Ras/MAPK Modifier Loci Revealed by eQTL in Caenorhabditis elegans

Mark G. Sterken,* Linda van Bemmelen van der Plaat,* Joost A. G. Riksen,* Miriam Rodriguez,* Tobias Schmid,†‡ Alex Hajnal,† Jan E. Kammenga,*‡ and Basten L. Snoek*†

*Laboratory of Nematology, Wageningen University and Research, 6708PB, The Netherlands, †Institute of Molecular Life Sciences, and ‡PhD Program in Molecular Life Sciences, University of Zurich, CH-8057, Switzerland

ORCID IDs: 0000-0001-7119-6213 (M.G.S.); 0000-0002-4098-3721 (A.H.); 0000-0003-4822-4436 (J.E.K.); 0000-0001-5321-2996 (B.L.S.)

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

The Ras/MAPK pathway is highly conserved across s metazoans and regulates a wide range of physiological responses, such as cell proliferation, apoptosis, cell differentiation, and tissue morphogenesis (Gokhale and Han 1994). After phosphorylation, MPK-1 enters the nucleus where it regulates many genes by phosphorylation of transcription factors (Tan et al. 1998). Additionally, let-60 activation underlies programmed cell death in C. elegans (Jiang and Wu 2014).

In C. elegans almost all studies on let-60 activation have been conducted with mutant screens using single genetic backgrounds (i.e., a mutation in one genotype, in this case Bristol N2). However, the phenotype of induced mutations can vary widely depending on the genetic background (Duveau and Felix 2012; Chandler et al. 2014; Schmid et al. 2015; Kammenga 2017). Induced mutations in one genetic background do not reveal the allelic effects that segregate in natural populations and contribute to phenotypic variation (Kammenga et al. 2008). At the moment, we lack insight into how genetic background effects modulate activated Ras/MAPK signaling and which genetic modifiers are involved in the underlying genetic architecture of gene expression.
Here we go beyond mutant screens of let-60(gf) in a single genetic background in *C. elegans* by incorporating the effects of multiple genetic backgrounds. This provides the opportunity to evaluate the genetic variation for identifying novel modifiers of the Ras/MAPK pathway. We recently mapped modifiers affecting Ras/MAPK signaling associated with vulval development in a population of recombinant inbred lines derived from a cross between wild-type Hawaiian CB4856 and Bristol N2 (Schmid et al. 2015). The lines were sensitized by introgression of the G13E gain-of-function mutation (n1046) in the Ras gene *let-60* [mutation introgression recombinant inbred lines (miRILs)]. This mutation is analogous to mutations causing excess cell division in human tumors (Kyriakakis et al. 2015). Hawaiian CB4856 males were crossed with Bristol N2 let-60 mutants. Random segregation of the two parental genomes was allowed, except for the *let-60* mutation which was kept homozygous from the F2 generation onwards. After 10 generations of self-fertilization, to drive all regions to homozygosity, independent miRILs were successfully obtained, each carrying the mutation. Quantitative trait loci (QTL) analysis on the vulval index (VI) of these miRILs in combination with allele-swapping experiments revealed the polymorphic monoamine oxidase A (MAOA) gene amx-2 as a negative regulator of Ras/MAPK signaling (Schmid et al. 2015).

Here, we extended the study by Schmid et al. (2015) by studying the transcriptional architecture of the same *let-60(gf)-sensitized* miRILs. First, we measured the transcriptome of 33 miRILs using microarrays. Second, to gain further insight into the underlying molecular mechanisms, pathways, and new modifiers, we mapped gene expression QTL (eQTL). eQTL are polymorphic loci underlying variation in gene transcript abundances and can be used to detect loci which affect many transcript levels and transcriptional pathways (Jansen and Nap 2001). We found that the *let-60(gf)* mutation mainly reveals novel trans-eQTL, which are eQTL distant from the regulated gene. In contrast, cis-eQTL, genes that colocalize with the eQTL, were mostly independent of the *let-60(gf)* mutation. Of the trans-eQTL, 77 genes have an eQTL in a trans-band (eQTL hotspot) colocalizing with *amx-2*. Comparing the transcriptional profiles of transgenic lines with the N2 *amx-2* allele with the CB4856 *amx-2* allele showed the involvement of *amx-2* in 79% (61/77) of the genes mapping to this locus. Through network-assisted gene expression analysis, we found evidence that *amx-2* indirectly affects gene expression mapping to this trans-band.

**MATERIALS AND METHODS**

**General methods and strains used**

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

Strains were maintained on NGM agar seeded with OP50 bacteria at 20°, strains were age synchronized by bleaching and were harvested 72 hr post synchronization. One sample consisted of multiple pools, nonstarved populations from approximately four separate 9-cm NGM plates.

**RNA isolation, cDNA synthesis, and cRNA synthesis**

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

For gene expression analysis, 200 ng of RNA (as quantified by NanoDrop) was used in the “Two-Color Microarray-Based Gene Expression Analysis; Low Input Quick Amp Labeling” protocol, version 6.0, from Agilent (Agilent Technologies, Santa Clara, CA).

**Microarray hybridization and scanning**

The Agilent *C. elegans* (V2) Gene Expression Microarray 4X44K slides were used to measure gene expression. As recommended by the manufacturer, the samples were hybridized for 17 hr and scanned by an Agilent High Resolution C Scanner. The intensities were extracted using Agilent Feature Extraction Software (version 10.7.1.1). The data were normalized using the Limma package in “R” (version 3.3.0 x64). For within-array normalization, the “Loess” method was used and the “Quantile” method for between-array normalization (Zahrak et al. 2007; Smyth and Speed 2003). Unless mentioned otherwise, the log2 transformed intensities were used in subsequent analysis.

**Expanding the genetic map**

Previously, the strains were genotyped with 73 fragment length polymorphism (FLP) markers (as described in Schmid et al. 2015 and Zipperlen et al. 2005) (Table S2).

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

\[
R_{x,i} = \frac{Y_{x,i}}{0.5 \times (Y_{MT2124,i} + Y_{CB4856,i})} - 0.5
\]

where \( Y \) stands for the untransformed intensities of strain \( x \) and spot \( i \) (1, 2, 3, ..., 45220). The obtained values were correlated to the *cis-eQTL* effects from Rockman et al. (2010). Since the current study uses a newer version of the *C. elegans* microarray, only the spots that occurred in both designs were compared. The expression markers were generated following a procedure that is a variation of the method described in West et al. (2006). Per 20 consecutive *cis-eQTL*, the correlation between the N2-eQTL effect and the transformed intensities (R) were calculated. In this way, the correlation would be negative if the genotype stemmed from CB4856, and positive if the genotype stemmed from N2.

In this way, 424 gene expression markers were generated per strain. To control for quality, the markers were filtered for calling the correct genotype in the four MT2124 replicas, the four CB4856 replicas, and in >50% of the samples an absolute correlation >0.5 was required. This pruning resulted in 204 reliable gene expression markers. The quality of these markers was controlled by predicting the genotype by expression in RIL AH2244, in which the gene expression was measured twice. There, we found that 10/204 selected gene expression markers were in disagreement with the genotype. This could be reduced to 0 by comparing only markers with an absolute correlation >0.6. Therefore, all the correlations with an absolute value >0.6 were assigned the predicted marker, which corresponds to an error rate of <0.01 per strain. The genotypes of the remaining markers were manually inferred from the surrounding markers. Furthermore, the genotypes at the ends of the chromosomes were inferred from the distal
most-assigned marker. This brings the size of the resulting map to 289 markers. The correlations and assigned genotypes can be found in Table S3.

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

**eQTL mapping and threshold determination**
Mapping of eQTL mapping was done in "R" (version 3.3.0.x64) and the gene expression data were fitted to the linear model

$$y_{ij} \sim \sum_{j} + e_j,$$

where $y$ is the log$_2$-normalized intensity as measured by microarray of spot $i$ ($i=1, 2, ..., 45220$) of miRIL $j$. This is explained by the genotype (either CB4856 or N2) on marker location $x$ ($x=1, 2, ..., 247$) of miRIL $j$.

A permutation approach was used to determine an empirical false discovery rate (Snoek et al. 2017; Vinuela et al. 2010). First, the log$_2$-normalized intensities were randomly distributed per gene over the genotypes. This randomized data set was tested using the eQTL mapping model. This procedure was repeated 10 times. The threshold was determined using

$$\frac{\text{FDS}}{\text{RDS}} \leq \frac{m_0}{m} q \cdot \log(m),$$

where, at a specific significance level, the false discoveries (FDS) were the averaged outcome of the permutations and read discoveries (RDS) were the outcome of the eQTL mapping. The value of $m_0$, the number of true null hypotheses tested, was 45220 RDS; and for the value of $m$, the number of hypotheses tested, the number of spots (45220) was taken. Because this study only used a limited set of strains, a more lenient threshold of 0.1 was taken as the $q$-value (Benjamini and Yekutieli 2001), which resulted in a threshold of $-\log_{10}(p) > 3.2$.

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

**eQTL analysis**
The distinction between cis- and trans-eQTL was made on the distance between the gene and the eQTL peak. Given the relatively low number of unique recombinations (due to the limited set of strains), the cis-eQTL window was set at 2 Mb. This means that if a gene lies within 2 Mb of the QTL peak or the confidence interval, it is called a cis-eQTL. The confidence interval was based on a $-\log_{10}(p)$ drop of 1.5 compared to the peak.

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

Comparative analysis against previous eQTL experiments was done on remapped experiments downloaded from WormQTL (Snoek et al. 2013; van der Velde et al. 2014) and EleQTL (http://www.bioinformatics.nl/EleQTL). This data set consists of four different experiments representing nine different conditions. The first set contains eQTL in two temperature conditions, 16° and 24°, measured in the L3 stage (Li et al. 2006). The second set contains eQTL over three life stages: L4 juvenile animals grown at 24°, reproducing adult animals (96 hr) grown at 24°, and aging animals (214 hr) grown at 24° (Vinuela et al. 2010). The third set contains eQTL from a single experimental condition (young adults grown at 20°) measured on a large RIL panel (Rockman et al. 2010). The fourth set contains eQTL from three experimental conditions over the course of a heat-shock treatment: a control condition (L4 animals grown for 48 hr at 20°), a heat-shock condition (L4 animals grown for 48 hr at 20° and exposed to 2 hr of 35°), and a recovery condition (similar to heat shock, only followed by 2 hr at 20°) (Snoek et al. 2017). Each of these experiments was compared at FDR = 0.05 to the eQTL mapped at FDR = 0.10 in the let-60(gf)-sensitized miRILs.

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

**Enrichment analysis**
Gene group enrichment analysis was done using a hypergeometric test on the unique genes (not on the spots) with the following criteria: Bonferroni-corrected $P < 0.05$; size of the category, $n > 3$; and size of the overlap, $n > 2$.

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

**amx-2 allelic comparison**
Four independent transgenic lines, two of amx-2[amx-2(Hawaii)]; let-60(gf) and two of amx-2[amx-2(Bristol)]; let-60(gf), were used to investigate the effect of the amx-2 CB4856 and N2 allele on gene expression. These lines contain single-copy insertions on LGII of the N2 or
CB4856 alleles of amx-2 in an N2 amx-2(gf); let-60(gf) background. The effect of different alleles on gene expression was measured by microarray for each independent transgenic line and compared to the effects of the eQTL mapping to the trans-band closest to the position of amx-2.

Connectivity network analysis
To investigate the connectivity and function of the affected genes we used WormNet (version 3) (Cho et al. 2014) and GeneMANIA plug-in for Cytoscape (version 3.4.0) (Montojo et al. 2010; Shannon et al. 2003). WormNet was used to investigate enrichment in connectivity within groups of genes. For example, groups of coexpressed genes mapping to the same trans-band were assumed to share the same regulator. Further evidence for this coexpression and regulation can be found in WormNet if these genes are significantly more connected than by chance. GeneMANIA was used to find the closest neighbors, those genes with the most connections, of lying within 2 Mb of the regulated gene), the green represents trans-eQTL lying elsewhere). The gray horizontal lines indicate the location into account, we only found ~24% (214 out of 898) of the genes had a trans-eQTL in at least one of the previous experiments (Figure 2).
and Figure S3C). To further compare the trans-eQTL overlap, it was investigated whether the trans-eQTL colocalized on the same chromosome. Using that criterion, ~9% (80 out of 898) of the genes with a trans-eQTL were found in a previous study. This showed that the majority of the genes with a trans-eQTL were specifically detected in the sensitized miRIL population. This observation also extended to the trans-bands. To test the specificity of the trans-bands, we counted the number of times a gene within the miRIL trans-bands had a colocalizing trans-eQTL in other studies. All trans-bands were found to be specific for the let-60(gf) miRILs (Poisson distribution, P < 0.001; Table 2). This provides additional evidence that these trans-bands are specific for interaction between the genetic background and the let-60(gf) mutation.

Allogetic differences

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

The genes with cis-eQTL were enriched for genes in the gene classes math, bath, btb, and fbx. These classes contain many genes that function as intra- and extracellular sensors involved in gene–environment interactions, and were found highly polymorphic between N2 and CB4856, but also between other wild isolates (Thompson et al. 2015; Volkers et al. 2013). The genes with trans-eQTL with higher expression to CB4856 loci were enriched for genes related to development in general and gonad development specifically. These enrichments were likely to stem from the trans-bands on chromosome I: 12–15 Mb and II: 3–6 Mb specifically. These two trans-bands mainly consisted of genes with a higher expression linked to the CB4856 locus and were enriched for genes associated with reproduction and/or development.

### Table 2 Trans-bands detected in the sensitized miRILs

| Trans-band Position | Number of Genes with a Trans-eQTL (spots) | Detected in Previous Experiments | Uniqueness |
|---------------------|------------------------------------------|----------------------------------|------------|
| I: 12.0–15.0 Mb     | 107 (127)                                | 9/107 (8.4%)                     | P = 1.6 x 10^-4 |
| II: 3.0–6.0 Mb      | 166 (205)                                | 6/166 (3.6%)                     | P = 4.1 x 10^-11 |
| II: 7.0–8.0 Mb      | 45 (56)                                  | 0/45 (0%)                        | P = 6.7 x 10^-6 |
| V: 13.0–14.0 Mb     | 70 (77)                                  | 0/70 (0%)                        | P = 4.5 x 10^-3 |
| X: 4.0–5.0 Mb       | 87 (107)                                 | 3/87 (3.4%)                      | P = 1.7 x 10^-6 |
| X: 15.0–17.0 Mb     | 85 (104)                                 | 1/85 (1.2%)                      | P = 3.4 x 10^-8 |

*Based on Li et al. (2006), Vinuela et al. (2010), Rockman et al. (2010), and Snoek et al. (2017), if the same gene had an eQTL on the same chromosome.

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

| Figure 2 eQTL detected in previous experiments. The eQTL mapped in the let-60(gf) RIL population were compared to eQTL mapped in three published C. elegans eQTL studies, representing nine different conditions (Li et al. 2006; Rockman et al. 2010; Vinuela et al. 2010; Snoek et al. 2017). The histogram shows how many of the eQTL mapped in this study were found back in these independent conditions. Whereas the majority of cis-eQTL was mapped in previous studies, there was only a minority of trans-eQTL that had been mapped.

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

The trans-band on chromosome I: 12–15 Mb colocalizes with a previously identified QTL for vulval induction, for which amx-2 was identified as polymorphic regulator (Schmid et al. 2015). To determine if the trans-eQTL mapping to the amx-2 locus were affected by the allelic difference between N2 and CB4856, we measured gene expression in transgenic lines containing single-copy insertions of either allele of amx-2 in a amx-2(gf) background. For each allele, two independent strains were created and the transcriptome was measured using microarrays (Schmid et al. 2015). From these measurements, the allelic effects of amx-2 on gene expression were determined and compared with the trans-eQTL effects measured in the miRIL population. These were found to correlate significantly (Figure S4; R^2 ~0.34; P < 2 x 10^-4). Specifically, 61 of the 77 genes with an eQTL colocating with amx-2 showed the same allelic effect in the transgenic strains carrying the two amx-2 variants. Furthermore, analysis of the effect directions showed that the trans-band on chromosome I (Table 2) consists of two separate trans-bands (see Figure S4); the first colocating with amx-2 and the second lying more distal on chromosome I.

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

**DISCUSSION**

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

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

To our knowledge, this is the first study where a mutation, introgressed into a panel of RILs, is used to investigate the genetic architecture of transcript variation by eQTL analysis. Unfortunately, the direct effect of the $let-60(gf)$ allele on the genetic background is impossible to determine in this population, since there is no population with the same genetic background lacking the $let-60(gf)$ mutation. Therefore, we used previously published eQTL studies in $C. elegans$ as an estimation for the miRIL-specific eQTL. We found that introgression of $let-60(gf)$ leads to an overrepresentation of specific trans-eQTL, while the cis-eQTL showed a high overlap with eQTL observed in the absence of $let-60(gf)$ allele. Although our study is based on a relatively small number of strains, the verification of identified eQTL against other eQTL studies in $C. elegans$ shows that at least 73% of the identified cis-eQTL were previously discovered. This supports our power analysis, showing that our study has enough statistical power to detect eQTL with larger effects.

Given that cis-eQTL are most likely genes with polymorphisms in or near the gene itself, it is expected that most have been previously discovered. In support of this, most of the cis-eQTL genes show higher expression in $N2$, making it likely that a relatively large proportion of these cis-eQTL indeed stem from polymorphisms causing mishybridization on the microarrays instead of true expression differences (Rockman et al. 2010; Alberts et al. 2007). Therefore, most of the cis-eQTL will not represent gene expression changes that can be linked to the $let-60(gf)$ mutation in the miRIL population. That is also exactly what we found when constructing the regulatory network by connecting cis- and trans-eQTL to $let-60$; cis-eQTL are only loosely, or not at all, connected to the network.

Compared to cis-eQTL, the trans-eQTL identified in the miRIL population were hardly found in previous eQTL studies in $C. elegans$ (only 24% were previously detected). However, it should be noted that trans-eQTL are themselves also environment and age dependent (Vinuela et al. 2010; Francesconi and Lehner 2014; Li et al. 2006; Snoek et al. 2017), therefore it is possible that not all the trans-eQTL are indeed $let-60$ induced. For example, they might be age related (Vinuela et al. 2010). Another aspect that influences detection of trans-eQTL is the power of the study, as trans-eQTL explain less of the variation than cis-eQTL. However, by being stringent in the overlap (requiring only one match over nine different environments), it is likely that most of these eQTL would be detected. This makes it likely that many of the identified trans-eQTL are specifically linked to an effect of the $let-60(gf)$ mutation in the genetic background. This extends the idea that environmental
perturbations can reveal additional genetic variation (Li et al. 2008) by perturbing the genetic environment (Kammenga et al. 2008; Duveau and Felix 2012; Schmid et al. 2015). The advantage of mutational perturbation and perturbation via induced responses in eQTL studies (for example, see Snoek et al. 2012 and Snoek et al. 2017) is that it contextualizes the transcriptional response. Yet, as the phenotypic influence of let-60(gf) in the miRIL population was investigated before, we know it results in a phenotype. In the study presented here, the context is the let-60(gf) mutation, which results in a phenotype where VI is affected. It is therefore interesting that the trans-bands colocalize with QTL mapped for VI in the same miRIL population: trans-band I: 12–15 Mb overlaps with QTL1b, trans-band II: 7–8 Mb with QTL2, and trans-band V: 13–14 Mb with QTL3 (Schmid et al. 2015). This colocalization adds to the plausibility that the novel trans-eQTL detected in the miRIL population are indeed due to the let-60(gf) mutation and its interaction with the genetic background.

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

Using the same let-60(gf) miRIL population, we have previously identified the polymorphic MAOA amx-2 as a background modifier that negatively regulates the Ras/MAPK pathway and the VI phenotype (Schmid et al. 2015). The trans-band on chromosome I colocalizes with the amx-2 QTL for VI (QTL1b (Schmid et al. 2015)), it is therefore possible that the modifier amx-2 also affects gene expression. More specifically, the trans-band on chromosome I originally identified in the miRILs consists of two separate regulatory loci. The N2 allelic effect at the amx-2 locus is mostly negative; whereas for the more distal part of the original trans-band, the N2 allelic effect is mostly positive. Yet, it seems unlikely that amx-2 is a direct gene expression effector. As amx-2 is a mitochondrial MAOA, a catalyzer of neuropeptide oxidative deamination, it will probably not influence gene expression directly (Tipton et al. 2004). Placing the trans-eQTL of the amx-2 trans-band in a gene interaction network supports this line of reasoning: although let-60 is in the center of the trans-eQTL, amx-2 is only peripheral.

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

As this places \textit{amx-2} in the causal chain of events, we think it is most likely that \textit{amx-2} does not affect gene expression directly but via another gene involved in Ras/MAPK signaling, possibly directly linked to \textit{let-60}. Therefore, we hypothesize that \textit{amx-2} affects Ras/MAPK signaling via one or more of the six previously found genes with a genetic interaction with \textit{let-60}: \textit{lin-40} (also called \textit{egrl}), \textit{egl-27}, \textit{trr-1}, \textit{pbrm-1}, \textit{ceh-26} (also called \textit{pros-1}), and \textit{emb-5}. For example, \textit{egl-27} and \textit{lin-40} both have a \textit{trans-eQTL} mapping to the \textit{amx-2} locus, and are the two MTA (metastasis-associated protein) homologs found in \textit{C. elegans} (Ch’ng and Kenyon 1999; Herman et al. 1999; Solari et al. 1999). The proteins act in the NURD chromatin remodeling complex, which has previously been shown to antagonize Ras-induced vulval development (Solari and Ahringer 2000). Chromatin remodeling provides a more likely mechanism of action than the molecular role of \textit{amx-2} itself.

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

How can our results help form a better understanding of the Ras/MAPK pathway? Our previous investigation on VI in the miRIL population resulted in the identification of three QTL harboring polymorphic Ras-signaling modifiers. Expanding our research to genetic variation affecting gene expression in a \textit{let-60(gf)}-sensitized RIL population uncovered six \textit{trans}-bands (or “eQTL hotspots”). As mentioned before, these \textit{trans}-bands overlap with the QTL mapped for VI in Schmid et al. (2015). There are three more \textit{trans}-bands that do not overlap with QTL for VI, but are also likely to represent modifier loci of the Ras/MAPK pathway, possibly underlying other Ras/MARK-associated phenotypic differences.

The main advantage of studying the genetic architecture of gene expression in the miRIL population is that it creates more insight in the pathways is important for gaining insight into individual-based differences.

**Conclusion**

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

**ACKNOWLEDGMENTS**

The authors thank Harm Nijveen for making our data available in EleQTL. M.R., T.S., and J.E.K. were funded by the European Community’s Health Seventh Framework Program (FP7/2007-2013) under grant 222936. B.L.S. was funded by ERA-sybio-plus ZonMW project GRAPPLE: Iterative modeling of gene regulatory interactions underlying stress, disease and ageing in \textit{C. elegans} (project 90201066); and The Netherlands Organization for Scientific Research (project no. 823.01.001).

Author contributions: A.H., J.E.K., and B.L.S. conceived and designed the experiments. J.A.G.R., M.R., and T.S. conducted the experiments. M.G.S., L.v.B.v.d.P., and B.L.S. conducted transcriptome and main analyses. M.G.S., A.H., J.E.K., and B.L.S. wrote the manuscript.

**LITERATURE CITED**

Alberts, R., P. Terpstra, Y. Li, R. Breitling, J. P. Nap et al., 2007 Sequence polymorphisms cause many false cis eQTLs. PLoS One 2: 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: 503–509.

Benjami, Y., and D. Yekutieli, 2001 The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 29: 1165–1188.

Byrne, A. B., M. T. Weirach, V. Wong, M. Koeva, S. J. Dixon et al., 2007 A global analysis of genetic interactions in Caenorhabditis elegans. J. Biol. 6: 8.

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: 1321–1336.

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: 3303–3312.

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: W76–W82.

Cowley, S., H. Paterson, P. Kemp, and C. J. Marshall, 1994 Activation of map kinase kinase is necessary and sufficient for Pcl2 differentiation and for transformation of NIH 3T3 cells. Cell 77: 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: 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: 208–211.

Gerstein, M. B., Z. J. Lu, C. Cheng, B. I. Arshinoff et al., 2010 Integrative analysis of the Caenorhabditis elegans genome by the modEncode project. Science 330: 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: 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: 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: D789–D793.

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: 1055–1064.

Jansen, R. C., and J. P. Nap, 2001 Genetical genomics: the added value from segregation. Trends Genet. 17: 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: e1004513.

Kammen, J. E., 2017 The background puzzle: how identical mutations in the same gene lead to different disease symptoms. FEBS J. DOI:10.1111/febs.14080.

Kammena, 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 Genet. 24: 178–185.

Kyriakakis, E., M. Markaki, and N. Tavernarakis, 2015 Caenorhabditis elegans as a model for cancer research. Mol. Cell. Oncol. 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: 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. Nat. Genet. 38: 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: e222.

Li, Y., R. Breitling, and R. C. Jansen, 2008 Generalizing genetical genomics: getting added value from environmental perturbation. Trends Genet. 24: 518–524.

Mansour, S. J., W. T. Matten, A. S. Hermann, J. M. Candia, S. Rong et al., 1999 Transformation of mammalian cells by constitutively active MAP kinase kinase. Science 265: 966–970.

Mentojo, J., K. Zuberi, H. Rodriguez, F. Kazi, G. Wright et al., 2010 GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. Bioinformatics: 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: 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: 29–34.

Paaby, A. B., and M. V. Rockman, 2014 Cryptic genetic variation: evolution's hidden substrate. Nat. Rev. Genet. 15: 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: e09178.

Prior, I. A., and J. F. Hancock, 2012 Ras trafficking, localization and compartmentalized signalling. Semin. Cell Dev. Biol. 23: 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: 372–376.