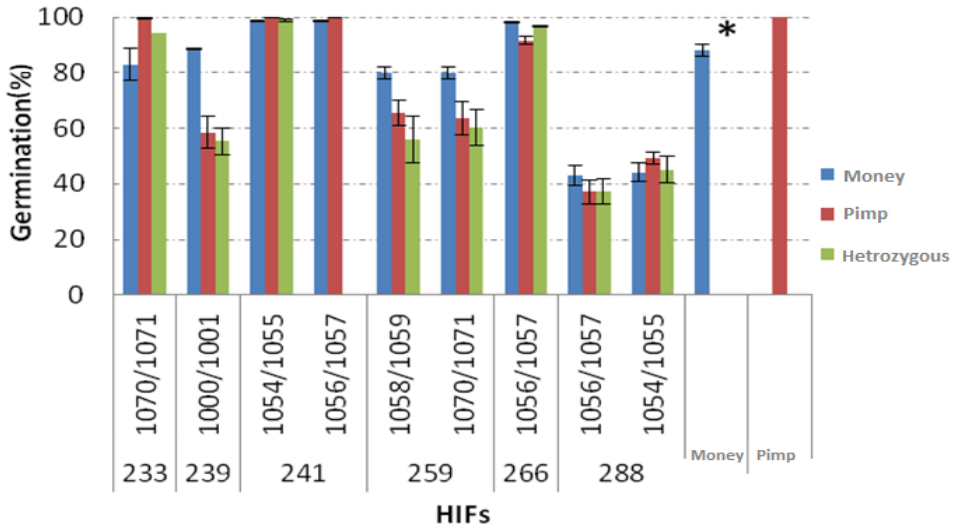


Figure 15. Average germination percentage of HIFs when clustered per marker locus. The HIFs, parents and markers placed on the x-axis. * = significant difference ($p<0.05$). The error bars represent the standard error of minimum of 2 biological replicates.

4. Discussion and conclusion

This minor thesis project aimed to confirm previously detected QTLs governing seedling traits on chromosome 6 and QTLs responsible for seed quality traits on chromosome 9 using a HIF approach. Seeds of HIFs on chromosome 6 were grown for 14 days after germination and evaluated for seedling traits. HIFs on chromosome 9 were grown in the greenhouse under controlled condition and the plants were evaluated for plant growth traits. Seeds of these plants were used for measuring and analysing seed quality traits.

4.1. QTLs on HIF chromosome 6

QTL for hypocotyl length on chromosome 6 which was detected by Khan and Kazmi et al. (2012) was selected for confirmation. In this study we found significant difference ($p<0.05$) for hypocotyl length between HIF233 lines with Money and Pimp background. This result confirm the QTL for this trait. Regarding the fresh root weight QTL, variation among the lines carrying different parental alleles at the same locus was not statistically significant. This result supports the previous study (Khan et al., 2012) as there was no QTL detected for fresh root weight at the locus of interest and proves our hypothesis that the locus is not linked to this trait. For the remaining seedling traits (Fresh shoot weight, dry shoot weight and dry root weight), QTLs were not detected in the previous study as well as in the current study at the selected marker position.

4.2. QTLs on chromosome 9

Plant phenotype QTL

HIFs generated at a specific position on chromosome 9 were genotyped, evaluated for plant growth traits and seed quality traits. Plant growth traits such as plant height, number of truss per plant and number of fruits per each truss were measured to evaluate whether the specific locus of the HIFs are linked to these plant phenotypes or not. Apparently we found significant variation in plant height among HIF241 for both marker loci providing a statistical evidence for a QTL at the marker position on chromosome 9 with higher value for lines with Money background. In the other HIFs no evidence was shown for the genetic locus to control the plant height as well as the remaining plant phenotypes (number of fruits and truss).

Seed quality QTLs

Seed quality traits such as seed weight, dry seed area, dry seed length, imbibed seed area and imbibed seed length were evaluated in this study in order to confirm the QTLs detected in the study by Khan and Kazmi et al. (2012). Germination percentage of the seed batch was also scored alongside to check the quality of seeds and make sure that there is no initial variation in the seed batch that can affect the traits of interest. The results confirmed that the germination rate of the HIFs grouped per marker locus do not vary significantly.

Seed weight

The seed weight of progenies in each HIF was measured to determine the association between seed weight and the segregating markers. In HIF259 segregating for markers STW1058/1059 and STW1070/1071, lines carrying the allele derived from parent Money showed significantly higher seed weight than lines carrying the allele from Pimp. The loci covers 94.46-113.627 cM on Chromosome 9. Significantly higher seed weight was also shown in HIF241 segregating for markers STW1054/1055 and 1056/1057 (78-94.2 cM) for the lines carrying Money allele. Interestingly we confirmed seed weight QTL at specific marker loci of STW1058/1059, STW1070/1071, STW1054/1055 and STW1056/1057 on chromosome 9. This provides an evidence for the studied loci to control seed weight. The Money parent lines and progenies of HIFs carrying the allele derived from Money showed higher value for seed weight than pimp. HIF266 which has a heterozygous part at position 90 to 116 cM seemed to have higher values for seed weight in lines carrying Money compared to Pimp background at that locus but the difference was not significant ($p<0.05$). This could be due to fewer replicates (2) to reveal statistical significance or a possible epistasis effect.

Seed size

Seed size was evaluated by measuring seed area and length at dry and imbibed condition. Seed area measurements taken from dry seeds and imbibed seeds of HIFs and parental lines showed the variation among progenies carrying different parental alleles. The parents differ significantly for both dry and imbibed seed area with higher value for Money lines. Higher values for dry and imbibed seed area were observed in progenies of HIFs carrying alleles derived from the Money parent. For HIF259 segregating for markers STW1058/1059 and STW1070/1071 significant difference for both dry and imbibed seed area was shown between lines carrying different parental alleles. This indicates that this locus has an association with

seed size. The significant difference found in both traits supports the consistency of the QTL. HIF241 segregating for markers STW1054/1055 and STW1056/1057 showed significant difference for imbibed seed area but not for dry seed area measurements. This could be due to variation in water uptake between the lines. HIF241 progenies carrying pimp allel at both marker loci showed larger dry seed area but smaller imbibed seed area when compared to progenies carrying Money allele. This could be due to seeds of Money background take up more water than Pimp background seeds and there could be variation in seed thickness and this might be relevant to study more in detail.

Dry and imbibed seed length measurements showed highly significant differences among the parental lines with higher seed length values for Money. In all HIFs, longer seeds were associated with the progenies possessing alleles derived from the Money parent. The significant differences in both dry and imbibed seed measurements shown in HIF259 segregating for markers STW1058/1059 and STW1070/1071 and for imbibed seed area in HIF 266 for marker locus STW1056/1057 confirmed the QTLs for these traits.

5. Further study

Number of progenies were limited per HIF when they were grouped according to their background , resulting in lower statistical power. Therefore it would help to include more lines for each HIF in the analysis.

Discoloration (brown/blackish) of seeds was observed mainly in HIFs (HIF266) with money maker background and also in the seeds of Moneymaker parental lines. These discolored seeds failed to germinate. Thus it would be interesting to further study the genetic mechanism underlying this phenotypes.

With the lines that confirm the QTLs, fine mapping can be a follow up approach towards finding the genes that are involved in the studied traits and studying the underlying mechanism.

Reference

Alonso-Blanco C, Aarts MG, Bentsink L, Keurentjes JJ, Reymond M, Vreugdenhil D and Koornneef M (2009). What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell 21:1877-1896

Bewley J D and Black M (1994) SEEDS: Physiology of Development and Germination. Second edition. ISBN 0-306-44747-9 Plenum press, New York

Cheung WY, Hubert N, Landry BS (1993) A simple and rapid DNA microextraction method for plant, animal, and insect suitable for RAPD and PCR analyses. PCR Methods and Applications 3:69-70

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts Euphytica 142: 169-196

Dognlar S, Frary A, Tanksley SD (2000) The genetic basis of seed-weight variation: tomato as a model system. Theor Appl Genet 100:1267-1273

Ellis RH. (1992) Seed and seedling vigour in relation to crop growth and yield. Plant growth regulation 11:249-255

Hilhorst HW and Toorop, PE. (1997) Review on dormancy, germinability, and germination in crop and weed seeds. Advances in Agronomy 61:111-165

Joosen RVL, Kodde J, Willems LAJ, Ligterink W, van der Plas LHW and Hilhorst HWM (2010) GERMINATOR: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination. The Plant Journal doi: 10.1111/j.1365-313X.2009.04116.x

Joosen RV, Arends D, Willems LA, Ligterink W, Jasen RC and Hilhorst HW. (2012) Visualizing the Genetic Landscape of Arabidopsis Seed Performance. Plant physiology 158:570-589

Kazmi RH, Khan N, Willems LAJ, van Heusden AW, Ligterink W and Hilhorst HWM (2012) Cpmplex genetics controls natural variation among seed quality phenotypes in a recombinant inbred population of an interspecific cross between *Solanum lycopersicum* X *Solanum pimpinellifolium*. Plant cell and environment 35:929-951

Khan N, Kazmi RH, Willems LA, Heusden AW, Ligterrink W, Hilhorst HW (2012) Exploring the natural variation for seedling traits and their link with seed dimensions in Tomato. PLOS ONE 7(8):e43991

Koieczny A and Ausubel FM (1993) A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. The plant journal 4(2):403-410

Koornneef M, Bentsink L and Hilhorst HW (2002) Seed dormancy and germination. Current opinion in Plant Biology 5:33-36

Monforte AJ and Tanksley SD (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the *Lycopersicon hirsutum* genome in a *L. esculentum* genetic background: A tool for gene mapping and gene discovery. Genome 43: 803-813

Nieuwhof M, Garretsen F, van Oeveren JC (1989) Maternal and genetic effects on seed weight of tomato, and effects of seed weight on growth of genotypes of tomato (*Lycopersicon esculentum* Mill). Plant Breed 102:248–254

Pet G and Garretsen F (1983) Genetical and environmental factors influencing seed size of tomato (*Lycopersicon esculentum* Mill.) and the effect of seed size on growth and development of tomato plants. Euphytica 32:711-718

Thiel T, Kota R, Grosse I, Stein N and Graner A (2004) SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. Nucleic Acids Research Vol. 32, No. 1 e5 DOI: 10.1093/nar/gnh006

Tuinstra MR, Ejeta G and Goldsbrough PB (1997) Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. Theor Appl Genet 95:1005-1011

Appendices

Appendix 1: DNA extraction protocol

Things to prepare before starting:

- Ice
- Clean and sharp forceps
- Clean 1ml tubes and put one stainless tube inside

Procedures:

Step 1. Collect small or part of leaf in 1ml tube

Step 2. Add 150 µl extraction buffer and grind on a shaker at 30Hz for 1 minute.

Step 3. Keep the grinded sample at 60^0C for one hour

Step 4. Spin for 10 minutes at 2500 rpm

Step 5. Transfer 75 µl supernatant to new tube or microtiter plate

Step 6. Add 75 µl iso-propanol and 30 µl 10M NH_4Ac

Step 7. Leave the mixture at room temperature for 15 minutes

Step 8. Spin for 20 minutes at 2400 rpm

Step 9. Remove the liquid with capillary and leave the pellet

Step 10. Add 70% ethanol to wash the pellet

Step 11. Spin for 5 minutes at 2400 rpm, remove the ethanol with capillary and dry the pellet for 5 minutes at room temperature

Step 12. Dissolve the pellet in 50 µl with MilliQ water

Marker information:

Appendix 2. Markers, their position on the chromosome primer sequences

Markers	Position (bp)	Map distance (cM)	Forward primer	Reverse primer
STW998/ 999	58369026	54,142	ACAAGAGGAGCTGGATAC	AGGGGCAAAGGGAGAAAA
STW1000/1001	60987757	65,002	TTGGCCAGCATTGTTTATG	GACCCTGAGATGTACTAGTT
STW1054/1055	63638559	82,424	GTTGGGAGGTTTTTGAATTG	GTACTTGGTCGGGAAATG
STW1056/1057	64993126	92,585	AAAAATGTAATCTGCAGG	TGCATAGGAGTTTAAGGT
STW1058/1059	65492233	98,225	ATAAAAGAGAGGTCGGGG	GAAAAGGAGTGATATCAAGGG
STW1070/1071	66097878	105,399	GTGAGTTTGATTTTGCACC	AATAACCTGCTGTGGAGA

Seedling measurement

Appendix 3. Mean values of seedling traits measured from HIF 233 (chromosome 6)

		Money	Pimp	Hetero	p-value§
Hypo (cm)		2.87	2.64	2.65	0.0388
	SE	0.08	0.05	0.06	
FrSH (gm)		23.13	21.82	22.19	0.2613
	SE	0.76	0.73	1.40	
FrRT (gm)		15.70	14.38	13.20	0.24339
	SE	0.64	0.80	0.63	
DrSH (gm)		1.43	1.66	2.93	0.61131
	SE	0.09	0.41	1.37	
DrRT (gm)		1.15	0.99	1.17	0.16827
	SE	0.09	0.03	0.09	

§ Statistical probability test is performed between HIF lines of Moneymaker and Pimp background excluding heterozygous lines. SE= standard error

Plant phenotypes (HIF Chromosome 9):

Plant height

Appendix 4. Mean plant height (cm). HIFs are clustered per marker according to their background

					p-value
		M	P	H	
parents	M	154.83			9.45E-08
	P		231.40		
233	1070/1071	146.33	147.67	160.00	0.94
239	1000/1001	145.00	143.00	158.43	0.79
241	1054/1055	201.00	163.50	212.67	0.007
	1056/1057	206.83	163.50	-	0.001
259	1058/1059	162.80	158.40	162.67	0.64
	1070/1071	162.80	159.75	160.25	0.77
266	1056/1057	160.00	164.33	169.50	0.77
288	1056/1057	174.17	182.50	172.00	0.16
	1054/1055	172.88	176.00	183.00	0.46

Fruits and Truss

Appendix 5. Mean number of truss and fruits per truss. HIFs are clustered according to their background per marker.

	fruit/truss				p-value	Truss			p-value
		M	P	H		M	P	H	
Parents	M	7.33			7E-05	5.33			4.89E-05
	P		35.60				8.00		
233	1070/1071	21.00	12.33	11.00	0.07	7.00	7.33	7.00	0.68
239	1000/1001	15.00	17.00	16.57	0.61	7.50	7.33	7.86	0.87
241	1054/1055	16.00	21.33	19.67	0.10	7.67	7.50	7.33	0.87
	1056/1057	17.83	21.33		0.35	7.50	7.50		1.00
259	1058/1059	16.20	13.40	9.67	0.49	7.20	7.20	6.33	1.00
	1070/1071	16.20	13.00	11.00	0.49	7.20	7.25	6.50	0.94
266	1056/1057	11.00	16.00	16.75	0.14	6.50	7.00	7.25	0.27
288	1056/1057	11.83	11.50	11.75	0.85	6.83	7.00	6.75	0.78
	1054/1055	12.25	11.67	10.67	0.61	6.75	7.67	7.00	0.08