

Sterken et al. *Let-60* eQTL

1 **Ras/MAPK modifier loci revealed by eQTL in *C. elegans***

2

3 Mark G. Sterken<sup>\*</sup>, Linda van Bemmelen van der Plaat<sup>\*</sup>, Joost A. G. Riksen<sup>\*</sup>, Miriam  
4 Rodriguez<sup>\*</sup>, Tobias Schmid<sup>†,‡</sup>, Alex Hajnal<sup>†</sup>, Jan E. Kammenga<sup>\*1</sup>, Basten L. Snoek<sup>\*1</sup>

5

6 <sup>\*</sup>Laboratory of Nematology, Wageningen University, Droevendaalsesteeg 1, 6708PB,

7 Wageningen, The Netherlands

8 <sup>†</sup>University of Zurich, Institute of Molecular Life Sciences, Winterthurerstrasse 190

9 CH-8057 Zurich, Switzerland.

10 <sup>‡</sup>PhD Program in Molecular Life Sciences

11

12

13 The raw transcriptomic data is available at ArrayExpress (E-MTAB-5856)

14

Sterken et al. *Let-60* eQTL

15 Running title: Ras/MAPK modifier loci

16

17 Key words: *C. elegans*, RAS/MAPK, eQTL, genetic background

18

19 <sup>1</sup>Corresponding authors: Jan Kammenga, Laboratory of Nematology, Wageningen

20 University, Droevendaalsesteeg 1, 6708PB, Wageningen, The Netherlands. Tel: +31

21 317 482998. Email: [jan.kammenga@wur.nl](mailto:jan.kammenga@wur.nl);

22 Basten Snoek, Laboratory of Nematology, Wageningen University,

23 Droevendaalsesteeg 1, 6708PB, Wageningen, The Netherlands. Tel: +31 317

24 482998. Email: [l.b.snoek@uu.nl](mailto:l.b.snoek@uu.nl)

25

26

27 **ABSTRACT**

28 The oncogenic Ras/MAPK pathway is evolutionarily conserved across metazoans.  
29 Yet, almost all our knowledge on this pathway comes from studies using single  
30 genetic backgrounds, whereas mutational effects can be highly background  
31 dependent. Therefore, we lack insight in the interplay between genetic backgrounds  
32 and the Ras/MAPK-signaling pathway. Here, we used a *Caenorhabditis elegans* RIL  
33 population containing a gain-of-function mutation in the Ras/MAPK pathway gene  
34 *let-60* and measured how gene expression regulation is affected by this mutation. We  
35 mapped eQTL and found that the majority (~73%) of the 1516 detected *cis*-eQTL  
36 were not specific for the *let-60* mutation, whereas most (~76%) of the 898 detected  
37 *trans*-eQTL were associated with the *let-60* mutation. We detected 6 eQTL *trans*-  
38 bands specific for the interaction between the genetic background and the mutation,  
39 one of which co-localized with the polymorphic Ras/MAPK modifier *amx-2*.  
40 Comparison between transgenic lines expressing allelic variants of *amx-2* showed  
41 the involvement of *amx-2* in 79% of the *trans*-eQTL for genes mapping to this *trans*-  
42 band. Together, our results have revealed loci hidden loci affecting Ras/MAPK  
43 signaling using sensitized backgrounds in *C. elegans*. These loci harbor putative  
44 polymorphic modifier genes that would not have been detected using mutant screens  
45 in single genetic backgrounds.

## 46 INTRODUCTION

47 The Ras/MAPK pathway is highly conserved across metazoans and regulates a wide  
48 range of physiological responses, such as cell proliferation, apoptosis, cell  
49 differentiation, and tissue morphogenesis (Gokhale and Shingleton 2015). In  
50 humans, activating (“gain-of-function”) mutations in HRas and KRas are strong tumor  
51 initiating mutations (Prior and Hancock 2012). Activation of MAP kinase components  
52 in model organisms has been shown to cause cell transformation and is implicated in  
53 tumorigenesis (Mansour et al. 1994; Cowley et al. 1994). As a key pathway in  
54 vertebrates and invertebrates, Ras/MAPK-signaling has been thoroughly studied in  
55 model organisms. Genetic studies in the model nematode *Caenorhabditis elegans*  
56 have provided insight into *let-60* Ras/MAPK signaling. Activated LET-60, a member  
57 of the GTP-binding RAS family (Beitel et al. 1990; Han and Sternberg 1990), induces  
58 the phosphorylation of LIN-45 (a Raf ortholog), MEK-2 (a MAPK kinase), and MPK-1  
59 (an ERK ortholog) (Wu and Han 1994). After phosphorylation, MPK-1 enters the  
60 nucleus where it regulates many genes by phosphorylation of transcription factors  
61 (Tan et al. 1998). Additionally, *let-60* activation underlies programmed cell death in *C.*  
62 *elegans* (Jiang and Wu 2014).

63 In *C. elegans* almost all studies on *let-60* activation have been conducted with  
64 mutant screens using single genetic backgrounds (*i.e.* a mutation in one genotype, in  
65 this case Bristol N2). However, the phenotype of induced mutations can vary widely  
66 depending on the genetic background (Duveau and Felix 2012; Chandler et al. 2014;  
67 Schmid et al. 2015; Kammenga 2017). Induced mutations in one genetic background  
68 do not reveal the allelic effects that segregate in natural populations and contribute to  
69 phenotypic variation (Kammenga et al. 2008). At the moment we lack insight into how  
70 genetic background effects modulate activated Ras/MAPK signaling and which

Sterken et al. *Let-60* eQTL

71 genetic modifiers are involved in the underlying genetic architecture of gene  
72 expression.

73 Here we go beyond mutant screens of *let-60(gf)* in a single genetic  
74 background in *C. elegans* by incorporating the effects of multiple genetic  
75 backgrounds. This provides the opportunity to explore the genetic variation for  
76 identifying novel modifiers of the Ras/MAPK pathway. We recently mapped modifiers  
77 affecting Ras/MAPK signaling associated with vulval development in a population of  
78 recombinant inbred lines derived from a cross between wildtype Hawaiian CB4856  
79 and Bristol N2 (Schmid et al. 2015). The lines were sensitized by introgression of the  
80 G13E gain-of-function mutation (*n1046*) in the Ras gene *let-60* (mutation  
81 introgression recombinant inbred lines, miRILs). This mutation is analogous to  
82 mutations causing excess cell division in human tumors (Kyriakakis et al. 2015).  
83 Hawaiian CB4856 males were crossed with Bristol N2 *let-60* mutants. Random  
84 segregation of the two parental genomes was allowed, except for the *let-60* mutation  
85 which was kept homozygous from the F2 generation onwards. After ten generations  
86 of self-fertilization, to drive all regions to homozygosity, independent miRILs were  
87 successfully obtained, each carrying the mutation. Quantitative trait locus (QTL)  
88 analysis on the vulval index (VI) of these miRILs in combination with allele swapping  
89 experiments revealed the polymorphic monoamine oxidase A (MAOA) gene *amx-2*  
90 as a negative regulator of Ras/MAPK signaling (Schmid et al. 2015).

91 Here, we extended the study by Schmid *et al.* by studying the transcriptional  
92 architecture of the same *let-60(gf)* sensitized miRILs. First, we measured the  
93 transcriptome of 33 miRILs using microarrays. Second, to gain further insight in the  
94 underlying molecular mechanisms, pathways, and new modifiers we mapped gene  
95 expression QTL (eQTL). eQTL are polymorphic loci underlying variation in gene

Sterken et al. *Let-60* eQTL

96 transcript abundances and can be used to detect loci which affect many transcript  
97 levels and transcriptional pathways (Jansen and Nap 2001). We found that the *let-*  
98 *60(gf)* mutation mainly reveals novel *trans*-eQTL, which are eQTL distant from the  
99 regulated gene. In contrast, *cis*-eQTL – genes that co-localize with the eQTL – were  
100 mostly independent of the *let-60(gf)* mutation. Of the *trans*-eQTL, 77 genes have an  
101 eQTL in a *trans*-band (eQTL hotspot) co-localizing with *amx-2*. Comparing the  
102 transcriptional profiles of transgenic lines with the N2 *amx-2* allele with the CB4856  
103 *amx-2* allele showed the involvement of *amx-2* in 79% (61/77) of the genes mapping  
104 to this locus. Through network-assisted gene expression analysis, we found evidence  
105 that *amx-2* indirectly affects gene expression mapping to this *trans*-band.

## 106 MATERIAL AND METHODS

### 107 General methods and strains used

108 The strains MT2124 (N2 background with the *let60(n1046)* gain of function mutation)  
109 and CB4856 were used, as were 33 miRILs described previously (Schmid et al.  
110 2015). A file with the strains and the genetic map has been included (**Supplemental**  
111 **table S1 1**). Furthermore transgenic lines *amx-2(lf);Si[amx-2(Bristol)];let-60(gf)* and  
112 *amx-2(lf);Si[amx-2(Hawaii)];let-60(gf)* containing the N2- or CB4856-allele of *amx-2* in  
113 a *amx-2(lf);let-60(gf)* background were used in a confirmation experiment. For details  
114 regarding the construction of these lines, see (Schmid et al. 2015).

115 Strains were maintained on NGM agar seeded with OP50 bacteria at 20°C,  
116 strains were age-synchronized by bleaching and were harvested 72 hours post  
117 synchronization. One sample consisted of multiple pooled non-starved populations  
118 from ~4 separate 9 cm NGM plates.

119

### 120 RNA isolation, cDNA synthesis, and cRNA synthesis

121 The RNA of the samples was isolated using the RNEasy Micro Kit from Qiagen  
122 (Hilden, Germany), following the 'Purification of Total RNA from Animal and Human  
123 Tissues' protocol with a modified lysing procedure. As prescribed, 150 µl RLT buffer  
124 and 295 µl RNase-free water were used to lyse the samples, but with an addition of  
125 800 µg/ml proteinase K and 1% β-mercaptoethanol. This suspension was incubated  
126 for 30 minutes at 55°C and 1000 rpm in a Thermomixer (Eppendorf, Hamburg,  
127 Germany). Thereafter the protocol as supplied by the manufacturer was followed.

128 For gene-expression analysis 200 ng of RNA (as quantified by NanoDrop) was  
129 used in the 'Two-Color Microarray-Based Gene Expression Analysis; Low Input

Sterken et al. *Let-60* eQTL

130 Quick Amp Labeling' -protocol, version 6.0 from Agilent (Agilent Technologies, Santa  
131 Clara, CA, USA).

132

### 133 **Microarray hybridisation and scanning**

134 The Agilent *C. elegans* (V2) Gene Expression Microarray 4X44K slides were used to  
135 measure gene expression. As recommended by the manufacturer, the samples were  
136 hybridized for 17 hours and scanned by an Agilent High Resolution C Scanner. The  
137 intensities were extracted using Agilent Feature Extraction Software (version  
138 10.7.1.1). The data was normalized using the Limma package in "R" (version 3.3.0  
139 x64). For within-array normalization the 'Loess' method was used and the 'Quantile'  
140 method for between-array normalization (Zahurak et al. 2007; Smyth and Speed  
141 2003). Unless mentioned otherwise, the log<sub>2</sub> transformed intensities were used in  
142 subsequent analysis.

143

### 144 **Expanding the genetic map**

145 Previously, the strains were genotyped with 73 fragment length polymorphisms (FLP)  
146 markers (as described in (Schmid et al. 2015; Zipperlen et al. 2005)) (Supplemental  
147 Table S2)

148 In order to increase the resolution of the genetic map, *cis*-eQTL of a large *C.*  
149 *elegans* eQTL experiment were used to further pinpoint the crossovers (Rockman et  
150 al. 2010). The gene expression of the miRILs was transformed to the mean of the two  
151 parental lines (R) by

$$152 \quad R_{x,i} = \frac{Y_{x,i}}{0.5 * (Y_{MT2124,i} + Y_{CB4856,i})}$$

153 where Y stands for the untransformed intensities of strain x and spot i (1, 2, 3, ...,  
154 45220). The obtained values were correlated to the *cis*-eQTL effects from Rockman



Sterken et al. *Let-60* eQTL

155 *et al.*, 2010. Since the current study uses a newer version of the *C. elegans*  
156 microarray, only the spots that occurred in both designs were compared. The  
157 expression markers were generated following a procedure that is a variation of the  
158 method described in (West et al. 2006). Per 20 consecutive *cis*-eQTL, the correlation  
159 between the N2-eQTL effect and the transformed intensities (R) were calculated. In  
160 this way, the correlation would be negative if the genotype stemmed from CB4856  
161 and positive if the genotype stemmed from N2.

162 In this way, per strain 424 gene expression markers were generated. In order  
163 to control for quality, the markers were filtered for calling the correct genotype in the  
164 four MT2124 replicas, the four CB4856 replicas, and in more than 50% of the  
165 samples an absolute correlation  $> 0.5$  was required. This pruning resulted in 204  
166 reliable gene expression markers. The quality of these markers was controlled by  
167 predicting the genotyped by expression of RIL AH2244, of which the gene expression  
168 was measured twice. There, we found that 10/204 selected gene expression markers  
169 were in disagreement of the genotype. This could be reduced to 0 by comparing only  
170 markers with an absolute correlation  $> 0.6$ . Therefore, all the correlations with an  
171 absolute value  $> 0.6$  were assigned the predicted marker, which corresponds to an  
172 error rate  $< 0.01$  per strain. The genotypes of the remaining markers were manually  
173 inferred from the surrounding markers. Furthermore, the genotypes at the ends of the  
174 chromosomes were inferred from the distal most assigned marker. This brings the  
175 size of the resulting map to 289 markers. The correlations and assigned genotypes  
176 can be found in **Supplemental table S3**.

177

178 **Evaluating the genetic map**

Sterken et al. *Let-60* eQTL

179 The 289 marker set was analysed by correlation analysis for markers describing  
180 unique crossover events and to see if there are any strong linkages between  
181 chromosomes (**Supplemental figure S1**). This led to the selection of 247 markers  
182 indicating the border of a crossover event. To reduce the chances of false positives,  
183 we only used markers where the frequency of the least occurring genotype in the  
184 population was >15%. It has to be noted that this excluded most of chromosome IV  
185 (as the genotype was predominantly N2 due to selection for strains including the *let-*  
186 *60(gf)* mutation (Schmid et al. 2015). Furthermore, also the *peel-1/zeel-1* region on  
187 chromosome I was excluded for the same reason (Seidel et al. 2008).

188

### 189 **eQTL mapping and threshold determination**

190 Mapping of eQTL mapping was done in “R” (version 3.3.0 x64), the gene-expression  
191 data was fitted to the linear model,

192

$$y_{i,j} \sim x_j + e_j$$

193 where  $y$  is the log2 normalized intensity as measured by microarray of spot  $i$  ( $i = 1, 2,$   
194 ..., 45220) of miRIL  $j$ . This is explained over the genotype (either CB4856 or N2) on  
195 marker location  $x$  ( $x = 1, 2, \dots, 247$ ) of miRIL  $j$ .

196 A permutation approach was used to determine an empirical false discovery  
197 rate (Snoek et al. 2017; Vinuela et al. 2010). First, the log2 normalized intensities  
198 were randomly distributed per gene over the genotypes. This randomized dataset  
199 was tested using the eQTL mapping model. This procedure was repeated 10 times.

200 The threshold was determined using

$$201 \quad \frac{FDS}{RDS} \leq \frac{m_0}{m} q \cdot \log(m)$$

202 Where, at a specific significance level, the false discoveries (FDS) were the averaged  
203 outcome of the permutations and read discoveries (RDS) were the outcome of the

Sterken et al. *Let-60* eQTL

204 eQTL mapping. The value of  $m_0$ , the number of true null hypotheses tested, was  
205 45220-RDS and for the value of  $m$ , the number of hypotheses tested, the number of  
206 spots (45220) was taken. Because this study only used a limited set of strains, a  
207 more lenient threshold of 0.1 was taken as the  $q$ -value (Benjamini and Yekutieli  
208 2001), which resulted in a threshold of  $-\log_{10}(p) > 3.2$ .

209

### 210 **Statistical power calculations**

211 The statistical power of the mapping was determined at the significance threshold of  
212  $-\log_{10}(p) > 3.2$ . Using the genetic map (33 strains with 247 markers), 10 QTL were  
213 simulated per marker location, explaining 20-80% of the variation (in increments of  
214 5%). Next to a peak, also random variation was introduced, based on a standard  
215 normal distribution ( $\mu = 0$  and  $\sigma = 1$ ). Peaks with corresponding explanatory  
216 power were simulated in this random variation (e.g. a peak size of 1 corresponds to  
217 20% explained variation). Furthermore, also random data without peaks was  
218 generated. This simulated data was mapped as described above and for each  
219 simulated peak it was determined: (i) if it was detected correctly, (ii) how precise the  
220 effect size estimation was, and (iii) how precise the location was determined. An  
221 overview of the results can be found in **Supplemental table S4**.

222

### 223 **eQTL analysis**

224 The distinction between *cis*- and *trans*-eQTL was made on the distance between the  
225 gene and the eQTL-peak. Given the relatively low number of unique recombinations  
226 (due to the limited set of strains), the *cis*-eQTL window was set at 2 Mb. This means  
227 that if a gene lies within 2 Mb of the QTL peak or the confidence interval, it is called a

Sterken et al. *Let-60* eQTL

228 *cis*-eQTL. The confidence interval was based on a  $-\log_{10}(p)$  drop of 1.5 compared to  
229 the peak.

230 *Trans*-bands were identified based on a Poisson-distribution of the mapped  
231 *trans*-eQTL (as in (Rockman et al. 2010)). The number of *trans*-eQTL was counted  
232 per 1 Mb bin, leading to the identification of 52 bins with *trans*-eQTL. Since we  
233 mapped 1149 *trans*-eQTL peaks (spots) to 52 bins, we expected 22.10 *trans*-eQTL  
234 per bin with *trans*-eQTL assigned to it. Based on a Poisson-distribution, any bin  
235 linking more than 30 spots with a *trans*-eQTL would have a significance  $p < 0.05$  and  
236 more than 38 *trans*-eQTL would yield a significance of  $p < 0.001$ . In this way, 6 *trans*-  
237 band loci were identified.

238 Comparative analysis against previous eQTL experiments was done on re-  
239 mapped experiments downloaded from WormQTL (Snoek et al. 2013; van der Velde  
240 et al. 2014) and EleQTL (<http://www.bioinformatics.nl/EleQTL>). This dataset consists  
241 of 4 different experiments representing 9 different conditions. The first set contains  
242 eQTL in two temperature conditions, 16°C and 24°C, measured in the L3 stage (Li et  
243 al. 2006). The second set contains eQTL over 3 life stages: L4 juvenile animals  
244 grown at 24°C, reproducing adult animals (96h) grown at 24°C, and aging animals  
245 (214h) grown at 24°C (Vinuela et al. 2010). The third set contains eQTL from a single  
246 experimental condition (young adults grown at 20°C) measured on a large RIL panel  
247 (Rockman et al. 2010). The fourth set contains eQTL from three experimental  
248 conditions over the course of a heat-shock treatment: a control condition (L4 animals  
249 grown for 48h at 20°C), a heat-shock condition (L4 animals grown for 46h at 20°C  
250 and exposed to 2h of 35°C), and a recovery condition (similar to heat-shock, only  
251 followed by 2h at 20°C) (Snoek et al. 2017). Each of these experiments was

Sterken et al. *Let-60* eQTL

252 compared at FDR = 0.05 to the eQTL mapped at FDR = 0.10 in the *let-60(gf)*  
253 sensitized miRILs.

254 Since the dataset of Rockman *et al.*, 2010 was used for expanding the genetic  
255 map, analysis was also conducted excluding this study. The reason for excluding this  
256 set from analysis was that a bias could be introduced in overlap with the *cis*-eQTL.  
257 The main conclusion that *cis*-eQTL are less unique than *trans*-eQTL also stands with  
258 this analysis. Excluding the Rockman *et al.* data, resulted in detection of 999/1516  
259 (65.9%) *cis*-eQTL and 171/898 (19.0%) *trans*-eQTL present in previous experiments.

260

### 261 **Enrichment analysis**

262 Gene group enrichment analysis was done using a hypergeometric test on the  
263 unique genes (not on the spots) with the following criteria: Bonferroni corrected p-  
264 value < 0.05, size of the category n>3, size of the overlap n>2.

265 The following databases were used: The Wormbase ([www.wormbase.org](http://www.wormbase.org))  
266 WS220 gene class annotations, the WS256 GO-annotation, anatomy terms,  
267 phenotypes, RNAi phenotypes, developmental stage expression, and disease related  
268 genes (Harris et al. 2014); the MODENCODE release 32 transcription factor binding  
269 sites ([www.modencode.org](http://www.modencode.org); (Gerstein et al. 2010; Niu et al. 2011)), which were  
270 mapped to transcription start sites (according to (Tepper et al. 2013)); the KEGG  
271 pathway release 65.0 (Kyoto Encyclopedia of Genes and Genomes  
272 ([www.genome.jp/kegg/](http://www.genome.jp/kegg/)) (Ogata et al. 1999).

273

### 274 ***amx-2* allelic comparison**

275 Four independent transgenic, 2 of *amx-2(lf);Si[amx-2(Bristol)];let-60(gf)* and 2 of *amx-*  
276 *2(lf);Si[amx-2(Hawaii)];let-60(gf)*, were used to investigate the effect of the *amx-2*

Sterken et al. *Let-60* eQTL

277 CB4856 and N2 allele on gene expression. These lines contain single-copy  
278 insertions on LGII of the N2- or CB4856-alleles of *amx-2* in an N2 *amx-2(lf); let-60(gf)*  
279 background. The effect of different alleles on gene expression was measured by  
280 micro-array for each independent transgenic and compared to the effects of the  
281 eQTL mapping to the *trans*-band closest to the position of *amx-2*.

282

### 283 **Connectivity network analysis**

284 To investigate the connectivity and function of the affected genes we used WormNet  
285 (version 3) (Cho et al. 2014) and GeneMania plugging for Cytoscape (version 3.4.0  
286 (Montoyo et al. 2010; Shannon et al. 2003)). WormNet was used to investigate  
287 enrichment in connectivity within groups of genes. For example, groups of co-  
288 expressed genes mapping to the same *trans*-band were assumed to share the same  
289 regulator. Further evidence for this co-expression and regulation can be found in  
290 WormNet if these genes are significantly more connected than by chance.  
291 GeneMania was used to find the closest neighbours, those genes with the most  
292 connections, of *amx-2* and *let-60* to locate the genes by which *amx-2* modifies RAS  
293 signalling.

294 **RESULTS**

295

296 **Expression-QTL mapping in miRIL population**

297 Using microarrays, we measured gene expression levels in 33 miRILs sensitized for  
298 RAS/MAPK signalling by introgression of a *let-60* (gf) mutation in a segregated  
299 N2/CB4856 genetic background (**Supplemental figure S2**) (Schmid et al. 2015).  
300 Analysis of the statistical power of this population with a genetic map of 247  
301 informative markers showed that we can detect 77% of the QTLs explaining 40% of  
302 the variation (**Supplemental table S4**). We detected 2303 genes (represented by  
303 3226 array spots) with at least one eQTL (FDR=0.1,  $-\log_{10}(p) > 3.2$ ; **Table 1, Figure**  
304 **1, Supplemental table S5**). For 1516 of these genes a *cis*-eQTL was found,  
305 indicating regulation from within a window of 2 Mb around the affected gene. For 898  
306 genes a *trans*-eQTL was found, showing regulation more distal from the affected  
307 transcript. Most *cis*-eQTL (1074 out of 1516; ~71%) show a positive effect for the N2  
308 allele, whereas for the *trans*-eQTL the positive and negative effects were equally  
309 distributed over the N2- and CB4856-allele (**Table 1**).

310 The eQTL were not equally distributed across the genome. We detected  
311 different clustered *trans*-eQTL as “hotspots” or *trans*-bands. *Trans*-bands are  
312 frequently found in eQTL experiments and indicate a locus with one or multiple  
313 alleles affecting the expression of many distant genes. To identify the *trans*-bands in  
314 our experiment, we calculated the chance of overrepresentation of *trans*-eQTL per  
315 marker, using a Poisson distribution (as in (Rockman et al. 2010)). This method was  
316 applied per 1 Mb window on the genome. The *trans*-bands that were identified in  
317 adjacent windows were merged. This way, 6 *trans*-bands were identified (Poisson  
318 distribution,  $p < 0.001$ ; **Table 2**).

319

### 320 **Specificity of *let-60* eQTL**

321 By comparative analysis of different environments and/or populations, it has been  
322 shown that eQTL *trans*-bands can be population, environment, or age specific (for  
323 example, see (Li et al. 2006)). To investigate if and which eQTL are specifically found  
324 in the sensitized miRILs, we compared the detected eQTL to the eQTL found in non-  
325 sensitized RILs (Li et al. 2006; Vinuela et al. 2010; Rockman et al. 2010; Snoek et al.  
326 2017). These datasets contain eQTL mapped over 9 different conditions (see  
327 Methods), which were compared at an FDR = 0.05 with the eQTL mapped in the  
328 sensitized miRILs. We found that ~73% (1112 out of 1516) of the genes with a *cis*-  
329 eQTL in the miRIL population were also found in one or more of the previous studies  
330 (**Figure 2, Supplemental figure S3B**). When taking the allelic effects in account, of  
331 the 1074 genes with a *cis*-eQTL that gave a higher expression linked to the N2 allele,  
332 816 (~76%) were found back in other studies. For the 456 genes with a *cis*-eQTL that  
333 gave a higher expression linked to the CB4856 allele, 310 (~68%) were found back in  
334 other studies. This shows that the majority of the *cis*-eQTL detected in the miRIL  
335 population can be also be detected in other experiments.

336 In contrast, for the *trans*-eQTL detected in our experiment - not taking location  
337 into account - we only found ~24% (214 out of 898) of the genes had a *trans*-eQTL in  
338 at least one of the previous experiments (**Figure 2 and Supplemental figure S3C**).  
339 In order to further compare the *trans*-eQTL overlap, it was investigated whether the  
340 *trans*-eQTL co-localized on the same chromosome. Using that criterion, ~9% (80 out  
341 of 898) of the genes with a *trans*-eQTL were found in a previous study. This showed  
342 that the majority of the genes with a *trans*-eQTL were specifically detected in the  
343 sensitized miRIL population. This observation also extended to the *trans*-bands. To



Sterken et al. *Let-60* eQTL

344 test the specificity of the *trans*-bands, we counted the number of times a gene within  
345 the miRIL *trans*-bands had a co-locating *trans*-eQTL in other studies. All *trans*-bands  
346 were found to be specific for the *let-60(gf)* miRILs (Poisson distribution,  $p < 0.001$ ;  
347 **table 2**). This provides additional evidence that these *trans*-bands are specific for  
348 interaction between the genetic background and the *let-60(gf)* mutation.

349 Together, we identified 404 genes with a novel *cis*-eQTL and 684 genes with a  
350 novel *trans*-eQTL (**Figure 3A**). As our method can be biased by the way the genetic  
351 map was constructed (using expression differences, see **Materials and methods**),  
352 we conducted the same analyses using the original FLP map. Our conclusions on the  
353 detection of novel *cis*- and *trans*-eQTL were not affected depending on the genetic  
354 map (for details, see **Supplemental text S1**). This implies that a substantial majority  
355 of *trans*-eQTL showed a mutation dependent induction. This is also illustrated by  
356 showing the eQTL that were consistently found (**Figure 3B**), which mainly shows the  
357 *cis*-diagonal and a few *trans*-eQTL.

358

### 359 **Functional allelic differences**

360 Enrichment analysis on groups of genes with an eQTL mapping to the same *trans*-  
361 band can be used to uncover the functional consequences of the allelic variation at  
362 the selected locus. Enrichment analysis was done on GO-terms, KEGG, Anatomy-  
363 terms, disease phenotypes, development expression, phenotypes, transcription-  
364 factor binding sites, and RNAi phenotypes (**Supplemental table S6**).

365 The genes with *cis*-eQTL were enriched for genes in the gene classes: *math*,  
366 *bath*, *btb*, and *fbx*. These classes contain many genes that function as intra- and  
367 extra-cellular sensors involved in gene-environment interactions and were found  
368 highly polymorphic between N2 and CB4856, but also between other wild-isolates

Sterken et al. *Let-60* eQTL

369 (Thompson et al. 2015; Volkers et al. 2013). The genes with *trans*-eQTL with higher  
370 expression linked to CB4856 loci were enriched for genes related to development in  
371 general and gonad development specifically. These enrichments were likely to stem  
372 from the *trans*-bands on chromosome I:12-15 Mb and II: 3-6 Mb specifically. These  
373 two *trans*-bands mainly consisted of genes with a higher expression linked to the  
374 CB4856 locus and were enriched for genes associated with reproduction and/or  
375 development.

376

### 377 **Allelic effect of *amx-2* on gene expression**

378 The *trans*-band on chromosome I:12-15Mb co-localizes with a previously identified  
379 QTL for vulval induction, for which *amx-2* was identified as polymorphic regulator  
380 (Schmid et al. 2015). To determine if the *trans*-eQTL mapping to the *amx-2* locus  
381 were affected by the allelic difference between N2 and CB4856, we measured gene  
382 expression in transgenic lines containing single-copy insertions of either allele of  
383 *amx-2* in a *amx-2(lf); let-60(gf)* background. For each allele, two independent strains  
384 were created and the transcriptome was measured using microarrays (Schmid et al.  
385 2015). From these measurements the allelic effects of *amx-2* on gene expression  
386 were determined and compared with the *trans*-eQTL effects measured in the miRIL  
387 population. These were found to correlate significantly (**Supplemental figure S4**;  $R^2$   
388  $\sim 0.34$ ;  $p < 2 \cdot 10^{-8}$ ). Specifically, 61 of the 77 genes with an eQTL co-localizing with  
389 *amx-2* showed the same allelic effect in the transgenic strains carrying the two *amx-2*  
390 variants. Furthermore, analysis of the effect directions showed that the *trans*-band on  
391 chromosome I (**Table 2**) consists of two separate *trans*-bands (see **Supplemental**  
392 **figure S4**); the first co-localizing with *amx-2* and the second lying more distal on  
393 chromosome I.

394 To further investigate the mechanism by which *amx-2* modifies *let-60*  
395 Ras/MAPK signalling, we used the well-established *C.elegans* connectivity networks  
396 WormNet (Cho et al. 2014) and GeneMANIA (Montejo et al. 2010; Shannon et al.  
397 2003). WormNet showed that the genes mapping to the *amx-2* locus were more  
398 connected than expected by chance ( $p < 3 \cdot 10^{-13}$ ). Visualising the connectivity of  
399 these genes together with *amx-2* and *let-60* using GeneMANIA identified one major  
400 gene-cluster (**Figure 4**). As expected, genes affected by the *trans*-band were found  
401 to be highly connected and at the core of the cluster, indicating a shared biological  
402 function. In contrast, 25% of the genes with a *cis*-eQTL at the *trans*-band location  
403 were not connected and those that were connected had only one to three  
404 connections. Interestingly, *let-60* was at the core of the cluster of genes with a *trans*-  
405 eQTL, whereas *amx-2* was at the periphery. We found six genes, *lin-40* (also called  
406 *egr-1*), *egl-27*, *trr-1*, *pbrm-1*, *ceh-26* (also called *pros-1*) and *emb-5*, previously found  
407 to have a genetic interactions with *let-60* (Lee et al. 2010; Lehner et al. 2006; Byrne  
408 et al. 2007). These could be the genes through which *amx-2* affects RAS/MAPK  
409 signalling.

## 410 **DISCUSSION**

411 Ras/MAPK signalling is strongly affected by gain of function mutations in *let-60* as  
412 shown by the strong effect on vulva development in *C. elegans* (Beitel et al. 1990).  
413 The allele *let-60(n1046)* hyperactivates Ras/MAPK signalling, leading to a multivulva  
414 phenotype and the differentiation of more than three vulval precursor cells. In the  
415 miRIL population, which carries a *let-60(n1046)* mutation in a segregating  
416 N2/CB4856 background, a wide range of vulval induction (VI) was found. While in the  
417 full N2 background VI of the *let-60(gf)* mutation was 3.7 induced vulval cells on  
418 average, the VI in the miRIL population varied from 3.0 to 5.7 induced vulval cells,  
419 illustrating the strong modulatory effect of the genetic background on the mutant  
420 phenotype (Schmid et al. 2015). By measuring gene expression across the different  
421 miRILs, we gained insight into the genetic background effects on the transcriptional  
422 architecture of *let-60(gf)* sensitized miRILs.

423

### 424 **The *let-60(gf)* mutation affects gene expression *in trans***

425 To our knowledge this is the first study where a mutation, introgressed into a panel of  
426 RILs, is used to investigate the genetic architecture of transcript variation by eQTL  
427 analysis. Unfortunately, the direct effect of the *let-60(gf)* allele on the genetic  
428 background is impossible to determine in this population, since there is no population  
429 with the same genetic background lacking the *let-60(gf)* mutation. Therefore we used  
430 previously published eQTL studies in *C. elegans* as estimation for the miRIL specific  
431 eQTL. We found that introgression of *let-60(gf)* leads to an overrepresentation of  
432 specific *trans*-eQTL, while the *cis*-eQTL showed a high overlap with eQTL observed  
433 in the absence of *let-60(gf)* allele. Although our study is based on a relatively small  
434 number of strains, the verification of identified eQTL against other eQTL studies in *C.*

Sterken et al. *Let-60* eQTL

435 *C. elegans* shows that at least 73% of the identified *cis*-eQTL were discovered before.  
436 This supports our power analysis, showing that our study has enough statistical  
437 power to detect eQTL with larger effects.

438         Given that *cis*-eQTL are most likely genes with polymorphisms in or near the  
439 gene itself, it is expected that most are discovered before. In support, most of the *cis*-  
440 eQTL genes show higher expression in N2, making it likely that a relatively large part  
441 of these *cis*-eQTL indeed stem from polymorphisms causing mis-hybridization on the  
442 microarrays instead of true expression differences (Rockman et al. 2010; Alberts et  
443 al. 2007). Therefore, most of the *cis*-eQTL will not represent gene expression  
444 changes that can be linked to the *let-60(gf)* mutation in the miRIL population. That is  
445 also exactly what we found when constructing the regulatory network by connecting  
446 *cis*- and *trans*-eQTL to *let-60*; *cis*-eQTL are only loosely - or not at all - connected to  
447 the network.

448         Compared to *cis*-eQTL, the *trans*-eQTL identified in the miRIL population were  
449 hardly found in previous eQTL studies in *C. elegans* (only 24% was detected  
450 previously). However, it should be noted that *trans*-eQTL on themselves are also  
451 environment- and age-dependent (Vinuela et al. 2010; Francesconi and Lehner  
452 2014; Li et al. 2006; Snoek et al. 2017), therefore it possible that not all the *trans*-  
453 eQTL are indeed *let-60* induced. For example, they might be age-related (Vinuela et  
454 al. 2010). Another aspect that influences detection of *trans*-eQTL is the power of the  
455 study, as *trans*-eQTL explain less of the variation than *cis*-eQTL. However, by being  
456 stringent in the overlap (requiring only one match over nine different environments) it  
457 is likely that most of these eQTL would be detected. This makes it likely that many of  
458 the identified *trans*-eQTL are specifically linked to an effect of the *let-60(gf)* mutation  
459 in the genetic background. This extends the idea that environmental perturbations

Sterken et al. *Let-60* eQTL

460 can reveal additional genetic variation (Li et al. 2008), by perturbing the genetic  
461 environment (Kammenga et al. 2008; Duveau and Felix 2012; Schmid et al. 2015).  
462 The advantage of mutational perturbation and perturbation via induced responses in  
463 eQTL studies (for example, see (Snoek et al. 2012) and (Snoek et al. 2017)), is the  
464 placement of the transcriptional response into a context. Yet, as the phenotypic  
465 influence of *let-60(gf)* in the miRIL population was investigated before, we know it  
466 results in a phenotype In the study presented here, the context is the *let-60(gf)*  
467 mutation, which results in a phenotype where vulval induction is affected. It is  
468 therefore interesting that the *trans*-bands co-localize with QTL mapped for index VI  
469 index in the same miRIL population: *trans*-band I:12-15 Mb overlaps with QTL1b,  
470 *trans*-band II:7-8 Mb with QTL2, and *trans*-band V:13-14 Mb with QTL3 (Schmid et al.  
471 2015). This co-localization adds to the plausibility that the novel *trans*-eQTL detected  
472 in the miRIL population are indeed due to the *let-60(gf)* mutation and its interaction  
473 with the genetic background.

474

#### 475 **A *trans*-band on chromosome I links to variation in the *amx-2* gene**

476 Using the same *let-60(gf)* miRIL population, we have previously identified the  
477 polymorphic monoamine oxidase *amx-2* as a background modifier that negatively  
478 regulates the Ras/MAPK pathway and the VI phenotype (Schmid et al. 2015). The  
479 *trans*-band on chromosome I co-localizes with the *amx-2* QTL for VI (QTL1b,  
480 (Schmid et al. 2015)), it is therefore possible that the modifier *amx-2* also affects  
481 gene expression. More specifically, the *trans*-band on chromosome I originally  
482 identified in the miRILs consists of two separate regulatory loci. The N2 allelic effect  
483 at the *amx-2* locus is mostly negative, whereas for the more distal part of the original  
484 *trans*-band the N2 allelic effect is mostly positive. Yet, it seems unlikely that *amx-2* is

Sterken et al. *Let-60* eQTL

485 a direct gene expression effector. As *amx-2* is a mitochondrial monoamine oxidase  
486 type A, a catalyser of neuropeptide oxidative deamination, it will probably not  
487 influence gene expression directly (Tipton et al. 2004). Placing the *trans*-eQTL of the  
488 *amx-2 trans*-band in a gene interaction network supports this line of reasoning:  
489 although *let-60* is in the centre of the *trans*-eQTL, *amx-2* is only peripheral.

490 By measuring gene expression in transgenic lines expressing the N2 or  
491 CB4856 allelic variants of *amx-2* in an *amx-2(lf);let-60(gf)* genetic background, we  
492 attempted to link the allelic effect of *amx-2* to the *trans*-eQTL mapping to the *amx-2*  
493 locus. As a significant correlation between the expression differences in the  
494 transgenic lines and the *trans*-eQTL effects mapping to the *amx-2* locus is found,  
495 these *trans*-eQTL can be confidentially linked to allelic variation in *amx-2*. As  
496 discussed before, it is unlikely that *amx-2* is the direct cause of the transcriptional  
497 variation, but rather acts through an indirect mechanism. One route of effect might be  
498 through *amx-2* mediated degradation of serotonin (5-HT) to 5-HIAA, which both affect  
499 VI (as discussed in (Schmid et al. 2015)). Subsequently, the affected VI might result  
500 in different gene expression levels. It is interesting to remark that the *amx-2 trans*-  
501 band is characterized by down-regulation of expression related to the N2 genotype  
502 as well as a decreased VI for that genotype.

503 As this places *amx-2* in the causal chain of events, we think it is most likely  
504 that *amx-2* does not affect gene expression directly but via another gene involved in  
505 Ras/MAPK signalling, possibly directly linked to *let-60*. Therefore, we hypothesize  
506 that *amx-2* affects Ras/MAPK signalling via one or more of the six previously found  
507 genes with a genetic interaction with *let-60*: *lin-40* (also called *egr-1*), *egl-27*, *trr-1*,  
508 *pbrm-1*, *ceh-26* (also called *pros-1*), and *emb-5*. For example, *egl-27* and *lin-40* both  
509 have a *trans*-eQTL mapping to the *amx-2* locus, and are the two MTA (metastasis-

Sterken et al. *Let-60* eQTL

510 associated protein) homologs found in *C. elegans* (Herman et al. 1999; Ch'ng and  
511 Kenyon 1999; Solari et al. 1999). The proteins act in the NURD chromatin  
512 remodelling complex, which has previously been shown to antagonize Ras-induced  
513 vulval development (Solari and Ahringer 2000). Chromatin remodelling provides a  
514 more likely mechanism of action than the molecular role of *amx-2* itself.

515

### 516 **Implications for understanding the Ras/MAPK pathway**

517 How can our results help a better understanding of the Ras/MAPK pathway? Our  
518 previous investigation on VI in the miRIL population resulted in the identification of  
519 three QTL harbouring polymorphic Ras-signalling modifiers. Expanding our research  
520 to genetic variation affecting gene expression in a *let-60(gf)* sensitized RIL population  
521 uncovered six *trans*-bands (or 'eQTL hotspots'). As mentioned before, these *trans*-  
522 bands overlap with the QTL mapped for VI in Schmid *et al.*, 2015. There are three  
523 more *trans*-bands that do not overlap with QTL for VI, but are also likely to represent  
524 modifier loci of the Ras/MAPK pathway, possibly underlying other Ras/MARK  
525 associated phenotypic differences.

526         The main advantage of studying the genetic architecture of gene expression in  
527 the miRIL population is that it creates more insight in the genes and pathways  
528 affected by allelic variation acting on the Ras/MAPK pathway. It identifies hidden  
529 genetic variation; genetic variants that are unlocked under altered environmental  
530 conditions or when the genetic background is modified (for a review, see (Paaby and  
531 Rockman 2014) and (Kammenga 2017)). Identification of background modifiers of  
532 disease pathways is important for gaining insight into individual based differences of  
533 disease contraction. Mutant phenotypes can be strongly affected by the genetic  
534 background. For example, large variation in traits between different backgrounds in



Sterken et al. *Let-60* eQTL

535 many different mutated genes has been observed in *C. elegans* (Vu et al. 2015;  
536 Paaby et al. 2015; Duvéau and Felix 2012). A major discovery was that this variation  
537 in mutant phenotypes could be predicted from gene expression variation (Vu et al.  
538 2015). Those results are in line with the discovery of *trans*-bands at the location of  
539 each VI QTL. Furthermore, results presented here (and previously (Elvin et al. 2011))  
540 link variations in individual gene expression levels to enhancing or diminishing the  
541 severity of a Mendelian disorder caused by a so-called “major gene”. More  
542 specifically, these differences seem to stem from a couple of modifier loci harbouring  
543 polymorphic regulators.

544         So far, we only explored the *amx-2 trans*-band, and it is likely that the other  
545 loci also contain polymorphic modifiers of the Ras/MAPK pathway. Genetic  
546 perturbation by *let-60(gf)* leads to a strong increase in specific *trans*-acting eQTL  
547 organized over six *trans*-bands, thus supporting the involvement of multiple genes as  
548 modifiers. Given the fact that we detected *amx-2* both via an indirect transcriptional  
549 response, but also mechanistically, we feel confident that the detected loci indeed  
550 harbour other modifiers. A single gene mutation apparently has the capacity to unlock  
551 a huge number of novel interactions controlled by many genes across different  
552 genetic backgrounds, further studies should be conducted to identify and  
553 characterize these underlying polymorphic regulators.

554 **CONCLUSION**

555 Introgression of a mutation in a segregating genetic background allows for  
556 identification of polymorphic modifier loci. By introducing a *let-60(gf)* mutation in a  
557 N2/CB4856 genetic background, in the form of the miRIL population and measuring  
558 the transcriptome, we identified six *trans*-bands specific for the *let-60(gf)* miRIL  
559 population. The majority of *trans*-eQTL are specific for the miRIL population, showing  
560 that genetic variation in gene expression can be specifically modified by a  
561 background mutation. Therefore, genetic perturbation can be viewed as analogous to  
562 environmental perturbation, which also results in specific *trans*-eQTL. We  
563 demonstrated the involvement of *amx-2* and the allelic variation between N2 and  
564 CB4856 in gene expression variation originating from chromosome I. Yet, we think it  
565 is unlikely that *amx-2* directly affects gene expression variation. Instead, we prefer  
566 the hypothesis that allelic variation in *amx-2* indirectly affects gene expression,  
567 possibly through the NURD complex.

Sterken et al. *Let-60* eQTL

568 **AVAILABILITY OF DATA AND MATERIALS**

569 All strains used can be requested from the authors. The transcriptome datasets  
570 generated and the mapped eQTL profiles can be interactively accessed via  
571 <http://www.bioinformatics.nl/EleQTL>. Moreover, the raw transcriptomic data is also  
572 available at ArrayExpress (E-MTAB-5856)

573

574

Sterken et al. *Let-60* eQTL

575 **SUPPLEMENTAL DATA**

576

577 **Supplemental text S1:** Text explaining the analysis of eQTL in the original FLP-  
578 based genetic map.

579

580 **Supplemental figure S1:** Between-marker correlations of the genotypes in the *let-*  
581 *60(gf)* sensitized RIL population. Per marker pair, the correlations are shown in a  
582 heat-plot. On the x-axis marker 1 and on the y-axis marker 2 is shown. Most in-  
583 between chromosome correlations are  $|R| < 0.6$ . However, between chromosome  
584 I:2.6-2.7 Mb and chromosome IV3.6-4.2 Mb and chromosome II:3.1-3.4 Mb and  
585 chromosome III:0.0-1.4 Mb higher correlations were detected.

586

587 **Supplemental figure S2:** A figure of the genetic map of the sensitized RIL  
588 population. On the x-axis the physical position on the genome is given and on the y-  
589 axis the genotypes used in this study. The genotypic origin at each genomic position  
590 is indicated by colour (orange for N2 and blue for CB4856).

591

592 **Supplemental figure S3:** Comparison of eQTL mapped in the *let-60(gf)* miRIL  
593 population with previous genetical genomics studies in *C. elegans*. **(A)** eQTL effect  
594 sizes per condition per study, split out for *cis*- and *trans*-eQTL. **(B)** Constitutively  
595 found *cis*-eQTL. On the x-axis the QTL position in the sensitized RIL population is  
596 shown, on the y-axis the QTL position in the compared experiment is shown.  
597 Therefore, a point on the diagonal indicates a QTL mapped to the same position. The  
598 lines indicate the 1.5 LOD-drop confidence interval of the QTL position. **(C)**  
599 Constitutively found *trans*-eQTL, annotations are the same as in **(B)**.

Sterken et al. *Let-60* eQTL

600

601 **Supplemental figure S4:** Comparison between eQTL effects on chromosome I and  
602 the difference between N2 or CB4856 allelic variants of *amx-2* in an *amx-2(lf);let-*  
603 *60(gf)*. **(A)** Position of the peaks of the eQTL on chromosome I and eQTL N2 allelic  
604 effects. Red to blue gradient shows the log<sub>2</sub> expression ratio between the allelic  
605 variants of *amx-2* in the transgenic lines. QTLs detected in *Schmid et al.* 2015  
606 (QTL1a and QTL1b) and newly identified sub-*trans*-band QTL1c are shown on top.  
607 Position of *amx-2* is shown in red. **(B)** Direct comparison of the eQTL effects and  
608 transgene effects for all eQTL on chromosome I. **(C)** Direct comparison of the eQTL  
609 effects and transgene effects for eQTL in QTL1b/*amx-2* locus. **(D)** Direct comparison  
610 of the eQTL effects and transgene effects for eQTL in QTL1c.

611

612 **Supplemental table S1:** The genetic map of the *let-60(gf)* sensitized RIL population  
613 and the two parental strains used in the gene expression experiment. The CB4856  
614 genotype is denoted with -1 and the N2 genotype is denoted with 1.

615

616 **Supplemental table S2:** FLP makers used for this study

617

618 **Supplemental table S3:** The assigned gene expression markers, organized per  
619 strain per sample per marker. The number of spots on which the correlation was  
620 based and the correlation value is given, as well as the assigned genotype.

621

622 **Supplemental table S4:** A table summarizing the results of the power analysis. Per  
623 simulated peak size (sigma) the variation explained is shown. These peaks were  
624 simulated 10 times at each marker location in noise simulated by a standard normal

Sterken et al. *Let-60* eQTL

625 distribution. The number of correctly detected, false-positive, and undetected QTL is  
626 shown (at  $-\log_{10}(p) > 3.2$ ). Also the fractions of the total are given. The quantiles of  
627 the effect estimation are listed as well as the quantiles of the 'detected peak'-'true  
628 peak' distance.

629

630 **Supplemental table S5:** A table with the eQTL mapped in the *let-60(gf)* miRIL  
631 population. eQTL are listed per trait (Spot) and QTL type. The peak location and its  
632 confidence interval are given (based on a 1.5 LOD drop), the peak significance, its  
633 effect, and the variation explained by the QTL. The effect direction indicates higher in  
634 N2 (positive numbers) or higher at CB4856 (negative numbers) loci. Furthermore,  
635 information about the affected gene represented by the microarray spot is shown  
636 (name, and location).

637

638 **Supplemental table S6:** eQTL enrichments split out by eQTL type. The genotype  
639 behind an eQTL class (e.g. *cis\_CB4856*) indicates it concerns an eQTL effect with  
640 high expression associated with that genotype. The database used for enrichment  
641 (Annotation), the category (Group), and the number of genes on the array that are in  
642 the group (Genes\_in\_group) are indicated. Furthermore, the overlap with the cluster  
643 (Overlap) and the Bonferroni-corrected significance of that overlap are shown.

Sterken et al. *Let-60* eQTL

644 **ACKNOWLEDGEMENTS**

645 The authors thank Harm Nijveen for making our data available in EleQTL.

646

647 **FUNDING**

648 MR, TS, AH, and JEK were funded by the European Community's Health Seventh

649 Framework Programme (FP7/2007-2013) under grant 222936. LBS was funded by

650 ERASysbio-plus ZonMW project GRAPPLE - Iterative modelling of gene regulatory

651 interactions underlying stress, disease and ageing in *C. elegans* (project 90201066)

652 and The Netherlands Organization for Scientific Research (project no. 823.01.001).

653

654 **AUTHOR CONTRIBUTIONS**

655 AH, JK, and LBS conceived and designed the experiments. JAGR, MR, and TS

656 conducted the experiments. MGS, LBP, and LBS conducted transcriptome and main

657 analyses. MGS, AH, JK, and LBS wrote the manuscript.

658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728

References

- Alberts, R., P. Terpstra, Y. Li, R. Breitling, J.P. Nap *et al.*, 2007 Sequence polymorphisms cause many false cis eQTLs. *PLoS One* 2 (7):e622.
- Beitel, G.J., S.G. Clark, and H.R. Horvitz, 1990 *Caenorhabditis elegans* ras gene *let-60* acts as a switch in the pathway of vulval induction. *Nature* 348 (6301):503-509.
- Benjamini, Y., and D. Yekutieli, 2001 The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* 29 (4):1165-1188.
- Byrne, A.B., M.T. Weirauch, V. Wong, M. Koeva, S.J. Dixon *et al.*, 2007 A global analysis of genetic interactions in *Caenorhabditis elegans*. *J Biol* 6 (3):8.
- Ch'ng, Q., and C. Kenyon, 1999 *egl-27* generates anteroposterior patterns of cell fusion in *C.elegans* by regulating Hox gene expression and Hox protein function. *Development* 126 (15):3303-3312.
- Chandler, C.H., S. Chari, D. Tack, and I. Dworkin, 2014 Causes and consequences of genetic background effects illuminated by integrative genomic analysis. *Genetics* 196 (4):1321-1336.
- Cho, A., J. Shin, S. Hwang, C. Kim, H. Shim *et al.*, 2014 WormNet v3: a network-assisted hypothesis-generating server for *Caenorhabditis elegans*. *Nucleic Acids Res* 42 (W1):W76-W82.
- Cowley, S., H. Paterson, P. Kemp, and C.J. Marshall, 1994 Activation of Map Kinase Kinase Is Necessary and Sufficient for Pc12 Differentiation and for Transformation of Nih 3t3 Cells. *Cell* 77 (6):841-852.
- Duveau, F., and M.A. Felix, 2012 Role of pleiotropy in the evolution of a cryptic developmental variation in *Caenorhabditis elegans*. *PLoS Biol* 10 (1):e1001230.
- Elvin, M., L.B. Snoek, M. Frejno, U. Klemstein, J.E. Kammenga *et al.*, 2011 A fitness assay for comparing RNAi effects across multiple *C. elegans* genotypes. *BMC Genomics* 12:510.
- Francesconi, M., and B. Lehner, 2014 The effects of genetic variation on gene expression dynamics during development. *Nature* 505 (7482):208-211.
- Gerstein, M.B., Z.J. Lu, E.L. Van Nostrand, C. Cheng, B.I. Arshinoff *et al.*, 2010 Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science* 330 (6012):1775-1787.
- Gokhale, R.H., and A.W. Shingleton, 2015 Size control: the developmental physiology of body and organ size regulation. *Wiley Interdiscip Rev Dev Biol* 4 (4):335-356.
- Han, M., and P.W. Sternberg, 1990 *let-60*, a gene that specifies cell fates during *C. elegans* vulval induction, encodes a ras protein. *Cell* 63 (5):921-931.
- Harris, T.W., J. Baran, T. Bieri, A. Cabunoc, J. Chan *et al.*, 2014 WormBase 2014: new views of curated biology. *Nucleic Acids Res* 42 (Database issue):D789-793.
- Herman, M.A., Q. Ch'ng, S.M. Hettenbach, T.M. Ratliff, C. Kenyon *et al.*, 1999 EGL-27 is similar to a metastasis-associated factor and controls cell polarity and cell migration in *C.elegans*. *Development* 126 (5):1055-1064.
- Jansen, R.C., and J.P. Nap, 2001 Genetical genomics: the added value from segregation. *Trends in Genetics* 17 (7):388-391.
- Jiang, H.S., and Y.C. Wu, 2014 LIN-3/EGF promotes the programmed cell death of specific cells in *Caenorhabditis elegans* by transcriptional activation of the pro-apoptotic gene *egl-1*. *PLoS Genet* 10 (8):e1004513.
- Kammenga, J.E., 2017 The background puzzle: how identical mutations in the same gene lead to different disease symptoms. *FEBS J.*
- Kammenga, J.E., P.C. Phillips, M. De Bono, and A. Doroszuk, 2008 Beyond induced mutants: using worms to study natural variation in genetic pathways. *Trends in Genetics* 24 (4):178-185.
- Kyriakakis, E., M. Markaki, and N. Tavernarakis, 2015 *Caenorhabditis elegans* as a model for cancer research. *Mol Cell Oncol* 2 (2):e975027.
- Lee, I., B. Lehner, T. Vavouri, J. Shin, A.G. Fraser *et al.*, 2010 Predicting genetic modifier loci using functional gene networks. *Genome Res* 20 (8):1143-1153.
- Lehner, B., C. Crombie, J. Tischler, A. Fortunato, and A.G. Fraser, 2006 Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse signaling pathways. *Nature Genetics* 38 (8):896-903.
- Li, Y., O.A. Alvarez, E.W. Gutteling, M. Tijsterman, J. Fu *et al.*, 2006 Mapping determinants of gene expression plasticity by genetical genomics in *C. elegans*. *PLoS Genet* 2 (12):e222.
- Li, Y., R. Breitling, and R.C. Jansen, 2008 Generalizing genetical genomics: getting added value from environmental perturbation. *Trends in Genetics* 24 (10):518-524.
- Mansour, S.J., W.T. Matten, A.S. Hermann, J.M. Candia, S. Rong *et al.*, 1994 Transformation of mammalian cells by constitutively active MAP kinase kinase. *Science* 265 (5174):966-970.
- Montejo, J., K. Zuberi, H. Rodriguez, F. Kazi, G. Wright *et al.*, 2010 GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* 26 (22):2927-2928.
- Niu, W., Z.J. Lu, M. Zhong, M. Sarov, J.I. Murray *et al.*, 2011 Diverse transcription factor binding features revealed by genome-wide ChIP-seq in *C. elegans*. *Genome Res* 21 (2):245-254.
- Ogata, H., S. Goto, K. Sato, W. Fujibuchi, H. Bono *et al.*, 1999 KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 27 (1):29-34.
- Paaby, A.B., and M.V. Rockman, 2014 Cryptic genetic variation: evolution's hidden substrate. *Nat Rev Genet* 15 (4):247-258.
- Paaby, A.B., A.G. White, D.D. Riccardi, K.C. Gunsalus, F. Piano *et al.*, 2015 Wild worm embryogenesis harbors ubiquitous polygenic modifier variation. *Elife* 4.
- Prior, I.A., and J.F. Hancock, 2012 Ras trafficking, localization and compartmentalized signalling. *Semin Cell Dev Biol* 23 (2):145-153.
- Rockman, M.V., S.S. Skrovanek, and L. Kruglyak, 2010 Selection at linked sites shapes heritable phenotypic variation in *C. elegans*. *Science* 330 (6002):372-376.



Sterken et al. *Let-60* eQTL

- 729 Schmid, T., L.B. Snoek, E. Frohli, M.L. van der Bent, J. Kammenga *et al.*, 2015 Systemic Regulation of  
730 RAS/MAPK Signaling by the Serotonin Metabolite 5-HIAA. *PLoS Genet* 11 (5):e1005236.
- 731 Seidel, H.S., M.V. Rockman, and L. Kruglyak, 2008 Widespread genetic incompatibility in *C. elegans* maintained  
732 by balancing selection. *Science* 319 (5863):589-594.
- 733 Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang *et al.*, 2003 Cytoscape: A software environment for  
734 integrated models of biomolecular interaction networks. *Genome Res* 13 (11):2498-2504.
- 735 Smyth, G.K., and T. Speed, 2003 Normalization of cDNA microarray data. *Methods* 31 (4):265-273.
- 736 Snoek, B., M. Sterken, R. Bevers, R. Volkers, A. van't Hof *et al.*, 2017 Contribution Of Trans Regulatory eQTL  
737 To Cryptic Genetic Variation In *C. elegans*. *bioRxiv*.
- 738 Snoek, L.B., I.R. Terpstra, R. Dekter, G. Van den Ackerveken, and A.J. Peeters, 2012 Genetical Genomics  
739 Reveals Large Scale Genotype-By-Environment Interactions in *Arabidopsis thaliana*. *Front Genet*  
740 3:317.
- 741 Snoek, L.B., K.J. Van der Velde, D. Arends, Y. Li, A. Beyer *et al.*, 2013 WormQTL--public archive and analysis  
742 web portal for natural variation data in *Caenorhabditis* spp. *Nucleic Acids Res* 41 (Database  
743 issue):D738-743.
- 744 Solari, F., and J. Ahringer, 2000 NURD-complex genes antagonise Ras-induced vulval development in  
745 *Caenorhabditis elegans*. *Curr Biol* 10 (4):223-226.
- 746 Solari, F., A. Bateman, and J. Ahringer, 1999 The *Caenorhabditis elegans* genes *egl-27* and *egr-1* are similar to  
747 MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic  
748 patterning. *Development* 126 (11):2483-2494.
- 749 Tan, P.B., M.R. Lackner, and S.K. Kim, 1998 MAP kinase signaling specificity mediated by the LIN-1 Ets/LIN-31  
750 WH transcription factor complex during *C. elegans* vulval induction. *Cell* 93 (4):569-580.
- 751 Tepper, R.G., J. Ashraf, R. Kaletsky, G. Kleemann, C.T. Murphy *et al.*, 2013 PQM-1 Complements DAF-16 as a  
752 Key Transcriptional Regulator of DAF-2-Mediated Development and Longevity. *Cell* 154 (3):676-690.
- 753 Thompson, O.A., L.B. Snoek, H. Nijveen, M.G. Sterken, R.J.M. Volkers *et al.*, 2015 Remarkably Divergent  
754 Regions Punctuate the Genome Assembly of the *Caenorhabditis elegans* Hawaiian Strain CB4856.  
755 *Genetics* 200 (3):975-+.
- 756 Tipton, K.F., S. Boyce, J. O'Sullivan, G.P. Davey, and J. Healy, 2004 Monoamine oxidases: certainties and  
757 uncertainties. *Curr Med Chem* 11 (15):1965-1982.
- 758 van der Velde, K.J., M. de Haan, K. Zych, D. Arends, L.B. Snoek *et al.*, 2014 WormQTLHD--a web database for  
759 linking human disease to natural variation data in *C. elegans*. *Nucleic Acids Res* 42 (Database  
760 issue):D794-801.
- 761 Vinuela, A., L.B. Snoek, J.A. Riksen, and J.E. Kammenga, 2010 Genome-wide gene expression regulation as a  
762 function of genotype and age in *C. elegans*. *Genome Res* 20 (7):929-937.
- 763 Volkers, R.J., L.B. Snoek, C.J. Hubar, R. Coopman, W. Chen *et al.*, 2013 Gene-environment and protein-  
764 degradation signatures characterize genomic and phenotypic diversity in wild *Caenorhabditis elegans*  
765 populations. *BMC Biol* 11:93.
- 766 Vu, V., A.J. Verster, M. Schertzberg, T. Chuluunbaatar, M. Spensley *et al.*, 2015 Natural Variation in Gene  
767 Expression Modulates the Severity of Mutant Phenotypes. *Cell* 162 (2):391-402.
- 768 West, M.A., H. van Leeuwen, A. Kozik, D.J. Kliebenstein, R.W. Doerge *et al.*, 2006 High-density haplotyping  
769 with microarray-based expression and single feature polymorphism markers in *Arabidopsis*. *Genome*  
770 *Res* 16 (6):787-795.
- 771 Wu, Y., and M. Han, 1994 Suppression of activated *Let-60* ras protein defines a role of *Caenorhabditis elegans*  
772 *Sur-1* MAP kinase in vulval differentiation. *Genes Dev* 8 (2):147-159.
- 773 Zahurak, M., G. Parmigiani, W. Yu, R.B. Scharpf, D. Berman *et al.*, 2007 Pre-processing Agilent microarray  
774 data. *BMC Bioinformatics* 8:142.
- 775 Zipperlen, P., K. Nairz, I. Rimann, K. Basler, E. Hafen *et al.*, 2005 A universal method for automated gene  
776 mapping. *Genome Biol* 6 (2):R19.
- 777

778 **TABLE AND FIGURES LEGENDS**

779

780 **Figure 1:** eQTL mapping in the *let-60(gf)* miRIL population. On top the *cis-trans* plot  
781 is shown. On the horizontal axes the positions of the eQTL are plotted, along the six  
782 chromosomes. The location of the genes with an eQTL is plotted on the vertical axis.  
783 The black dots represent *cis*-eQTL (lying within 2 Mb of the regulated gene), the  
784 green dots represent *trans*-eQTL (eQTL lying elsewhere). The grey horizontal lines  
785 indicate the confidence interval of the QTL location (based on a 1.5 drop in –  
786  $\log_{10}(p)$ ). The bottom histogram shows the number of eQTL per genomic location.  
787 Note that on chromosome IV hardly any eQTL are mapped, this is because of linkage  
788 of the *let-60(gf)* mutation (which is located at IV:11.7 Mb).

789

790 **Figure 2:** eQTL detected in previous experiments. The eQTL mapped in the *let-*  
791 *60(gf)* RIL population were compared to eQTL mapped in three published *C. elegans*  
792 eQTL studies, representing 9 different conditions (Li et al. 2006; Vinuela et al. 2010;  
793 Rockman et al. 2010; Snoek et al. 2017). The histogram shows how many of the  
794 eQTL mapped in this study were found back in these independent conditions.  
795 Whereas the majority of *cis*-eQTL was mapped in previous studies, only a minority of  
796 the *trans*-eQTL was.

797

798 **Figure 3:** eQTL mapped in the *let-60(gf)* miRIL population compared to other studies  
799 (Li et al. 2006; Vinuela et al. 2010; Rockman et al. 2010; Snoek et al. 2017). As in  
800 **figure 2**, the *cis*- and *trans*-eQTL are plotted. **(A)** eQTL only found in the miRIL  
801 population. **(B)** eQTL found over multiple studies.

802

Sterken et al. *Let-60* eQTL

803 **Figure 4:** Interaction network of genes co-localizing with *amx-2*. Genes with an eQTL  
804 co-localizing with *amx-2* were selected together with *amx-2* (yellow) and *let-60* (pink).  
805 Interactions were obtained from Genemania (Montejo et al. 2010; Shannon et al.  
806 2003). Node colour indicates: *cis* in black and *trans* in green. Node shape indicates  
807 eQTL effect, N2 > CB4856 as a circle and CB4856 > N2 as a square. Node border  
808 colour gradient indicates expression in *amx-2* transgenic lines (blue N2-*amx-2*-allele  
809 > CB4856-*amx-2* allele, red CB4856-*amx-2* allele > N2 *amx-2* allele). Node size  
810 indicates number of edges connected. Edges in blue show co-expression links and  
811 edges in pink show genetic interactions as found by (Lee et al. 2010; Lehner et al.  
812 2006; Byrne et al. 2007). Genes connected by genetic interactions have bold gene  
813 names (as well as *amx-2*).

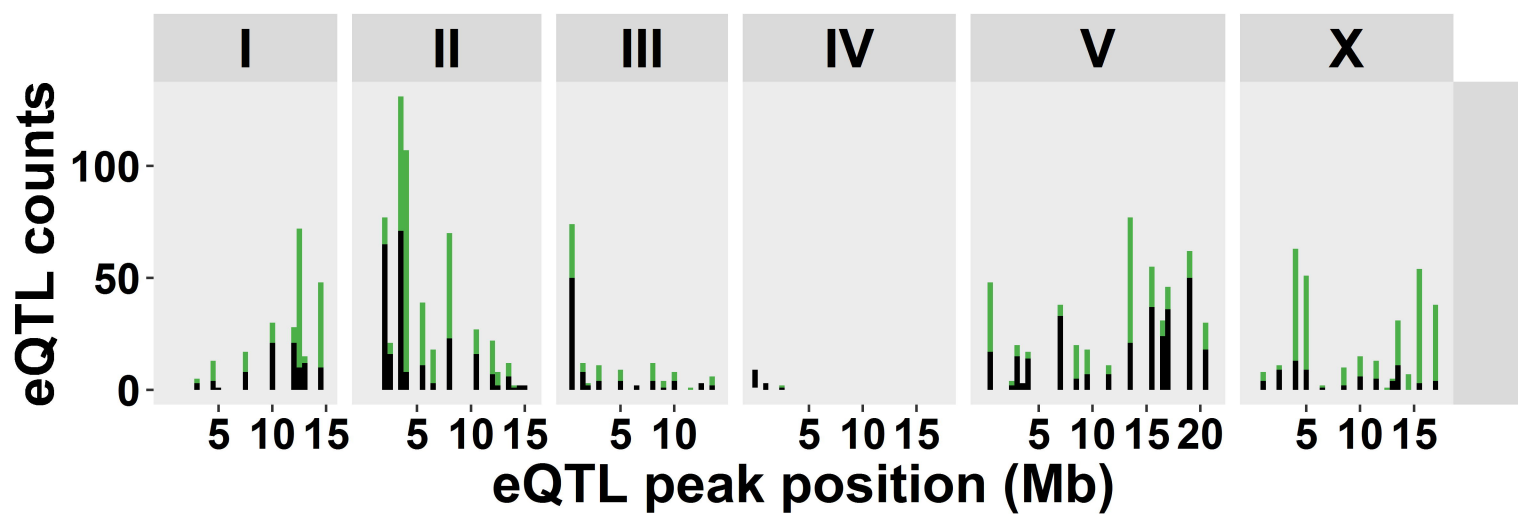
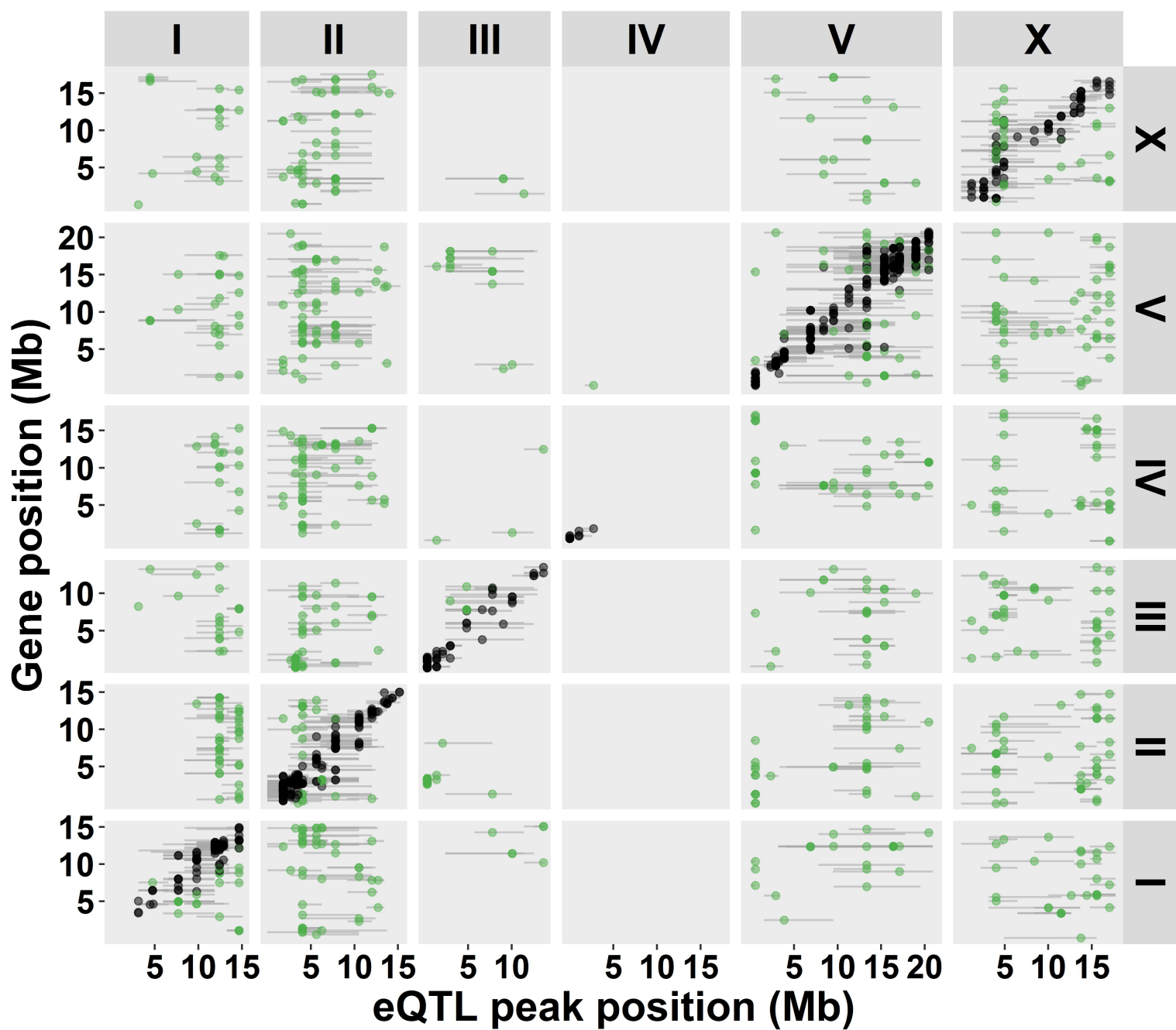
814

815 **Table 1:** The number of genes with an eQTL, in brackets the number of spots.

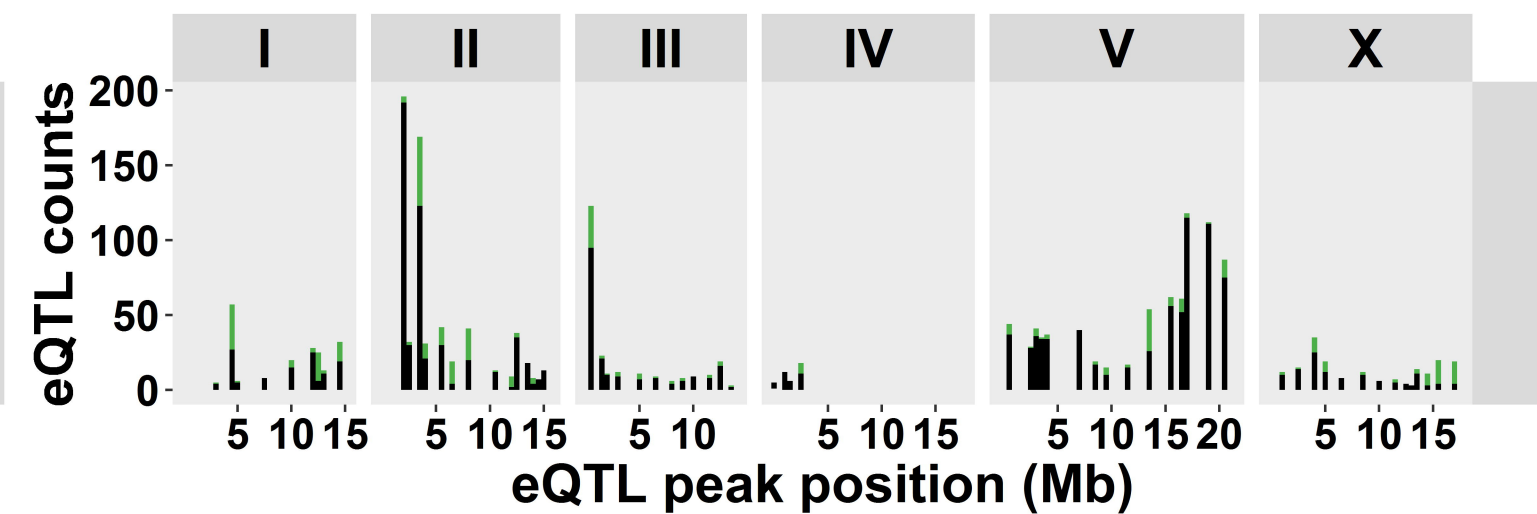
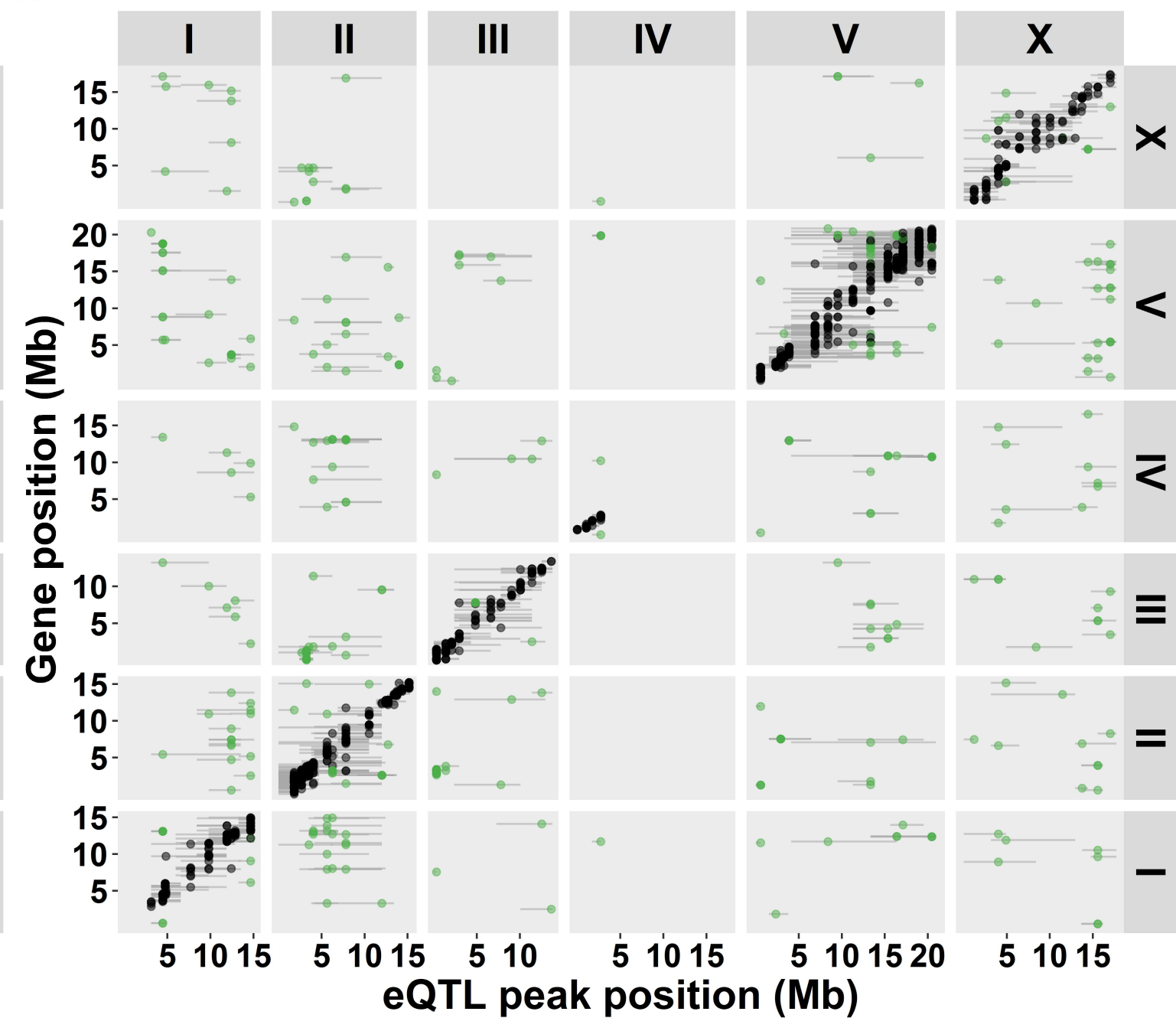
816

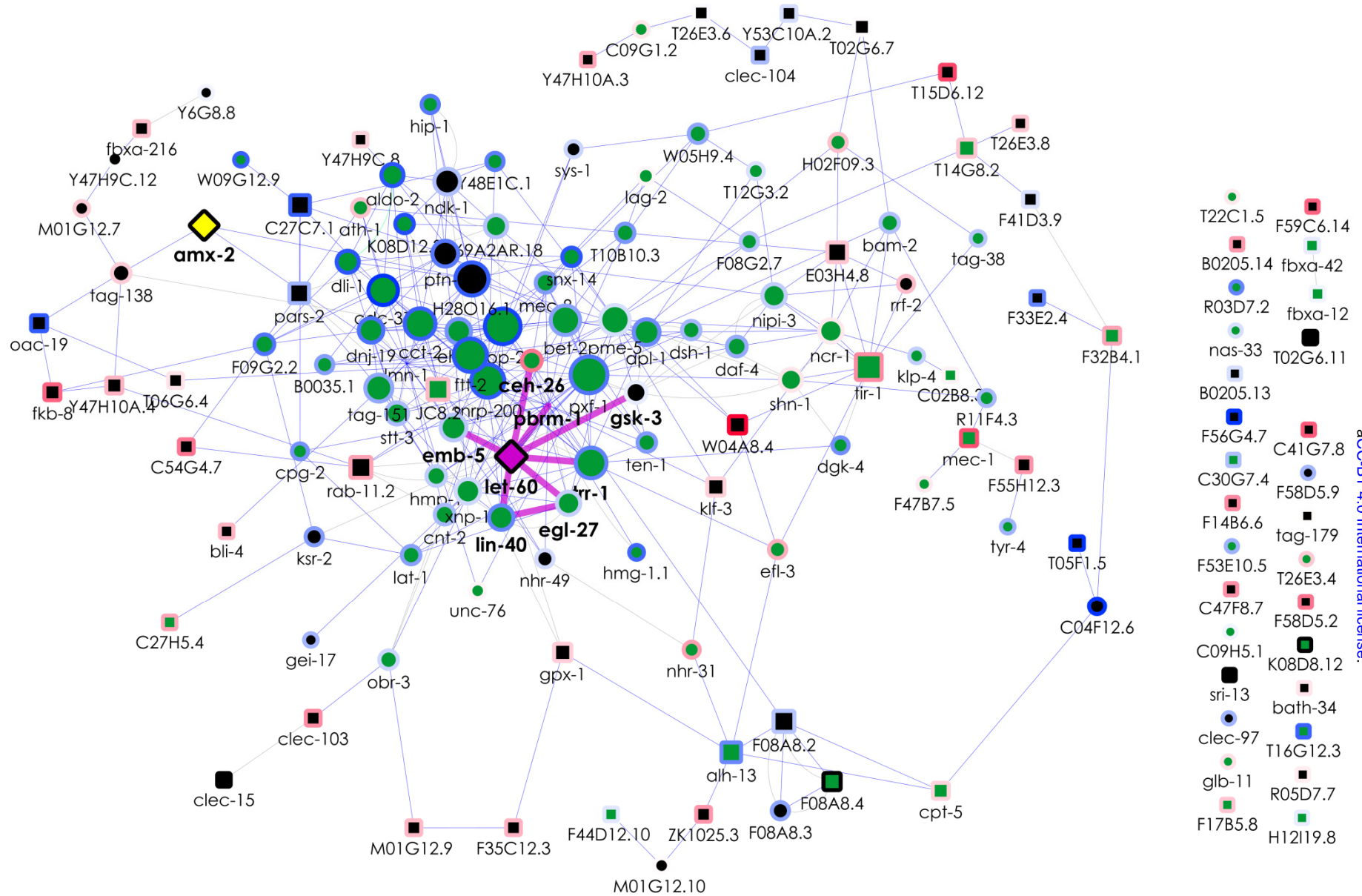
817 **Table 2:** *trans*-bands detected in the sensitized miRILs.

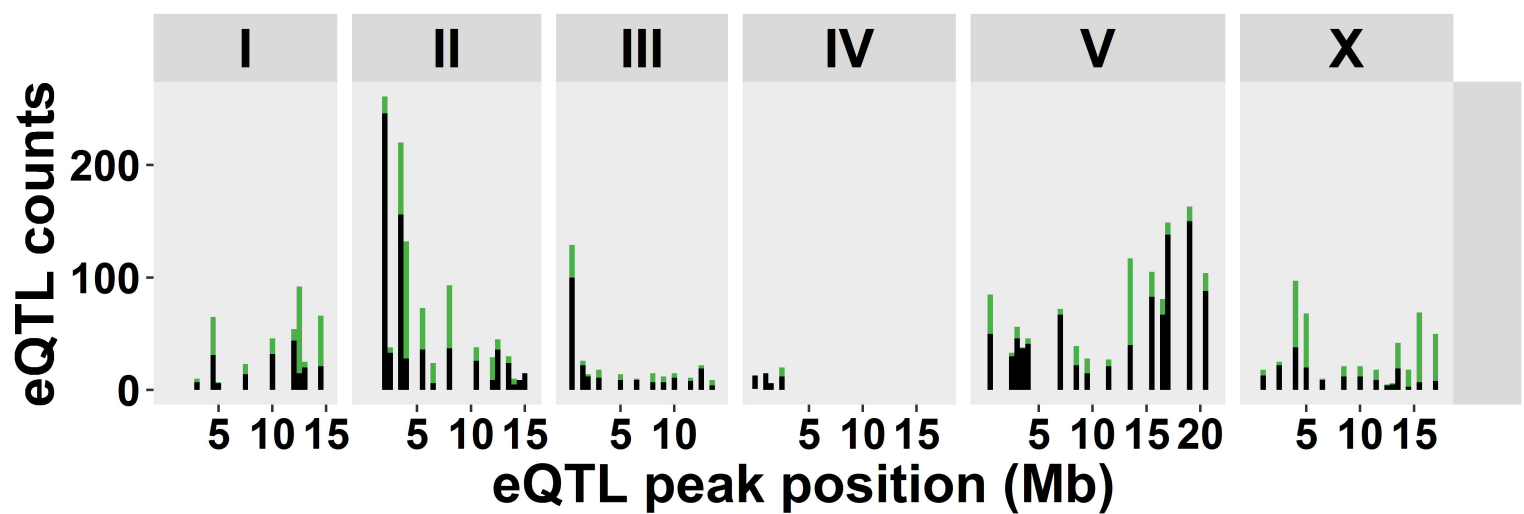
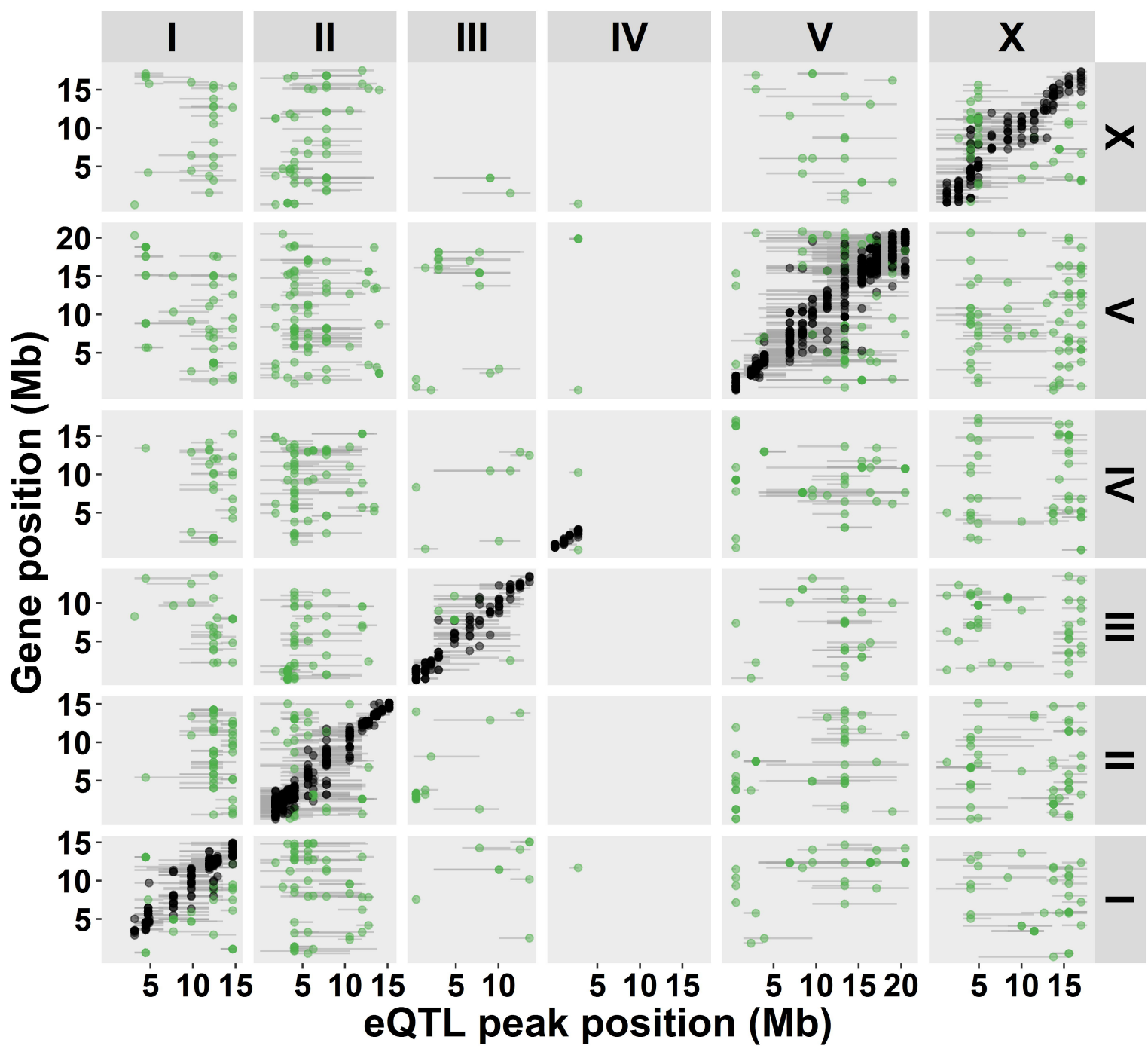
A

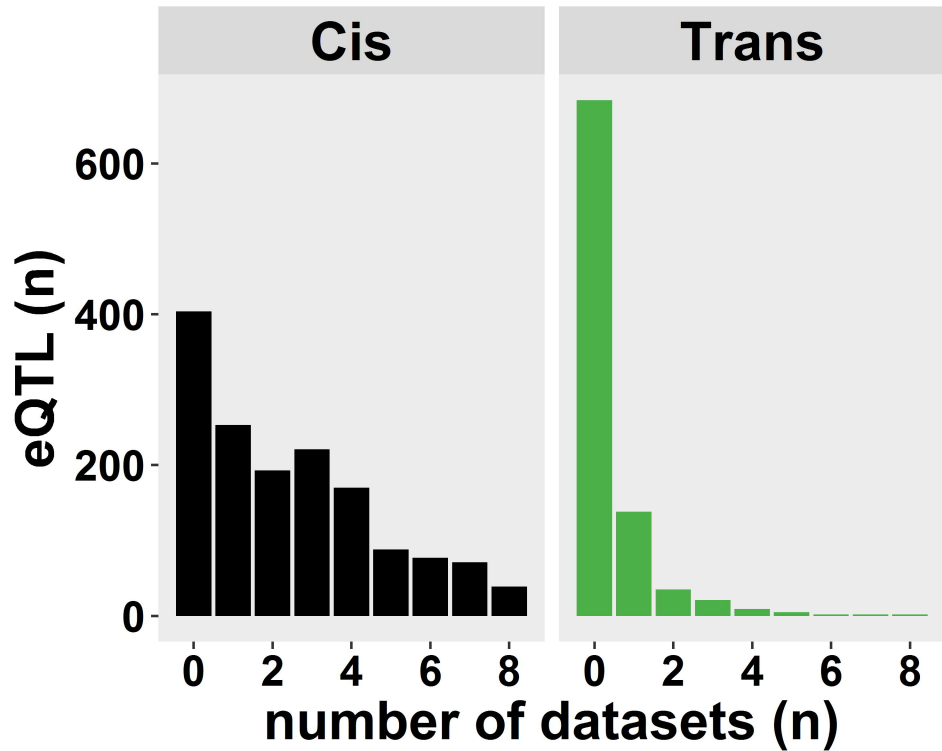


B









<b>Table 1</b>		
<b>The number of genes with an eQTL, in brackets the number of spots.</b>		
	<i>cis</i> -eQTL	<i>trans</i> -eQTL
N2 higher	1074 (1564)	448 (582)
CB4856 higher	456 (632)	464 (576)
Total <sup>1</sup>	1516 (2196)	898 (1149)
<sup>1</sup> : The discrepancy between the sum of N2 higher and CB4856 higher and the total is due to genes being represented by multiple spots, which often represent different splice variants.		



<b>Table 2</b>			
<b><i>Trans</i>-bands detected in the sensitized miRILs.</b>			
<b><i>Trans</i>-band position</b>	<b>Number of genes with a <i>trans</i>-eQTL (spots)</b>	<b>Detected in previous experiments<sup>1</sup></b>	<b>Uniqueness<sup>2</sup></b>
I:12.0-15.0 Mb	107 (127)	9/107 (8.4%)	$p = 1.6 \times 10^{-4}$
II:3.0-6.0 Mb	166 (205)	6/166 (3.6%)	$p = 4.1 \times 10^{-11}$
II:7.0-8.0 Mb	45 (56)	0/45 (0%)	$p = 6.7 \times 10^{-6}$
V:13.0-14.0 Mb	70 (77)	0/70 (0%)	$p = 4.5 \times 10^{-8}$
X:4.0-5.0 Mb	87 (107)	3/87 (3.4%)	$p = 1.7 \times 10^{-6}$
X:15.0-17.0 Mb	85 (104)	1/85 (1.2%)	$p = 3.4 \times 10^{-8}$

<sup>1</sup>Based on Li *et al*, Vinuela *et al*, Rockman *et al.*, and Snoek & Sterken *et al.*; if the same gene had an eQTL on the same chromosome.

<sup>2</sup>Based on a Poisson distribution, where it was expected that 23.8% of the *trans*-eQTL were replicated. P-value indicates the likeliness of an overrepresentation of *trans*-eQTL in the miRIL population.