Plasticity of maternal environment dependent expression-QTLs of tomato seeds

Mark G. Sterken[1,#], Harm Nijveen[2,3], Martijn van Zanten[4], Jose M. Jiménez-Gómez[5,6], Nafiseh Geshnizjani[2], Leo A. J. Willems[2], Juriaan Rienstra[2], Henk W. M. Hilhorst[2], Wilco Ligterink[2,*], Basten L. Snoek[1,7,#,*]

[1] Laboratory of Nematology, Wageningen University, 6708 PB Wageningen, The Netherlands
[2] Wageningen Seed Lab, Laboratory of Plant Physiology, Wageningen University, 6708 PB Wageningen, The Netherlands
[3] Laboratory of Bioinformatics, Wageningen University, 6708 PB Wageningen, The Netherlands
[4] Molecular Plant Physiology, Institute of Environmental Biology, Utrecht University, 3584 CH Utrecht, The Netherlands
[5] Department of Plant Breeding and Genetics, Max Planck Institute for Plant Breeding Research, Cologne, Germany
[6] Institut Jean-Pierre Bourgin, INRAE, AgroParisTech, Université Paris-Saclay, 78000, Versailles, France
[7] Theoretical Biology and Bioinformatics, Institute of Biodynamics and biocomplexity, Utrecht University, 3584 CH Utrecht, The Netherlands
[#] Equal contribution
[*] Corresponding authors

e-mail and ORCID:
mark.sterken@wur.nl; 0000-0001-7119-6213
harm.nijveen@wur.nl; 0000-0002-9167-4945
m.vanzanten@uu.nl; 0000-0002-2810-7374
Jose.Jimenez-Gomez@inrae.fr; 0000-0002-5033-7192
nafiseh.geshnizjani@gmail.com;
leo.willems@wur.nl; 0000-0002-9502-7745
juriaan.rienstra@wur.nl; 0000-0003-4870-8078
henk.hilhorst@wur.nl; 0000-0002-6743-583X
wilcoligterink@gmail.com; 0000-0002-0228-169X
l.b.snoek@uu.nl; 0000-0001-5321-2996

Key words: Tomato, eQTLs, High phosphate, Low nitrogen, Maternal environment, Seeds

Abstract

Seeds are essential for plant reproduction, survival, and dispersal. Germination ability and successful establishment of young seedlings strongly depends on seed quality and on environmental factors such as nutrient availability. In tomato (*Solanum lycopersicum*) and many other species, seed quality and seedling establishment characteristics are determined by genetic variation, as well as the maternal environment in which the seeds develop and mature. The genetic contribution to variation in seed and seedling quality traits and environmental responsiveness can be estimated at transcriptome level in the dry seed by mapping genomic loci that affect gene expression (expression QTLs) in contrasting maternal environments.

In this study, we applied RNA-sequencing to measure gene expression of seeds of a tomato RIL population derived from a cross between *S. lycopersicum* (cv. Moneymaker) and *S. pimpinellifolium* (G1.1554). The seeds matured on plants cultivated under different nutritional environments; i.e. on high phosphorus or low nitrogen. The obtained SNPs were subsequently used to construct a high-density genetic map. We show how the genetic landscape of plasticity in gene regulation in dry seeds is affected by the maternal nutrient environment. The combined information on natural genetic variation mediating (variation in) responsiveness to the environment may contribute to knowledge-based breeding programs aiming to develop crop cultivars that are resilient to stressful environments.

50 **Introduction**

51 Seeds are essential for reproduction and dispersal of plants and function as survival structures to
52 overcome harsh environmental conditions unfavourable for plant growth. Well-timed development
53 and ripening of seeds, to ensure optimal seed performance and the ability to germinate in a
54 permissive environment, are therefore essential for plant fitness. Successful germination strongly
55 depends on seed performance, which is affected by environmental conditions, such as temperature,
56 water availability, light conditions, and the nutrient status that the maternal plant experienced [1-
57 5]. More specifically, seed performance/germination in species such as tomato and the model plant
58 *Arabidopsis thaliana* is determined during seed development and maturation, and depends on
59 temperature [6-8], photoperiod [9, 10], nutrient composition and levels [8, 11, 12]. Seed quality,
60 germination and seedling establishment traits also have strong genetic determinants and (natural)
61 genetic variation in quality traits, including Quantitative Trait Loci (QTLs), have been reported [8,
62 12-17].

63 Phosphate and nitrate are essential plant nutrients with profound effects on plant growth
64 [18, 19] and seed performance/germination traits [8, 12, 15, 20]. In Arabidopsis it has been shown
65 that seeds produced by plants fertilized with higher-than-normal levels of phosphate showed
66 increased germination rates under stressful conditions [8]. Nitrate is known to have a strong effect
67 on seed germination and seed dormancy in multiple plant species [21], with high concentrations of
68 nitrate supplied to the mother plant leading to lower dormancy of the seeds [11]. This is attributed
69 to nitrogen effects on the gibberellin/abscisic acid (GA/ABA) balance in the seeds; with higher
70 endogenous nitrate levels resulting in lower ABA levels in seeds and hence, shallower dormancy
71 [22]. In Arabidopsis, altered nitrate levels experienced by the mother plant also has a substantial
72 effect on the levels of multiple metabolites and transcripts in the seeds, with a notable reduction in
73 nitrogen metabolism-related metabolites and genes [23].

74 Tomato (*Solanum lycopersicum*) is one of the most important vegetable crops worldwide
75 and is a model organism for research on fruit-bearing crops [24-27]. However, in the process of
76 domestication, breeding selection and propagation, a substantial fraction of the genetic variation of
77 the founder's germplasms has been lost [26-28]. Moreover, due to a focus on fruit quality,
78 resistance and yield traits, other desirable traits that have not been directly selected for have been
79 lost over time in modern varieties. This includes several seed quality traits [28-32]. Trait variation
80 loss can be restored by including wild cultivars/ancestors of modern commercial tomato such as
81 *Solanum pimpinellifolium,* that represent a rich source of genetic variation, in breeding programs
82 and in studies on tomato (quantitative) genetics [26, 33-37]. For instance, wild cultivars have been
83 used in genetic screens and genome wide association studies (GWAS) to discover genomic loci
84 and genes involved in variation in metabolic traits [38-42], insect resistance [43], floral meristem
85 identity [31], trichome formation [44], and fruit shape and size [28, 34, 45]. In addition to GWAS,
86 Recombinant Inbred Line (RIL) populations, derived from experimental crossing between *S.*
87 *lycopersicum* and *S. pimpinellifolium,* are frequently used to uncover the effect of genetic variation
88 on tomato traits [46-52], including various seed quality traits [12, 14, 15, 17, 53, 54].
89 The introduction and improved feasibility of diverse ~omics techniques have accelerated
90 studies into the molecular mechanisms underlying natural variation in tomato traits in the past two
91 decades [55]. In particular, advances in transcriptomics techniques such as microarray analysis and
92 later RNA-sequencing, have proven useful in this context, by enabling e.g. GWAS studies.
93 Moreover, measuring gene expression in RILs has enabled expression-QTL (eQTL) analysis as a
94 powerful tool to detect gene regulatory loci [56-61]. Combining the wealth of information obtained
95 by mapping eQTLs enables (re)construction of regulatory networks underlying plant traits [56, 60,
96 62]. In addition, comparison of eQTL profiles from multiple environments may aid our
97 understanding of how genetic variation shapes the effects the environment has on the appearance

98 of phenotypes [57, 63, 64]. In plant (Arabidopsis) and worm (*Caenorhabditis elegans*) model
99 systems it has been shown that especially *trans*- eQTLs are dynamic and can be highly specific for
100 a certain environment [57, 63-68].

101 Although seed quality and seedling establishment characteristics are determined by both
102 genetic variation and the maternal environment in which the seeds develop and mature [8, 12, 15],
103 it is currently unknown if the maternal environment causes a perturbated eQTL landscape in the
104 progeny seeds and how the nutrient environment of the mother plant affects these landscapes. We
105 therefore followed an RNA-seq approach and quantified natural variation in mRNA levels in the
106 dry seeds of a tomato RIL population from a cross derived from *S. lycopersicum* (cv. Moneymaker)
107 and *S. pimpinellifolium* (G1.1554) parents [17, 49], that were cultivated either in a low nitrogen or
108 a high phosphorus environment. In this work we first present a high-density RNA-seq-derived
109 genetic map of tomato and subsequently we demonstrate how the genetic landscape of gene
110 regulation of tomato dry seeds is affected by the nutritional environment of the mother plant.
111 Altogether, our detailed analysis of the genetic underpinning of plasticity in gene expression as
112 responsiveness to the maternal environment, attributed to the progeny seeds, may contribute to
113 knowledge-based breeding programs aiming to develop crop cultivars that are resilient to stressful
114 environments, including production of high-quality seeds under sub-optimal environmental
115 conditions.
116
117

118 **Results**

119 *An RNA-seq-derived genetic map of tomato*

120 We performed an RNA-sequencing experiment to uncover the interplay between genetic variation,
121 the nutritional status of the maternal environment and mRNA abundances in progeny tomato seeds.
122 The used seeds were derived from tomato RIL plants of a cross between *S. lycopersicum* (cv.
123 Moneymaker; MM) and *S. pimpinellifolium* (G1.1554; PI) [17, 49] and their parental lines. All
124 maternal plants were pre-cultivated on standard nutrient conditions and upon flowering transferred
125 to either low nitrogen (LN) or high phosphate (HP) nutrition (~100 RILs in total, ~50 RILs in each
126 environment) [15].

127 In addition to estimating expression differences among individuals, RNA-seq reads allowed for the
128 identification of single nucleotide polymorphisms (SNPs) in transcribed genes of the parental lines
129 and the RILs. These SNPs were subsequently used to construct high density genetic and physical
130 maps of the RIL population, to facilitate QTL and eQTL mappings [13, 69]. In total, we detected
131 43,188 consistent SNPs between the parental lines. These SNPs were subsequently used to
132 reconstruct the genotypes (*i.e.* determine the crossover locations) of the RILs in high detail (**Figure**
133 **1A**). Across our RIL set, a balanced distribution of the parental alleles was observed genome-wide,
134 with the notable exception of chromosome 2, which had a substantial higher frequency of PI alleles
135 (**Figure 1B**). Overall, 2,847 recombination (crossover) events were detected across the RIL
136 population. As expected, the crossovers were found almost exclusively in euchromatic regions of
137 the chromosomes, causing severe distortion between the physical and genetic maps, as described
138 before [70](**Figure 1C**). On average, two recombination events were detected per RIL per
139 chromosome. Altogether, the population size and recombination events provided 4,515 unique
140 genetic markers and 4,568 distinguishable genomic loci/bins suitable for mapping, improving the
141 previously available map [71] (**Supplementary table 1**). The detected loci had a size-range from

60 Mb to 1.7 Kb, with an average locus size of 180 Kb and a median of 11 Kb (**Supplementary table 2**). Given the high local recombination frequency, relatively small loci were overrepresented towards the chromosome tips (**Figure 1C**). Together, our dataset enables precise mapping of QTLs and eQTLs, especially towards the tips of the chromosomes.

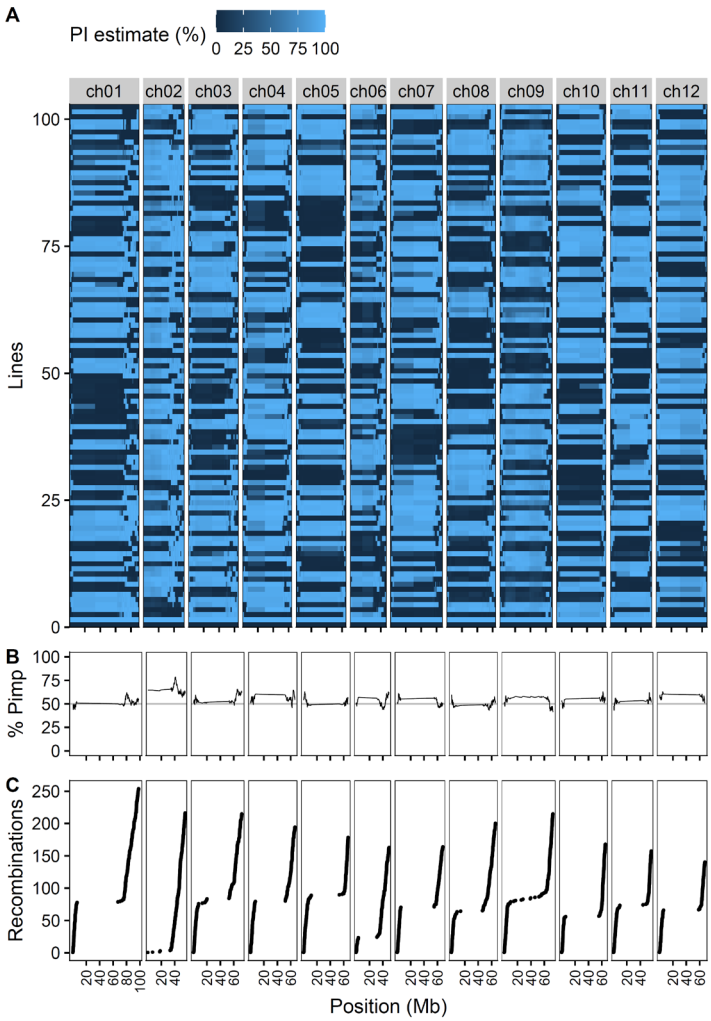


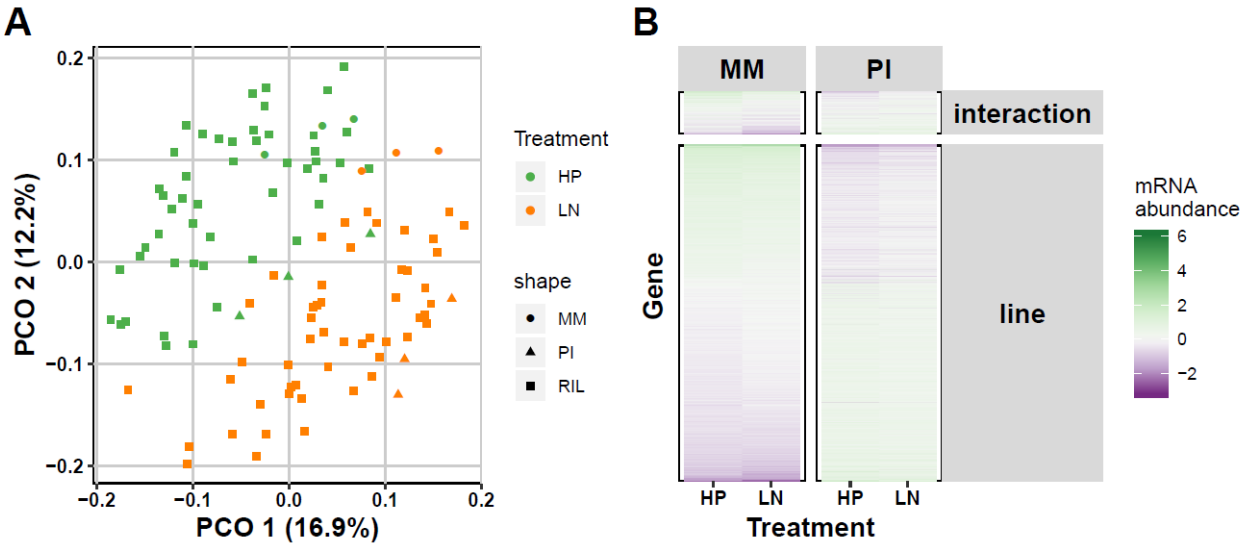
Figure 1: (**A**) Genetic map showing the genotype of the RILs and parental lines per chromosome (ch01 to ch12). Dark blue indicates MM (bottom horizontal line, with line number 0), light blue indicates PI (horizontal line above MM, with line number 100). Shades between dark and light blue visualize the certainty of the estimate that a locus corresponds to either MM or PI, depending on the SNPs identified (see legend above the panel; PI estimate). (**B**) Allele frequency (percentage) of *S. pimpinellifolium* (PI) alleles for each marker across the chromosomes, considering all RILs in the population. (**C**) Cumulative number of recombination events per chromosome for the whole population. Chromosome numbers are indicated above panels **A**, position on the chromosomes (in Mb) is shown on the x-axis below panel **C**.

157 *The maternal nutrient environment affects mRNA abundances in seeds.*

158 Next, we compared mRNA abundances in all HP-treated lines (RILs and parental lines) with the
159 mRNA abundances in LN-treated lines, to identify genes contributing to differences between the
160 two environments. Principal Component Analysis (PCA) demonstrated the presence of a
161 substantial effect of the maternal nutrient environment on transcript levels in seeds (**Figure 2A**). A
162 linear model was used to identify which mRNAs were differentially expressed between the two
163 maternal environments. A multiple-testing correction was applied and differential expression of
164 2,871 mRNAs (out of 14,772 detected mRNAs) was found (Bonferroni corrected p-value < 0.05)
165 to depend on the nutritional conditions the mother plant experienced during the seed maturation
166 phase (i.e. LN or HP) (**Supplementary table 3**). Of these 2,871 mRNAs, 922 were more abundant
167 in seeds developed and ripened in HP conditions compared to LN, and 1,949 mRNAs were
168 significantly more abundant in LN conditions compared to HP. The mRNAs of genes that were
169 more abundant after LN treatment were among others enriched for Gene Ontology (GO) terms:
170 'chloroplast', 'ATP binding', 'proteasome' and 'nitrate transport' (**Supplementary table 4**).
171 mRNAs that were more abundant in seeds grown in HP conditions were enriched for the GO terms:
172 'cellular response to hypoxia', 'pectin esterase activity' and 'glucosinolate metabolic process'
173 (**Supplementary table 4**).

174 We also inquired the differences of the mRNA abundances between the MM and PI parental
175 lines, within and between treatments. To this end, we again employed a linear model, but were less
176 stringent in the statistical thresholds (as there were no confounding effects). We found 2,976
177 mRNAs differentially expressed between the two parental lines regardless of treatment and 382
178 mRNAs that were differentially expressed between the lines due to treatment (linear model, FDR
179 ≤ 0.05; **Figure 2B** and **Supplementary table 5**). GO enrichment indicated that the 1,240 mRNAs
180 more abundant in MM compared to PI were, among other categories, enriched, for 'transcription

181 factor activity', 'oxidation-reduction', 'protein -binding', '-phosphorylation', '-ubiquitination',
182 'chloroplast', 'circadian rhythm', and 'metal ion binding' (**Supplementary table 6a**). The 1,736
183 mRNAs that were more abundant in PI compared to MM were, among other categories, enriched
184 for 'cytosol', 'chloroplast', 'nucleus', 'mitochondrion', 'cytoplasm', 'ribosome', 'translation',
185 'nucleolus', 'endoplasmic reticulum', 'oxidation-reduction', 'vacuole', and 'copper ion binding'
186 (**Supplementary table 6a**). The 382 genes showing a significant interaction effect between the
187 parental background and maternal environment showed an enrichment for the GO terms 'oxidation-
188 reduction', 'extracellular region', 'transcript regulation', 'iron ion binding', and 'response to
189 gibberellin' (**Supplementary table 6b**). Of note, the 'oxidation-reduction process' and 'transcript
190 regulation' GO terms are enriched in the upregulated genes of both MM and PI, which is not
191 surprising since both GO terms are quite general and each represents many genes. These results
192 show that the nutrition status of the mother plant (environment; E) as well as genotype (G), and the
193 interaction between the two (G x E), modulate mRNA abundances in dry seeds of tomato.



194

195 **Figure 2**: Nutrition status-related mRNA abundance differences between genotypes and seed
196 development and maturation environments. (**A**) The first two axes of a principal component
197 analysis on the $\log_2$ ratio with the mean transcripts per million (TPM) values. The first axis (PCO1)
198 explained 16.9% of the variance in the data, the second 12.2%. Square symbols represent individual
199 RILs, Moneymaker (MM) parental samples are represented by dots and *S. pimpinellifolium* (PI)

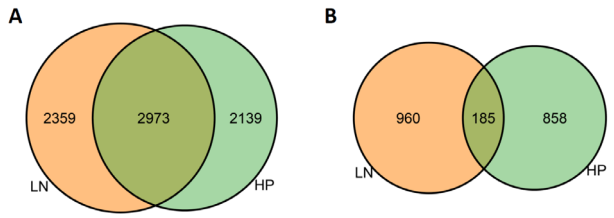
parental samples by triangles. The colours indicate high phosphorous (HP; green) or low nitrogen (LN; orange) treatments applied to the mother plants. (**B**) Differentially abundant mRNAs in the two parental lines that are either not affected (line) or affected by treatment (interaction). Lower mRNA abundance is shown in purple and higher in green (see legend).

Heritability and transgression in mRNA abundances

To estimate the contribution of genetic variation to differences in mRNA abundance between the genetic backgrounds (plant lines) and treatments (nutrient status), we calculated the Broad-Sense Heritability (BSH). In addition, replicated measurements in the parental lines were used to estimate non-genetic variance. We found 5,112 genes in HP and 5,332 genes in LN that showed significant heritability for mRNA abundance, of which 2,973 genes overlapped (39.8%; permutation, FDR < 0.05; **Figure 3A; Supplementary table 7a**). Subsequently, we checked if genes with significant heritable contribution to mRNA abundance differences were predominantly affected by the maternal nutrient environment. However, we did not find such an enrichment for any of the overlapping groups of genes (hypergeometric test, $p > 0.01$; **Supplementary figure 1A**). We thus conclude that, overall, the number of genes with significant heritability for mRNA abundance were not specifically responsive to the maternal nutrient treatments. The genes with heritable mRNA abundance in HP alone were enriched for the GO terms: 'translation', 'ribosome', 'mitochondrion', and more (**Supplementary table 7b**). Those that showed significant heritability only in LN were enriched for the GO terms: 'ABA metabolic process', and others (**Supplementary table 7b**). The genes that showed significant heritability in both environments were enriched for various GO terms: 'oxidation-reduction process', 'ribosome/translation', 'nucleolus', 'cell wall', 'heme binding', 'ion binding', and 'vacuole' (**Supplementary table 7b**).

We next assessed the complexity of the genetic regulation underlying mRNA abundance differences. To this end, the transgression was calculated, i.e., trait values in RILs that extent

226 beyond the parental means. We found significant transgression in mRNA abundance (trait) levels
227 for 1,043 genes in the maternal HP treatment and 1,145 genes in the maternal LN treatment
228 (permutation, FDR < 0.05; **Supplementary table 8a**). This suggests a polygenic genetic
229 architecture for mRNA abundance. Of these, the mRNA abundances of 185 genes showed
230 significant transgression beyond the parental means in both treatments (**Figure 3B**). Also, here, we
231 tested for significant overlap with treatment-related genes. Yet, with 18% response to treatment of
232 the transgressive mRNAs, there was no significant enrichment for transgressive mRNA
233 abundances with treatment-related differences (hypergeometric test, $p > 0.01$; **Supplementary**
234 **figure 1B**). So, alike heritability, transgression is apparently not linked to a reduction of nitrogen
235 or increase of phosphorus content in the maternal growth environment. Moreover, compared to
236 genes showing significant heritability, many fewer GO terms were enriched in the genes showing
237 transgression, and those GO terms that were enriched, generally had a lower level of significance.
238 For genes showing transgression in HP alone, the GO terms 'cell periphery', 'positive
239 gravitropism', 'cysteine biosynthetic process', 'symporter activity' and 'response to heat' were
240 enriched. Whereas for genes only showing transgression in LN the GO-terms 'beta-glucosidase
241 activity', 'preprophase band', and 'phragmoplast' were enriched. The GO term 'DNA-binding
242 transcription factor activity' was enriched in genes showing transgression in both environments
243 (**Supplementary table 8b**).
244
245
246



247
248

249 **Figure 3**: Venn diagrams showing the overlap and differences of **(A)** genes with significant
250 heritable variance (Supplementary table 7a) and **(B)** genes exhibiting significant transgression
251 (Supplementary table 8a), of mRNA abundance levels between LN (orange) and HP (green; FDR
252 < 0.05).
253

254 *The maternal nutrient environment produces specific eQTL landscapes.*

255 Altogether, our analyses revealed both a considerable effect of the maternal nutrient environment
256 (HP versus LN) and a significant influence of genetic variation in the RIL panel (heritability) on
257 the detected mRNA abundance levels. By combining our constructed high density SNP genetic
258 map (**Figure 1A, Supplementary table 1**) with the obtained mRNA abundance dataset (**Figure**
259 **2**), we were able to identify eQTLs that potentially contribute to the variation in mRNA abundance
260 (**Figure 4A-F)**. In other words, the identified eQTL loci have a high chance of harboring
261 polymorphic regulatory factors (e.g., genes or other genetic elements) for mRNA abundance,
262 prospectively explaining variation in the seed and germination trait phenotypes observed.

263 We detected a maternal environment-specific *trans*-eQTL landscape, as the distribution of
264 the position of the *trans*-eQTLs was very different between the two environments. For the HP
265 environment, 4,281 eQTLs for 3,833 genes were identified, of which 2,247 were *cis*-eQTLs and
266 2,034 were *trans*-eQTLs. For the LN environment, 7,487 eQTLs were detected for 6,815 genes, of
267 which 2,356 were *cis*-eQTLs and 5,131 were *trans*-eQTLs (FDR < 0.05; $-\log_{10}(p) > 3.9$; **Figure**
268 **4A-D; Supplementary Table 9**; **Supplementary Table 10**). A significant overlap between *cis*-
269 eQTLs of the two environments was noted (**Figure 4E;** 1,506 overlapping *cis*-eQTLs; 48.6%;

270 hypergeometric test, $p < 1*10^{-16}$). On the contrary, the *trans*-eQTLs were mainly specific for each
271 tested maternal environment (**Figure 4F**; 590 overlapping *trans*-eQTLs; 9.7%; hypergeometric
272 test, $p = 1.0$). However, both *cis*- and *trans*-eQTLs were not enriched for genes with differentially
273 abundant mRNA levels based on the maternal environment (hypergeometric test, $p > 0.01$;
274 **Supplementary figure 3A** and **B**). Together with the significant transgression (**Supplementary**
275 **table 8a)** and considerable heritability of mRNA abundances (**Supplementary table 7a; Figure**
276 **3A)**, this indicates that *trans*-eQTLs represent a genotype-specific interaction with the maternal
277 nutrient environment. Many different GO terms were found to be enriched in the genes with
278 environment specific eQTLs. For an overview see **Supplementary table 11**).

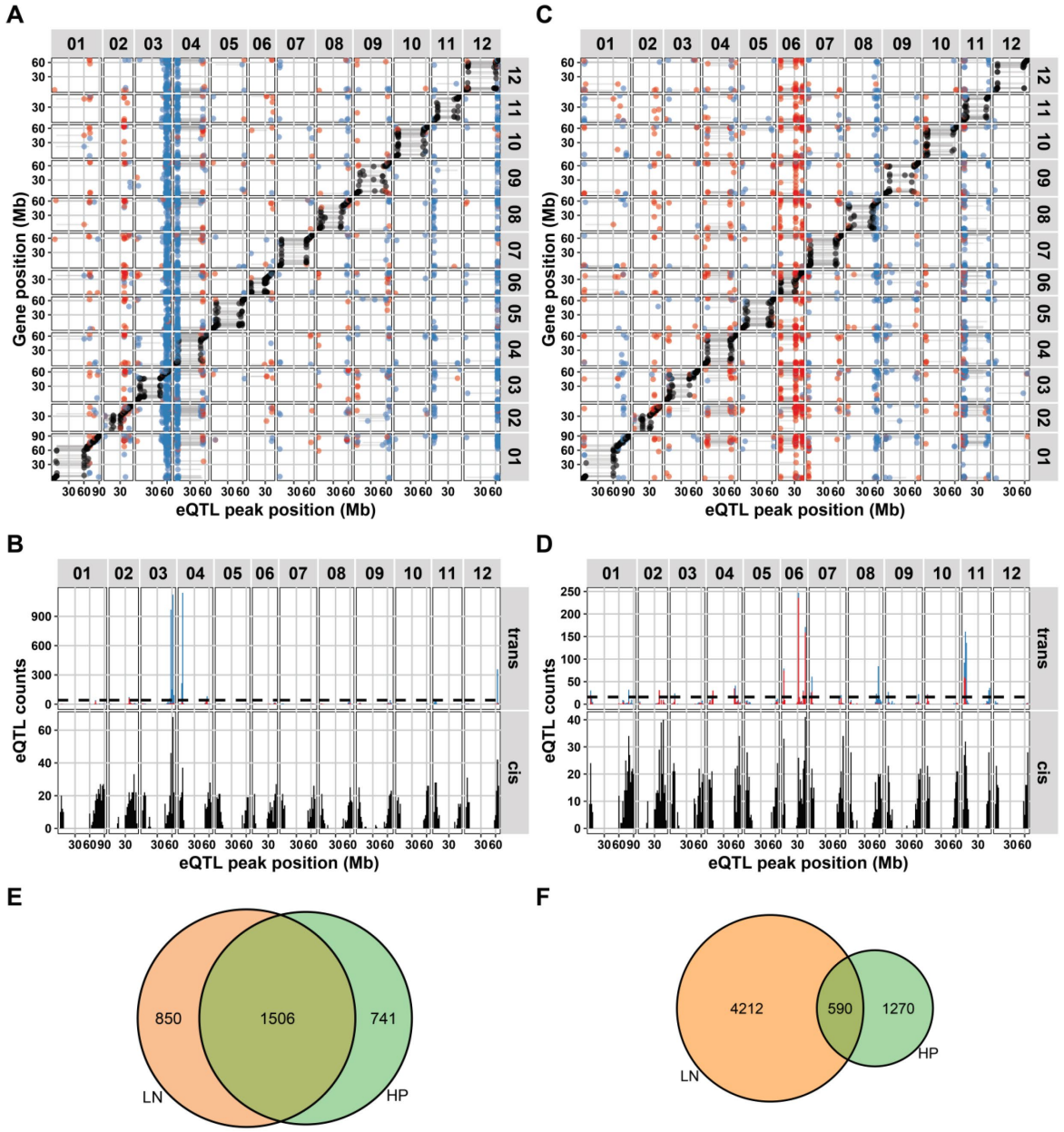


Figure 4: Characteristics of the detected eQTL landscapes in tomato dry seeds in (**A, B**) LN and (**C, D**) HP environments. (**A, C**) *Cis-trans* plots of eQTLs mapped ($-\log_{10}(p) > 3.9$). The positions (in Mb, per chromosome) of the eQTL peaks are plotted on the x-axis and the positions of the corresponding genes on the y-axis. Chromosome numbers are indicated on the top and right (grey labels). Colored dots indicate *cis*-eQTLs (black), eQTLs associated with higher mRNA abundance due to the MM allele (blue dots) or with higher abundance by the PI allele (red dots). (**B, D**) Histograms showing the distribution of the amount of *cis*- (lower panel) and *trans* (upper panel)-eQTLs over the chromosomes, arranged by eQTL peak location counted per 2 million bases (Mb) bins. The dashed lines in the *trans*-eQTL panels indicate the threshold for calling a *trans*-band (poisson distribution, $p < 0.0001$). (**E**) The overlap of *cis*-eQTLs in the two treatments and (**F**) the overlap of *trans*-eQTLs in the two maternal environments.

291
292 The majority of the *trans*-eQTLs clustered in maternal nutrient environment-specific eQTL
293 hotspots or *trans*-bands (**Figure 4A, C**). Hence, these genomic regions harbor the main loci
294 underlying the genetic variation in environment-specific gene expression regulation in our dataset.
295 A total of 13 *trans*-bands (9 in the HP treatment and 4 in the LN treatment; see Methods for the
296 *trans*-band criteria) were identified, which account for 1,206 of the *trans*-eQTLs in the HP
297 treatment (59.3% of HP total) and 4,181 of the *trans*-eQTLs in the LN treatment (81.5% of LN
298 total; **Table 1**).
299
300 **Table 1**: Overview of detected *trans*-band (TB) eQTLs. Indicated are given ID's, location on the
301 physical genome (map position in Mb), number of eQTLs underlying the *trans*-band (+ sign: MM
302 > PI; - sign PI > MM), GO terms enriched in the eQTLs underlying the *trans*-band in either MM
303 or PI and co-location with known phenotypic QTLs for germination [15, 17]; + sign: MM > PI; -
304 sign PI > MM).
305

TB ID	Position	eQTLs	GO enrichment (+ MM higher)	GO enrichment (- PI higher)	Germination QTL Refs: [15, 17]
LN_TB1	ch03: 56-64 Mb	2369 (2330+; 39-)	Translation; ribosome; nucleolus; RNA binding; mitochondrion; cell wall; and more	None	Th-I (LN-)
LN_TB2	ch04:4-8 Mb	1348 (1311+; 37-)	Telomere; nucleus; protein binding; ubiquitin; and more	None	Gmax water (LN+); T10 water (LN+); T10 NaCl, mann, HT (LN+); T50 water, mann, (LN+); AUC water, mann (LN+)
LN_TB3	ch04:54-56 Mb	99 (45+; 54-)	None	Secretory vesicle; and more	Gmax NaCl, mann (LN+); SW (LN+)
LN_TB4	ch12:62-66 Mb	365 (348+; 17-)	Golgi; endosome; glycosylation; ER; and more	None	None in LN
HP_TB1	ch01:2-4 Mb	41 (16+; 25-)	None	None	None in HP
HP_TB2	ch04:6-8 Mb	38 (7+; 31-)	None	Heme binding; oxidoreductase; iron ion binding	Gmax mann, HT (HP+); Th-I (HP+)
HP_TB3	ch06:2-4 Mb	96 (5+; 91-)	None	None	U8416 NaCl, HT (HP+)
HP_TB4	ch06:32-34 Mb	254 (12+; 242-)	None	RNA processing	Th-T (HP+)
HP_TB5	ch06:44-48 Mb	182 (13+; 169-)	None	Tricarboxylic acid cycle; plastid; vacuolar membrane; cell wall	T10 NaCl, mann, HT (HP+); T50 NaCl, HT (HP+); AUC NaCl (HP+)
HP_TB6	ch07:0-2 Mb	77 (32+; 45-)	None	None	None in HP
HP_TB7	ch08:58-60 Mb	83 (80+; 3-)	Vacuole; oxidoreductase; golgi	None	Th-D (HP+)

HP_TB8	ch11:0-6Mb	371 (249+; 122-)	chromosome, centromeric region; ubiquitin conjugating enzyme activity	Transferase activity; hydrolase activity; response to heat	None
HP_TB9	ch11:52-54Mb	64 (63+; 1-)	Ribosome; nucleolus; translation; and more	None	T10 mann (HP-)

306

307 Thus, *trans*-bands are a major explanatory factor for *trans*-eQTLs. In other words, a relatively large
308 proportion of *trans*-eQTLs are caused by a few pleiotropic major-effect loci. Remarkably, the MM
309 allele had a positive effect on mRNA abundance for the majority of the eQTLs of the *trans*-bands
310 in the LN soil environment, whereas this was not so prevalent in the HP environment (**Table 1;**
311 **Figure 4A, C**). Most of these *trans*-bands showed enrichment for specific GO terms, such as
312 'translation' and 'specific cellular organelles' for LN and, 'oxidoreductase' and 'vacuole' for HP.
313 (**Table 1**, **Supplementary table 12**). Moreover, many of the *trans*-bands co-locate with known
314 QTLs for germination and seed traits (**Table 1** [15, 17]). These eQTLs can therefore contribute to
315 uncovering the molecular genetic mechanisms underlying the germination and seed trait QTLs.

316 **Discussion**

317 Our RNA-sequencing data obtained from a Tomato RIL population (*S. lycopersicum* (cv.
318 Moneymaker; MM) x *S. pimpinellifolium* (G1.1554; PI)) [17, 49], allowed for the construction of
319 a detailed and high resolution genetic map, describing the genotypes using 4,515 SNP markers.
320 This is over five times more than previously reported in Kazmi *et al*., 2012 [71], which used 865
321 markers. However, intrinsic to RNA-seq data, only SNPs present in the coding parts of the genes
322 (mRNA's) could be used. Therefore, determining the exact locus where recombination took place
323 would need additional genome sequencing as described in [70].

324 By measuring transcript levels (i.e. mRNA abundances) in the seeds of a tomato RIL
325 population that had matured in different maternal nutrient environments, we show that the maternal
326 environment affects both regulation and the genetic architecture of gene expression in progeny
327 seeds. Especially, *trans* eQTLs proved environment specific, which is comparable to other species
328 [57, 63-67, 72-74]. We found 3,833 genes (~26% of all detected expressed genes in the RILs), with
329 an eQTL in HP and 6,815 genes (~46% of all expressed genes in the RILs) with an eQTL in LN.
330 This is comparable to the number detected by Ranjan *et al*. 2016 [75], who used the upper part of
331 5 day-old hypocotyls of introgression lines (ILs), developed from the wild desert-adapted species
332 *Solanum pennellii* and domesticated *Solanum lycopersicum* cv. M82 [76], and found 5,300 genes
333 (~25% of total expressed genes) to have an eQTL, with roughly half in *cis* and half in *trans*. We
334 also found this close to 50/50 ratio in the HP condition, whereas in the LN condition the ratio of
335 *cis/trans* eQTLs was increased to 30/70. Research in yeast indicated that the detection of *trans*-
336 acting eQTLs is more strongly affected by the power of the study than detection of *cis*-acting
337 eQTLs [73]. So, it is likely that in our study we would have even more *trans*-eQTLs relative to *cis*-
338 eQTLs.

339 By comparing two different maternal environments in a population originating from two different
340 genetic backgrounds, many different maternal environment specific eQTLs were detected. This
341 underlines the interplay between genetics and nutrient environment in our study. Yet, we expect
342 much of the variation caused by this interplay will be uncovered in future studies increasing
343 numbers of different timepoints, environments and genotypes. More detailed data on the number
344 and type of polymorphisms between tomato lines, such as frameshifts [77] and copy number
345 variations [35], could facilitate identification of the causal polymorphic genes in this and other
346 eQTL studies. Moreover, combining eQTLs with QTLs obtained using phenotypic trait data [12,
347 15, 17], as well as other molecular data such as proteomics and/or metabolomics [47], will
348 contribute to obtaining mechanistic insight on how genotypic variation leads to phenotypic
349 variation between individuals at a systemic level. Furthermore, these eQTLs could be used as a
350 lead in studies with a larger source of wild-genotypes and combined with GWAS [39-42, 44, 78],
351 to pinpoint causal polymorphisms underlying variation at both the molecular and phenotypic levels.

352 **Methods**

353

354 *Plant lines, growth conditions, and nutrient treatments*

355 The mother plants (maternal conditions) were cultivated as described in Kazmi *et al*. 2012 and
356 Geshnizjani *et al*. 2020 [15, 17], in the greenhouse at Wageningen University, the Netherlands. In
357 short; the parental lines *Solanum lycopersicum* cv. Money maker (MM) and *Solanum*
358 *pimpinellifolium* accession CGN14498 (PI) as well as the derived recombinant inbred lines (RILs;
359 [49]; **Supplementary table 1**) were grown on rockwool under standard nutrient conditions (14
360 mM Nitrate and 1 mM Phosphate) with a 16h light (25°C) and 8h darkness (15°C) photoperiod.
361 From the moment the first flower opened, the plants were fertilized with the specific nutrient
362 solutions, low nitrate (2.4 mM Nitrate, 1 mM Phosphate) and high phosphate (14 mM Nitrate, 5
363 mM Phosphate) in two biological replicates per environment. The seeds were collected from
364 healthy and ripe fruits and the pulp still attached to the seeds was removed with 1% hydrochloric
365 acid (HCl) and a mesh sieve. Water was used to remove the remaining HCl and pulp. For
366 disinfection, seeds were treated with trisodium phosphate (Na3PO4.12H2O). Subsequently, seeds
367 were dried at 20°C for 3 days on a clean filter paper in ambient conditions. The seeds were then
368 stored in paper bags at room temperature.

369

370 *RNA-isolation, library prep and RNA-seq*

371 We used 10 mg grinded powder derived from 30 whole, dry, brushed, after-ripened seeds (12
372 months after harvest) of parental lines and the RILs grown under the different nutrient
373 environments in a GGG design [79, 80] to extract total RNA. In total, 3 replicates per treatment for
374 the parental lines where sequenced and 49 single RIL seed pools for HP and 52 single RIL seed
375 pools for LN (**Supplementary table 13**). RNA was isolated using the NucleoSpin RNA plant

376 isolation kit (Macherey-Nagel 740949) with on-column DNA digestion and adding Plant RNA
377 isolation Aid (Life technologies) according to the manufacturer's protocol and instructions. Strand-
378 specific RNA-seq libraries were prepared from each RNA sample using the TruSeq RNA kit from
379 Illumina according to manufacturer's instructions. Poly-A-selected mRNA was sequenced using
380 the Illumina HiSeq2500 sequencer, producing strand-specific single-end reads of 100 nucleotides.
381 Raw sequence reads can be found in the Sequence Read Archive (SRA; www.ncbi.nlm.nih.gov/sra)
382 under ID: PRJNA704909
383
384 *Alignment and SNP calling*
385 Reads were trimmed using Trimmomatic (version 0.33, [81] to remove low quality nucleotides.
386 Trimmed reads were subsequently mapped to the Tomato SL4.0 reference genome with the
387 ITAG4.0 annotation [82] using the HISAT2 software (version 2.1.0, [83] with the --dta-cufflinks
388 option. The resulting SAM alignment files were sorted and indexed using samtools version1.9 [84].
389 SNPs were called using bcftools mpileup with a minimum read depth of 3.
390
391 *Generation of a genetic map from RNA-seq data*
392 The genetic map used for mapping the eQTLs was made from the RNA-seq data following the
393 protocol described in Serin & Snoek *et al*. 2017 [13] and Snoek *et al.* 2019 [69]. With the following
394 modifications: SNPs were filtered for those that were consistently found in all replicates of the
395 parental lines and observed in all RILs. Then the genotype per RIL was determined per sliding bin
396 of 100 SNPs where the mean position of those SNPs was taken as the physical position of the
397 obtained marker.
398
399

400 *Quantification of RNAseq*

401 Before mRNA abundance analysis, between 12M and 31M reads per sample were mapped to the
402 SL4.0 genome with ITAG4.0 annotation [82] using HISAT2 as described above. The mRNA
403 abundance was quantified to counts using Stringtie [85] with the options -e, -B and -G. In R, the
404 counts were used to calculate transcripts per million (TPM). The TPM values were $\log_2$-
405 transformed by

406 $$TPM_{log} = \log_2(TPM + 1)$$

407 Additionally, to use for statistics, also a ratio with the average was calculated, by

408 $$TPM_{rat,i,j} = \log_2(\frac{TPM_{i,j}}{\overline{TPM_j}})$$

409 Where the $\log_2$ was calculated for each transcript *i* of sample *j* by dividing over the average value
410 for that transcript $\overline{TPM}$ over all samples *j*. After transformation, the transcripts were filtered for
411 $TPM_{log} > 0$, and detection in all samples.

412

413 *mRNA abundance analysis and QTL analyses*

414 The analyses reported below were conducted in "R" (version 3.5.3, x64)[86] with custom written
415 scripts, accessible via https://git.wur.nl/published_papers/sterken_tomato-eqtl_2021. For analysis,
416 the dplyr and tidyr packages were used for data organization [87, 88], and plots were generated
417 using ggplot2 [89].

418

419 *Treatment related mRNA abundance differences*

420 The principal component analysis comparing the mRNA abundances was done on the TPM_{rat}-
421 transformed data, using the *prcomp* function in "R". The mRNA abundance differences between
422 treatments were tested between the LN and HP treatments using the linear model

423 $$TMP_{log,i} = T_i + e_i$$

424 where $TPM_{log,i}$ is the abundance level of transcript *i* (one of 14,772 transcripts) in RIL *j* (n = 55 for
425 the HP treatment and n = 58 for the LN treatment), *T* is the treatment (HP or LN), and *e* is the error
426 term. To reduce the chance of detecting differences due to genetic variation, a strict multiple-testing
427 correction was applied (Bonferroni) using *p.adjust*. The threshold for significance was $-\log_{10}(p) >$
428 5.47 (FDR = 0.05).

429 To determine the effect of treatment on the differences in mRNA abundance between the
430 parental lines, we ran a linear model explaining the differences due to treatment and line effects on
431 the MM and PI parental data. The model used was

432 $$TMP_{log,i,j} = T_{i,j} + L_{i,j} + T_{i,j} \times L_{i,j} + e_{i,j}$$

433 where $TPM_{log,i,j}$ is the abundance level of transcript *i* (one of 14,772 transcripts) in parental replicate
434 *j* (n = 3 for both treatments for MM and PI), *T* is the treatment (HP or LN), *L* is the line (MM or
435 PI), and *e* is the error term. Values were corrected for multiple testing using *p.adjust* following the
436 Benjamini Hochberg algorithm. The thresholds for FDR = 0.05 were: $-\log_{10}(p) = 1.71$ for line, -
437 $\log_{10}(p) = 2.08$ for treatment, and $-\log_{10}(p) = 2.89$ for the interaction between line and treatment.
438 We took the most stringent p value, $-\log_{10}(p) = 2.89$ as threshold to determine significance.

439

440 *Transgression*

441 Transgression was calculated by counting the number of lines with expression levels beyond three
442 standard deviations from the mean of the parental lines (as in RB Brem and L Kruglyak [90]); $\mu \pm$
443 $3*\sigma$. This was done for both treatments separately. The lower boundary was established by the
444 parental line with the lowest mean, and the upper boundary was established by the parental line

445 with the highest mean. The standard deviation used to determine transgression (σ) was calculated
446 as the pooled standard deviation of the two parental lines (n =3 for both).
447 Significance of the transgression was calculated by permutation. The expression values
448 were randomized over the line designations and the same test as above was conducted. This was
449 repeated 1000 times for each transcript, so the obtained values could be used as the by-chance
450 distribution. The 50th highest value was used as the false discovery rate (FDR) = 0.05 threshold.
451
452 *Heritability*
453 The heritability was calculated by estimating the genotypic variance in the RILs and the remaining
454 variance (e.g. measurement error) in the parental lines (as in JJ Keurentjes, J Fu, IR Terpstra, JM
455 Garcia, G van den Ackerveken, LB Snoek, AJ Peeters, D Vreugdenhil, M Koornneef and RC
456 Jansen [56]). This was done for both treatments separately, by

457 $$H^2_{RIL} = \frac{V_{RIL} - V_e}{V_{RIL}}$$

458 where V_{RIL} is the variance within the RIL population and V_e is the pooled variance of both parental
459 lines.
460 To establish whether the heritability was significant and not outlier-driven, we applied a
461 permutation approach (as in A Vinuela, LB Snoek, JA Riksen and JE Kammenga [91]). The trait
462 values were randomized over the line designations and the heritability calculation were repeated.
463 This was done 1000 times for each transcript to generate a by-chance distribution. The 50th highest
464 value was used as the FDR = 0.05 threshold.
465
466
467

468 *eQTL mapping*

469 For eQTL mapping a single marker model was used, and was applied separately for both treatments
470 (as in [65, 92]). QTLs were mapped using the model

471 $$TPM_{log,i,j} = x_j + e_j$$

472 where $TPM_{log,i,j}$ is the expression level of transcript i (one of 14,772 transcripts) in RIL j (n = 49
473 for the HP treatment and n = 52 for the LN treatment). The expression levels were explained over
474 the genotype on marker location x (x = 1, 2, ..., 4515) of RIL j.

475 To determine the reliability of the detected QTLs and correct for multiple testing, a
476 permutation approach was used. As in the other permutations, the expression levels were randomly
477 distributed over the lines and this randomized set was mapped again according to the procedure
478 described above, which was repeated 10 times. To determine the FDR, we applied a correction for
479 multiple testing under dependency [93]

480 $$\frac{FDS}{RDS} \leq \frac{m_0}{m} \times q \times \log(m)$$

481 where *FDS* (false discovery) is the number of eQTLs detected in the permutation and the *RDS* (real
482 discovery) is the number of eQTLs detected in the QTL mapping at a specific significance level.
483 The number of true null hypotheses tested (m_0), was 14,772 -*RDS*, where the number of hypotheses
484 tested (m) was the number of transcripts, 14,772. The q-value was set at 0.05, which led to a
485 threshold of $-\log_{10}(p) = 3.7$ for the LN treatment and $-\log_{10}(p) = 3.9$ for the HP treatment. To keep
486 comparisons straightforward (similar effect sizes), analyses were conducted at the most stringent
487 threshold ($-\log_{10}(p) > 3.9$).

488 The eQTL types (*cis* or *trans*) were called based on distance to the gene encoding the
489 affected transcript. A *trans*-eQTL had to be located at least 1 Mb from the gene. Furthermore, we
490 calculated the confidence interval of the QTL as a 1.5-drop from the highest $-\log_{10}(p)$. For a *trans*-

491 eQTL to be called, the location of the affect transcript was required to be outside of this confidence
492 interval as well.
493
494 *Trans-band identification*
495 Identification of regulatory hotspots (*trans*-bands) was based on assessing whether the number of
496 *trans*-eQTLs mapped to a locus exceeded the expected number based on an equal genome-wide
497 distribution (as in [65, 94]. We used a Poisson distribution to ascertain the significance of eQTL
498 abundances per 2 Mb bin. For the HP treatment, we expected 15.8 *trans*-eQTL per bin, and for the
499 LN treatment we expected 40.8 *trans*-eQTL per bin. We used a conservative threshold for calling
500 a bin enriched in *trans*-eQTL, $p < 0.0001$. After identifying significant bins, adjacent bins
501 (significant bins, with up to 1 non-significant bin in-between) were merged to a single *trans*-band.
502
503 *Enrichment*
504 GO enrichment was determined using the hypergeometric test in R on the GO annotation done for
505 ITAG2.4 downloaded from AgriGO (www.bioinfo.cau.edu.cn/agriGO) [95] combined with the
506 annotation for ITAG3.1 and expanded with the GO annotation of the Arabidopsis homologues. All
507 expressed genes were used as background genes in the enrichment test.
508
509 *Map and eQTL data in TomQTL*
510 The physical map of the RIL population and the eQTL –log10(p-value) scores are available for
511 download and online exploration in TomQTL at http://www.bioinformatics.nl/TomQTL/, an
512 interactive website based on AraQTL [63] and WormQTL2 [74].
513
514

515 **Acknowledgements**

516 We thank Prof. Dr. Jan Kammenga of the Laboratory of Nematology of Wageningen University,
517 Prof. Dr. Dick de Ridder of the Laboratory of Bioinformatics of Wageningen University and Prof.
518 Dr. Berend Snel of the Theoretical Biology and Bioinformatics department of Utrecht University
519 for their support.

520

521 **Author contributions**

522 WL and HWMH conceived the study, NG, JR and LW performed the experiments, MGS, HN and
523 LBS analyzed and visualized the data, MGS and LBS wrote the manuscript with input from MvZ,
524 HN, JJG, and WL. All authors approved the final version of this manuscript.

525

526 **Funding**

527 This work was supported by Technology Foundation (STW), which is part of the Netherlands
528 Organization for Scientific Research (NWO) (LW, JR, HN, WL). M.G.S. was supported by NWO
529 domain Applied and Engineering Sciences VENI grant (17282).

530 **Supplementary tables and figures**

531

532 **Supplementary table 1**: **Genetic map of the parental and Recombinant Inbred Lines used**.
533 Matrix of the 101 lines used, 3 F1 heterozygotes (columns) and the 4,515 detected markers listed
534 per chromosome (ch01-ch12) (lines). The genotypes are likelihood based where "0" indicates a
535 locus derived from MM and "1" indicates a locus derived from PI. Chromosome number and
536 genomic position (basepair) are given in the first two columns. Position is the average basepair
537 position of the 100 SNPs sliding bin used to determine the parental origin of the locus.

538

539 **Supplementary table 2**: **Introgression size statistics**. Minimum, maximum, mean and median
540 introgression sizes per chromosome (first column).

541

542 **Supplementary table 3**: Outcome of a linear model to detect differentially abundant mRNAs
543 between the HP versus LN-treated RILs. For the mRNAs, two identifiers are given in the columns:
544 identifier, and Name. Furthermore, the location (chromosome number and genome position of gene
545 start in basepairs), orientation (+ or – strand), and size (length in basepairs) are indicated. Then,
546 the outcome of the linear model is listed, first the significance in $-\log_{10}(p)$ followed by two types
547 of corrections for multiple testing: Bonferroni (conservative, as used in the main text) and
548 Benjamini Hochberg False-discovery rate (FDR; less conservative for comparison). The column
549 effect describes the difference between HP and LN treated maternal environment. A description of
550 the effect direction (treatment) is given in the last column.

551

552 **Supplementary table 4**: Gene Ontology (GO) enrichment data of maternal environment-related
553 mRNAs. Shown are the GO bin ID, GO bin category (GO name), GO aspect; molecular function

554 (F), cellular component (C), and biological process (P), P value (p.value), Total mRNA's identified
555 in mRNA bin and GO (in.set), total number of genes in GO (In.GO) and total mRNA set size
556 (set.size).
557
558
559 **Supplementary table 5**: Outcome of a linear model to detect differentially abundant mRNAs
560 between the MM and PI parental lines and their interaction with the environment. For the mRNAs,
561 two identifiers are given in the columns: identifier, and Name. Furthermore, the location
562 (chromosome number and genome position of gene start in basepairs), orientation (+ or – strand),
563 size (length in basepairs) are indicated. Then, the outcome of the linear model is listed, first the
564 tested factor, then significance in $-\log_{10}(p)$ and a correction for multiple testing (Benjamini
565 Hochberg (FDR)). The column effect describes the difference between the factors tested and the
566 interpretation of the effect direction is given in the last column.
567
568 **Supplementary table 6a**: Gene Ontology enrichment analysis of mRNAs that are higher in
569 parental lines MM (left), PI (middle) and their interaction (right). Significantly different genes were
570 taken from the model (see methods and material) only including the parental lines. See legend table
571 S4 for details and abbreviations.
572
573 **Supplementary table 6b**: Gene Ontology enrichment analysis of mRNA differences between the
574 parental lines and nutrient environment; higher in HP (left) or higher in LN (right). Significantly
575 different genes were taken from the model (see methods and material) including the parental lines
576 and the nutritional environment. See legend table S4 for details and abbreviations.
577

578 **Supplementary table 7a**: Heritability of mRNA abundances from the HP and LN maternal
579 environments. The treatment column indicates the maternal nutrient environment, the mRNA ID
580 is specified in the trait column. H2_keurentjes is the heritability, which was calculated using the
581 genotypic variance (Vg) and the residual variance (Ve) as described in Keurentjes *et al*. (2007)
582 [56]. The FDR column indicates the FDR = 0.05 threshold as determined by 1,000 permutations.
583 The last two columns specify if an mRNA abundance was significantly heritable and whether it
584 was specific for one or multiple maternal environments, or not (group).
585
586 **Supplementary table 7b**: Gene Ontology enrichment analysis of mRNAs with significant
587 heritability.
588
589 **Supplementary table 8a**: Transgression for mRNA abundances from the HP and LN maternal
590 environments in the RILs. The treatment column indicates the maternal environment, the mRNA
591 ID is specified in the trait column. The n_lines_transgression column specifies how many RILs
592 displayed transgression. The FDR column shows how many RILs showed transgression at the FDR
593 = 0.05 threshold as determined by 1,000 permutations. The last two columns specify if an mRNA
594 abundance was significantly transgressive or not and whether it was specific for one or multiple
595 maternal environments, or not (group).
596
597 **Supplementary table 8b**: Gene Ontology enrichment analysis of mRNAs with significant
598 transgression. See legend table S4 for details and abbreviations.
599
600 **Supplementary table 9**: Number of eQTLs detected per chromosome and treatment (LN; upper
601 table, HP; lower table). Indicated are for all eQTLs, QTL type (*cis* or *trans*) and QTL effect found

602 per chromosome per nutrient environment (+ or -). The last column indicates the number of eQTLs
603 in the *trans*-bands (TB).
604
605 **Supplementary table 10**: List with eQTLs mapped in both the LN and HP maternal nutrient
606 environments. First, the maternal environment is listed, second the mRNA ID (trait). Then columns
607 with the location information of the eQTL: chromosome number, location (bp; and the confidence
608 interval bp_left and bp_right), and the marker. Then, the significance in $-\log_{10}(p)$ is given and the
609 effect size (negative is higher in MM-derived loci; positive is higher in PI-derived loci).
610 Furthermore, the type of QTL is given (*cis* or *trans*) and whether the QTL is part of a *trans*-band.
611 Also, the variance explained by a single marker model is given (R2_sm). Subsequently, the name
612 and location (chromosome number and start of the gene in basepairs) of the mRNA is listed.
613
614 **Supplementary table 11**: Gene Ontology enrichment in genes with an eQTL. See legend table S4
615 for details and abbreviations.
616
617 **Supplementary table 12**: Gene Ontology enrichment in genes with eQTLs mapping to a *trans*-
618 band. See legend table S4 for details and abbreviations.
619
620 **Supplementary table 13:** Recombinant Inbred Lines per treatment.
621
622 **Supplementary figure 1**: Venn-diagrams showing **(A)** the overlap between all nutrient treatment-
623 affected mRNA abundances and HP and LN heritable mRNA abundances and (**B**) mRNAs
624 showing transgressive segregation in the HP and LN treatment.
625

626 **Supplementary figure 2**: Venn-diagrams showing the overlap between treatment-affected
627 mRNAs, **(A)** *trans*-eQTLs and (**B)** *cis*-eQTL mapped in the HP and LN nutrient environments.
628
629
630
631 **Figures**
632 **Figure 1**: (**A**) Genetic map showing the genotype of the RILs and parental lines. Dark blue
633 indicates MM (bottom horizontal line, with index 0), light blue indicates PI (horizontal line above
634 MM, with index 1). Shades between dark and light blue visualize the certainty of the estimate that
635 a locus corresponds to either MM or PI, depending on the SNPs identified (see legend above the
636 panel). (**B**) Allele frequency (percentage) of *S. pimpinellifolium* (PI) alleles for each marker across
637 the chromosomes, considering all RILs in the population. (**C**) Cumulative number of recombination
638 events per chromosome for the whole population. Chromosome numbers are indicated above
639 panels **A**, position on the chromosomes (in Mb) is shown on the x-axis below panel **C**.
640
641 **Figure 2**: Nutrition status-related mRNA abundance differences. (**A**) The first two axes of a
642 principal component analysis on the $\log_2$ ratio with the mean transcripts per million (TPM) values.
643 The first axis (PCO1) explained 16.9% of the variance in the data, the second 12.2%. Square
644 symbols represent individual RILs, Moneymaker (MM) parental samples are represented by dots
645 and *S. pimpinellifolium* (PI) parental samples by triangles. The colours indicate high phosphorous
646 (HP; green) or low nitrogen (LN; orange) treatments. (**B**) Differentially abundant mRNAs in the
647 two parental lines not affected by treatment (line) and affected by treatment (interaction). Lower
648 abundance is shown in purple and higher in green.
649

650 **Figure 3**: Venn diagrams showing the overlap and differences of **(A)** genes with significant
651 heritable variance and **(B)** genes exhibiting significant transgression, of mRNA abundance levels
652 between LN (orange) and HP (green; FDR < 0.05).
653
654 **Figure 4**: Characteristics of the detected eQTL landscapes in tomato dry seeds in (**A, B**) LN and
655 (**C, D)** HP environments. (**A, C**) *Cis-trans* plots of eQTLs mapped ($-\log_{10}(p) > 3.9$). The positions
656 of the eQTL peaks are plotted on the x-axis and the positions of the corresponding genes on the y-
657 axis. Chromosomes are indicated on the top and right in the grey labels. Colours indicate *cis*-eQTL
658 (black), eQTL associated with higher mRNA abundance due to the MM allele (blue) or with higher
659 abundance by the PI allele (red). (**B, D**) Histograms showing the distribution of the *cis*- and *trans*-
660 eQTL over the chromosomes, arranged by eQTL peak location counted per 2 million bases (Mb)
661 bins. The dashed lines in the *trans*-eQTL panels indicate the threshold for calling a *trans*-band
662 (poisson distribution, $p < 0.0001$). (**E**) The overlap of *cis*-eQTL in the two treatments and (**F**) the
663 overlap of *trans*-eQTL in the two maternal environments.
664
665 **Table 1**: Overview of detected *Trans*-band (TB) eQTLs. Indicated are given ID's, location on the
666 physical genome (map position in Mb), number of eQTLs underlying the *trans*-band (+ sign: MM
667 > PI; - sign PI > MM), GO terms enriched in the eQTLs underlying the *trans*-band in either MM
668 or PI and co-location with known phenotypic QTLs for germination [15, 17].
669
670

671 ## References

672 1. Rowse H, Finch-Savage W: **Hydrothermal threshold models can describe the germination**
673 **response of carrot (Daucus carota) and onion (Allium cepa) seed populations across both sub-**
674 **and supra-optimal temperatures**. *New Phytologist* 2003, **158**(1):101-108.

675 2. Delouche J, Baskin C: **Determinants of seed quality**. *SHORT COURSE FOR SEEDSMEN* 1971,
676 **14**:53-68.

677 3. Delouche JC: **Environmental effects on seed development and seed quality**. *HortScience* 1980,
678 **15**:775-780.

679 4. Donohue K: **Completing the cycle: maternal effects as the missing link in plant life histories**.
680 *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 2009,
681 **364**(1520):1059-1074.

682 5. de Souza Vidigal D, He H, Hilhorst HWM, Willems LAJ, Bentsink L: **Arabidopsis in the Wild-The**
683 **Effect of Seasons on Seed Performance**. *Plants* 2020, **9**(5).

684 6. Schmuths H, Bachmann K, Weber WE, Horres R, Hoffmann MH: **Effects of preconditioning and**
685 **temperature during germination of 73 natural accessions of Arabidopsis thaliana**. *Annals of*
686 *botany* 2006, **97**(4):623-634.

687 7. Demir I, Mavi K, Oztokat C: **Changes in germination and potential longevity of watermelon**
688 **(Citrullus lanatus) seeds during development**. *New Zealand Journal of Crop and Horticultural*
689 *Science* 2004, **32**(1):139-145.

690 8. He H, de Souza Vidigal D, Snoek LB, Schnabel S, Nijveen H, Hilhorst H, Bentsink L: **Interaction**
691 **between parental environment and genotype affects plant and seed performance in**
692 **Arabidopsis**. *J Exp Bot* 2014, **65**(22):6603-6615.

693 9. Munir J, Dorn L, Donohue K, Schmitt J: **The influence of maternal photoperiod on germination**
694 **requirements in Arabidopsis thaliana**. *American Journal of Botany* 2001, **88**:1240-1249.

695 10. Pourrat Y, Jacques R: **The influence of photoperiodic conditions received by the mother plant**
696 **on morphological and physiological characteristics of Chenopodium polyspermum L. seeds**.
697 *Plant Science Letters* 1975, **4**(4):273-279.
698 11. Alboresi A, Gestin C, LEYDECKER MT, Bedu M, Meyer C, TRUONG HN: **Nitrate, a signal relieving**
699 **seed dormancy in Arabidopsis**. *Plant, cell & environment* 2005, **28**(4):500-512.
700 12. Geshnizjani N, Sarikhani Khorami S, Willems LAJ, Snoek BL, Hilhorst HWM, Ligterink W: **The**
701 **interaction between genotype and maternal nutritional environments affects tomato seed and**
702 **seedling quality**. *J Exp Bot* 2019, **70**(10):2905-2918.
703 13. Serin EAR, Snoek LB, Nijveen H, Willems LAJ, Jimenez-Gomez JM, Hilhorst HWM, Ligterink W:
704 **Construction of a High-Density Genetic Map from RNA-Seq Data for an Arabidopsis Bay-0 x**
705 **Shahdara RIL Population**. *Front Genet* 2017, **8**:201.
706 14. Khan N, Kazmi RH, Willems LA, van Heusden AW, Ligterink W, Hilhorst HW: **Exploring the natural**
707 **variation for seedling traits and their link with seed dimensions in tomato**. *PLoS One* 2012,
708 **7**(8):e43991.
709 15. Geshnizjani N, Snoek BL, Willems LAJ, Rienstra JA, Nijveen H, Hilhorst HWM, Ligterink W:
710 **Detection of QTLs for genotype x environment interactions in tomato seeds and seedlings**.
711 *Plant, cell & environment* 2020.
712 16. Joosen RV, Arends D, Willems LA, Ligterink W, Jansen RC, Hilhorst HW: **Visualizing the genetic**
713 **landscape of Arabidopsis seed performance**. *Plant Physiol* 2012, **158**(2):570-589.
714 17. Kazmi RH, Khan N, Willems LA, Van Heusden AW, Ligterink W, Hilhorst HW: **Complex genetics**
715 **controls natural variation among seed quality phenotypes in a recombinant inbred population**
716 **of an interspecific cross between Solanum lycopersicum× Solanum pimpinellifolium**. *Plant, cell*
717 *& environment* 2012, **35**(5):929-951.

718 18. Schachtman DP, Reid RJ, Ayling SM: **Phosphorus Uptake by Plants: From Soil to Cell**. *Plant*
719 *Physiol* 1998, **116**(2):447-453.
720 19. Urbanczyk-Wochniak E, Fernie AR: **Metabolic profiling reveals altered nitrogen nutrient regimes**
721 **have diverse effects on the metabolism of hydroponically-grown tomato (Solanum**
722 **lycopersicum) plants**. *J Exp Bot* 2005, **56**(410):309-321.
723 20. Alboresi A, Gestin C, Leydecker MT, Bedu M, Meyer C, Truong HN: **Nitrate, a signal relieving**
724 **seed dormancy in Arabidopsis**. *Plant, cell & environment* 2005, **28**(4):500-512.
725 21. Duermeyer L, Khodapanahi, E., Yan, D., Krapp, A., Rothstein, S., & Nambara, E. : **Regulation of**
726 **seed dormancy and germination by nitrate**. *Seed Science Research* 2018, **28**(3):150-157.
727 22. Matakiadis T, Alboresi A, Jikumaru Y, Tatematsu K, Pichon O, Renou JP, Kamiya Y, Nambara E,
728 Truong HN: **The Arabidopsis abscisic acid catabolic gene CYP707A2 plays a key role in nitrate**
729 **control of seed dormancy**. *Plant Physiol* 2009, **149**(2):949-960.
730 23. He R, Yu D, Li X, Duan G, Zhang Y, Tang D, Zhao X, Liu X: **F-box gene FOA2 regulates GA- and**
731 **ABA- mediated seed germination in Arabidopsis**. *Science China Life sciences* 2016, **59**(11):1192-
732 1194.
733 24. Schauer N, Semel Y, Roessner U, Gur A, Balbo I, Carrari F, Pleban T, Perez-Melis A, Bruedigam C,
734 Kopka J: **Comprehensive metabolic profiling and phenotyping of interspecific introgression**
735 **lines for tomato improvement**. *Nature biotechnology* 2006, **24**(4):447-454.
736 25. Giovannoni J: **Molecular biology of fruit maturation and ripening**. *Annual review of plant*
737 *biology* 2001, **52**(1):725-749.
738 26. Tomato Genome C: **The tomato genome sequence provides insights into fleshy fruit evolution**.
739 *Nature* 2012, **485**(7400):635-641.

740 27. Tomato Genome Sequencing C, Aflitos S, Schijlen E, de Jong H, de Ridder D, Smit S, Finkers R,
741 Wang J, Zhang G, Li N *et al*: **Exploring genetic variation in the tomato (Solanum section**
742 **Lycopersicon) clade by whole-genome sequencing**. *Plant J* 2014, **80**(1):136-148.
743 28. Razifard H, Ramos A, Della Valle AL, Bodary C, Goetz E, Manser EJ, Li X, Zhang L, Visa S, Tieman D
744 *et al*: **Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin**
745 **America**. *Mol Biol Evol* 2020, **37**(4):1118-1132.
746 29. Doebley JF, Gaut BS, Smith BD: **The molecular genetics of crop domestication**. *Cell* 2006,
747 **127**(7):1309-1321.
748 30. McCouch S: **Diversifying selection in plant breeding**. *PLoS Biol* 2004, **2**(10):e347.
749 31. Bauchet G, Munos S, Sauvage C, Bonnet J, Grivet L, Causse M: **Genes involved in floral meristem**
750 **in tomato exhibit drastically reduced genetic diversity and signature of selection**. *BMC plant*
751 *biology* 2014, **14**:279.
752 32. Wang Z, Hong Y, Zhu G, Li Y, Niu Q, Yao J, Hua K, Bai J, Zhu Y, Shi H *et al*: **Loss of salt tolerance**
753 **during tomato domestication conferred by variation in a Na(+) /K(+) transporter**. *The EMBO*
754 *journal* 2020:e103256.
755 33. Yang J, Wang Y, Shen H, Yang W: **In silico identification and experimental validation of**
756 **insertion-deletion polymorphisms in tomato genome**. *DNA research : an international journal*
757 *for rapid publication of reports on genes and genomes* 2014, **21**(4):429-438.
758 34. Blanca J, Montero-Pau J, Sauvage C, Bauchet G, Illa E, Diez MJ, Francis D, Causse M, van der
759 Knaap E, Canizares J: **Genomic variation in tomato, from wild ancestors to contemporary**
760 **breeding accessions**. *BMC Genomics* 2015, **16**:257.
761 35. Razali R, Bougouffa S, Morton MJL, Lightfoot DJ, Alam I, Essack M, Arold ST, Kamau AA,
762 Schmockel SM, Pailles Y *et al*: **The Genome Sequence of the Wild Tomato Solanum**
763 **pimpinellifolium Provides Insights Into Salinity Tolerance**. *Front Plant Sci* 2018, **9**:1402.

764 36. Lin T, Zhu G, Zhang J, Xu X, Yu Q, Zheng Z, Zhang Z, Lun Y, Li S, Wang X *et al*: **Genomic analyses**
765 **provide insights into the history of tomato breeding**. *Nat Genet* 2014, **46**(11):1220-1226.
766 37. Pascual L, Albert E, Sauvage C, Duangjit J, Bouchet JP, Bitton F, Desplat N, Brunel D, Le Paslier
767 MC, Ranc N *et al*: **Dissecting quantitative trait variation in the resequencing era:**
768 **complementarity of bi-parental, multi-parental and association panels**. *Plant science : an*
769 *international journal of experimental plant biology* 2016, **242**:120-130.
770 38. Sauvage C, Segura V, Bauchet G, Stevens R, Do PT, Nikoloski Z, Fernie AR, Causse M: **Genome-**
771 **Wide Association in Tomato Reveals 44 Candidate Loci for Fruit Metabolic Traits**. *Plant Physiol*
772 2014, **165**(3):1120-1132.
773 39. Ye J, Li W, Ai G, Li C, Liu G, Chen W, Wang B, Wang W, Lu Y, Zhang J *et al*: **Genome-wide**
774 **association analysis identifies a natural variation in basic helix-loop-helix transcription factor**
775 **regulating ascorbate biosynthesis via D-mannose/L-galactose pathway in tomato**. *PLoS Genet*
776 2019, **15**(5):e1008149.
777 40. Zhang J, Zhao J, Xu Y, Liang J, Chang P, Yan F, Li M, Liang Y, Zou Z: **Genome-Wide Association**
778 **Mapping for Tomato Volatiles Positively Contributing to Tomato Flavor**. *Front Plant Sci* 2015,
779 **6**:1042.
780 41. Bauchet G, Grenier S, Samson N, Segura V, Kende A, Beekwilder J, Cankar K, Gallois JL, Gricourt J,
781 Bonnet J *et al*: **Identification of major loci and genomic regions controlling acid and volatile**
782 **content in tomato fruit: implications for flavor improvement**. *The New phytologist* 2017,
783 **215**(2):624-641.
784 42. Zhao J, Sauvage C, Zhao J, Bitton F, Bauchet G, Liu D, Huang S, Tieman DM, Klee HJ, Causse M:
785 **Meta-analysis of genome-wide association studies provides insights into genetic control of**
786 **tomato flavor**. *Nat Commun* 2019, **10**(1):1534.

787 43. Vosman B, Van't Westende WPC, Henken B, van Eekelen H, de Vos RCH, Voorrips RE: **Broad**
788 **spectrum insect resistance and metabolites in close relatives of the cultivated tomato**.
789 *Euphytica: Netherlands journal of plant breeding* 2018, **214**(3):46.
790 44. Chang J, Yu T, Yang Q, Li C, Xiong C, Gao S, Xie Q, Zheng F, Li H, Tian Z *et al*: **Hair, encoding a**
791 **single C2H2 zinc-finger protein, regulates multicellular trichome formation in tomato**. *Plant J*
792 2018, **96**(1):90-102.
793 45. Albert E, Segura V, Gricourt J, Bonnefoi J, Derivot L, Causse M: **Association mapping reveals the**
794 **genetic architecture of tomato response to water deficit: focus on major fruit quality traits**. *J*
795 *Exp Bot* 2016, **67**(22):6413-6430.
796 46. Zhang S, Yu H, Wang K, Zheng Z, Liu L, Xu M, Jiao Z, Li R, Liu X, Li J *et al*: **Detection of major loci**
797 **associated with the variation of 18 important agronomic traits between Solanum**
798 **pimpinellifolium and cultivated tomatoes**. *Plant J* 2018, **95**(2):312-323.
799 47. Kazmi RH, Willems LAJ, Joosen RVL, Khan N, Ligterink W, Hilhorst HWM: **Metabolomic analysis**
800 **of tomato seed germination**. *Metabolomics : Official journal of the Metabolomic Society* 2017,
801 **13**(12):145.
802 48. Celik I, Gurbuz N, Uncu AT, Frary A, Doganlar S: **Genome-wide SNP discovery and QTL mapping**
803 **for fruit quality traits in inbred backcross lines (IBLs) of solanum pimpinellifolium using**
804 **genotyping by sequencing**. *BMC Genomics* 2017, **18**(1):1.
805 49. Voorrips RE, Verkerke, W., Finkers, R., Jongerius, R. & Kanne, J.: **Inheritance of taste components**
806 **in tomato.** *Acta Physiologiae Plantarum* 2000(22):3.
807 50. Viquez-Zamora M, Caro M, Finkers R, Tikunov Y, Bovy A, Visser RG, Bai Y, van Heusden S:
808 **Mapping in the era of sequencing: high density genotyping and its application for mapping**
809 **TYLCV resistance in Solanum pimpinellifolium**. *BMC Genomics* 2014, **15**:1152.

810 51. Capel C, Yuste-Lisbona FJ, Lopez-Casado G, Angosto T, Cuartero J, Lozano R, Capel J: **Multi-**
811 **environment QTL mapping reveals genetic architecture of fruit cracking in a tomato RIL**
812 **Solanum lycopersicum x S. pimpinellifolium population**. *TAG Theoretical and applied genetics*
813 *Theoretische und angewandte Genetik* 2017, **130**(1):213-222.
814 52. Capel C, Fernandez del Carmen A, Alba JM, Lima-Silva V, Hernandez-Gras F, Salinas M, Boronat A,
815 Angosto T, Botella MA, Fernandez-Munoz R *et al*: **Wide-genome QTL mapping of fruit quality**
816 **traits in a tomato RIL population derived from the wild-relative species Solanum**
817 **pimpinellifolium L**. *TAG Theoretical and applied genetics Theoretische und angewandte Genetik*
818 2015, **128**(10):2019-2035.
819 53. de Souza Vidigal D, Willems L, van Arkel J, Dekkers BJW, Hilhorst HWM, Bentsink L: **Galactinol as**
820 **marker for seed longevity**. *Plant science : an international journal of experimental plant biology*
821 2016, **246**:112-118.
822 54. Geshnizjani N, Ghaderi-Far F, Willems LAJ, Hilhorst HWM, Ligterink W: **Characterization of and**
823 **genetic variation for tomato seed thermo-inhibition and thermo-dormancy**. *BMC plant biology*
824 2018, **18**(1):229.
825 55. Rothan C, Diouf I, Causse M: **Trait discovery and editing in tomato**. *Plant J* 2019, **97**(1):73-90.
826 56. Keurentjes JJ, Fu J, Terpstra IR, Garcia JM, van den Ackerveken G, Snoek LB, Peeters AJ,
827 Vreugdenhil D, Koornneef M, Jansen RC: **Regulatory network construction in Arabidopsis by**
828 **using genome-wide gene expression quantitative trait loci**. *Proc Natl Acad Sci U S A* 2007,
829 **104**(5):1708-1713.
830 57. Snoek LB, Terpstra IR, Dekter R, Van den Ackerveken G, Peeters AJ: **Genetical Genomics Reveals**
831 **Large Scale Genotype-By-Environment Interactions in Arabidopsis thaliana**. *Front Genet* 2012,
832 **3**:317.

833 58. Kawakatsu T, Huang SS, Jupe F, Sasaki E, Schmitz RJ, Urich MA, Castanon R, Nery JR, Barragan C,
834 He Y *et al*: **Epigenomic Diversity in a Global Collection of Arabidopsis thaliana Accessions**. *Cell*
835 2016, **166**(2):492-505.
836 59. West MA, Kim K, Kliebenstein DJ, van Leeuwen H, Michelmore RW, Doerge RW, St Clair DA:
837 **Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in**
838 **Arabidopsis**. *Genetics* 2007, **175**(3):1441-1450.
839 60. Jimenez-Gomez JM, Wallace AD, Maloof JN: **Network analysis identifies ELF3 as a QTL for the**
840 **shade avoidance response in Arabidopsis**. *PLoS Genet* 2010, **6**(9):e1001100.
841 61. Jansen RC, Nap JP: **Genetical genomics: the added value from segregation**. *Trends Genet* 2001,
842 **17**(7):388-391.
843 62. Terpstra IR, Snoek LB, Keurentjes JJ, Peeters AJ, van den Ackerveken G: **Regulatory network**
844 **identification by genetical genomics: signaling downstream of the Arabidopsis receptor-like**
845 **kinase ERECTA**. *Plant Physiol* 2010, **154**(3):1067-1078.
846 63. Nijveen H, Ligterink W, Keurentjes JJ, Loudet O, Long J, Sterken MG, Prins P, Hilhorst HW, de
847 Ridder D, Kammenga JE *et al*: **AraQTL - workbench and archive for systems genetics in**
848 **Arabidopsis thaliana**. *Plant J* 2017, **89**(6):1225-1235.
849 64. Hartanto M, Joosen RVL, Snoek BL, Willems LAJ, Sterken MG, de Ridder D, Hilhorst HWM,
850 Ligterink W, Nijveen H: **Network analysis prioritizes DEWAX and ICE1 as the candidate genes for**
851 **two major eQTL hotspots in seed germination**. *BioRxiv* 2020.
852 65. Snoek BL, Sterken MG, Bevers RPJ, Volkers RJM, Van't Hof A, Brenchley R, Riksen JAG, Cossins A,
853 Kammenga JE: **Contribution of trans regulatory eQTL to cryptic genetic variation in C. elegans**.
854 *BMC Genomics* 2017, **18**(1):500.
855 66. Vinuela A, Snoek LB, Riksen JA, Kammenga JE: **Genome-wide gene expression regulation as a**
856 **function of genotype and age in C. elegans**. *Genome Res* 2010, **20**(7):929-937.

857 67. Cubillos FA, Stegle O, Grondin C, Canut M, Tisne S, Gy I, Loudet O: **Extensive cis-regulatory**
858 **variation robust to environmental perturbation in Arabidopsis**. *Plant Cell* 2014, **26**(11):4298-
859 4310.
860 68. Sterken MG, Bevers RPJ, Volkers RJM, Riksen JAG, Kammenga JE, Snoek BL: **Dissecting the eQTL**
861 **micro-architecture in Caenorhabditis elegans**. *BioRxiv* 2019.
862 69. Snoek BL, Volkers RJM, Nijveen H, Petersen C, Dirksen P, Sterken MG, Nakad R, Riksen JAG,
863 Rosenstiel P, Stastna JJ *et al*: **A multi-parent recombinant inbred line population of C. elegans**
864 **allows identification of novel QTLs for complex life history traits**. *BMC Biol* 2019, **17**(1):24.
865 70. Demirci S, van Dijk AD, Sanchez Perez G, Aflitos SA, de Ridder D, Peters SA: **Distribution, position**
866 **and genomic characteristics of crossovers in tomato recombinant inbred lines derived from an**
867 **interspecific cross between Solanum lycopersicum and Solanum pimpinellifolium**. *Plant J* 2017,
868 **89**(3):554-564.
869 71. Kazmi RH, Khan N, Willems LA, AW VANH, Ligterink W, Hilhorst HW: **Complex genetics controls**
870 **natural variation among seed quality phenotypes in a recombinant inbred population of an**
871 **interspecific cross between Solanum lycopersicum x Solanum pimpinellifolium**. *Plant, cell &*
872 *environment* 2012, **35**(5):929-951.
873 72. Li Y, Alvarez OA, Gutteling EW, Tijsterman M, Fu J, Riksen JA, Hazendonk E, Prins P, Plasterk RH,
874 Jansen RC *et al*: **Mapping determinants of gene expression plasticity by genetical genomics in**
875 **C. elegans**. *PLoS Genet* 2006, **2**(12):e222.
876 73. Albert FW, Bloom JS, Siegel J, Day L, Kruglyak L: **Genetics of trans-regulatory variation in gene**
877 **expression**. *Elife* 2018, **7**.
878 74. Snoek BL, Sterken MG, Hartanto M, van Zuilichem AJ, Kammenga JE, de Ridder D, Nijveen H:
879 **WormQTL2: an interactive platform for systems genetics in Caenorhabditis elegans**. *Database*
880 *(Oxford)* 2020, **2020**.

881 75. Ranjan A, Budke JM, Rowland SD, Chitwood DH, Kumar R, Carriedo L, Ichihashi Y, Zumstein K,
882 Maloof JN, Sinha NR: **eQTL Regulating Transcript Levels Associated with Diverse Biological**
883 **Processes in Tomato**. *Plant Physiol* 2016, **172**(1):328-340.
884 76. Eshed Y, Zamir D: **An introgression line population of Lycopersicon pennellii in the cultivated**
885 **tomato enables the identification and fine mapping of yield-associated QTL**. *Genetics* 1995,
886 **141**(3):1147-1162.
887 77. Kevei Z, King RC, Mohareb F, Sergeant MJ, Awan SZ, Thompson AJ: **Resequencing at >/=40-Fold**
888 **Depth of the Parental Genomes of a Solanum lycopersicum x S. pimpinellifolium Recombinant**
889 **Inbred Line Population and Characterization of Frame-Shift InDels That Are Highly Likely to**
890 **Perturb Protein Function**. *G3 (Bethesda)* 2015, **5**(5):971-981.
891 78. Mata-Nicolas E, Montero-Pau J, Gimeno-Paez E, Garcia-Carpintero V, Ziarsolo P, Menda N,
892 Mueller LA, Blanca J, Canizares J, van der Knaap E *et al*: **Exploiting the diversity of tomato: the**
893 **development of a phenotypically and genetically detailed germplasm collection**. *Horticulture*
894 *research* 2020, **7**:66.
895 79. Li Y, Breitling R, Jansen RC: **Generalizing genetical genomics: getting added value from**
896 **environmental perturbation**. *Trends Genet* 2008, **24**(10):518-524.
897 80. Li Y, Swertz MA, Vera G, Fu J, Breitling R, Jansen RC: **designGG: an R-package and web tool for**
898 **the optimal design of genetical genomics experiments**. *BMC Bioinformatics* 2009, **10**:188.
899 81. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data**.
900 *Bioinformatics* 2014, **30**(15):2114-2120.
901 82. Prashant S. Hosmani MF-G, Henri van de Geest, Florian Maumus, Linda V. Bakker, Elio Schijlen,
902 Jan van Haarst, Jan Cordewener, Gabino Sanchez-Perez, Sander Peters, Zhangjun Fei, James J.
903 Giovannoni, Lukas A. Mueller and Surya Saha: **An improved de novo assembly and annotation**

904 **of the tomato reference genome using single-molecule sequencing, Hi-C proximity ligation and**
905 **optical maps** *BioRxiv* 2020.
906 83. Kim D, Langmead B, Salzberg SL: **HISAT: a fast spliced aligner with low memory requirements**.
907 *Nature methods* 2015, **12**(4):357-360.
908 84. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R,
909 Genome Project Data Processing S: **The Sequence Alignment/Map format and SAMtools**.
910 *Bioinformatics* 2009, **25**(16):2078-2079.
911 85. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL: **StringTie enables**
912 **improved reconstruction of a transcriptome from RNA-seq reads**. *Nature biotechnology* 2015,
913 **33**(3):290-295.
914 86. R-Core-Team: **R: A Language and Environment for Statistical Computing**. *R Foundation for*
915 *Statistical Computing* 2017, **Vienna, Austria**(https://www.R-project.org/).
916 87. Wickham HF, R.; Henry, L; Müller, K.: **dplyr: A Grammar of Data Manipulation**. 2018.
917 88. Wickham HH, L: **tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions**. 2018.
918 89. Wickham H: **GGplot2: elegant graphics for data analysis.** 2009.
919 90. Brem RB, Kruglyak L: **The landscape of genetic complexity across 5,700 gene expression traits in**
920 **yeast**. *Proc Natl Acad Sci U S A* 2005, **102**(5):1572-1577.
921 91. Vinuela A, Snoek LB, Riksen JA, Kammenga JE: **Aging Uncouples Heritability and Expression-QTL**
922 **in Caenorhabditis elegans**. *G3 (Bethesda)* 2012, **2**(5):597-605.
923 92. Sterken MG, van Bemmelen van der Plaat L, Riksen JAG, Rodriguez M, Schmid T, Hajnal A,
924 Kammenga JE, Snoek BL: **Ras/MAPK Modifier Loci Revealed by eQTL in Caenorhabditis elegans**.
925 *G3 (Bethesda)* 2017, **7**(9):3185-3193.
926 93. Benjamini Y, Daniel Yekutieli: **The control of the false discovery rate in multiple testing under**
927 **dependency**. *he annals of statistics* 2001, **29**(4):23.

928 94. Rockman MV, Skrovanek SS, Kruglyak L: **Selection at linked sites shapes heritable phenotypic**
929 **variation in C. elegans**. *Science* 2010, **330**(6002):372-376.
930 95. Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, Xu W, Su Z: **agriGO v2.0: a GO analysis toolkit for the**
931 **agricultural community, 2017 update**. *Nucleic Acids Res* 2017, **45**(W1):W122-W129.
932