Common loci underlie natural variation in diverse toxin responses

Shannon C. Brady*,†,§§, Kathryn S. Evans*,†,§§, Joshua S. Bloom‡,§,**, Robyn E. Tanny*, Daniel E. Cook*,†, Sarah E. Giuliani*, Stephen W. Hippleheuser*, Mostafa Zamanian††, Erik C. Andersen*,†,‡‡,1

*Molecular Biosciences, Northwestern University, Evanston, IL 60208
†Interdisciplinary Biological Sciences Program, Northwestern University, Evanston, IL 60208
‡Department of Human Genetics, University of California, Los Angeles, Los Angeles, CA 90095
§Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, CA 90095
**Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, CA 90095
††Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin, Madison, Madison, WI
‡‡Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago, IL 60611
§§ The first two authors contributed equally
1Corresponding author

Corresponding author:
Erik C. Andersen
Department of Molecular Biosciences
Northwestern University
3125 Cook Hall
2205 Tech Drive
Evanston, IL 60208
847-467-4382
erik.andersen@northwestern.edu

Shannon: 0000-0002-3043-1544
Katie: 0000-0002-1388-8155
Josh: 0000-0002-7241-1648
Robyn: 0000-0002-0611-3909
Dan: 0000-0003-3347-562X
Mostafa: 0000-0001-9233-1760
Erik: 0000-0003-0229-9651

Running title: (35 characters): Dissecting toxin-response QTL

Keywords: (5)
C. elegans, QTL, genetic interactions, toxin, pleiotropy

Article summary (100 words):
A panel of recombinant *Caenorhabditis elegans* lines was exposed to 16 toxins, and responses were quantified using a high-throughput fitness assay. We identified 114 distinct quantitative trait loci (QTL), including three hotspots where QTL were enriched across the conditions tested. These hotspots could represent common loci underlying toxin responses. Additionally, we found that both additive and epistatic genetic factors control toxin responses. We go on to validate additive QTL and inter- and intra-chromosomal interaction QTL based on toxin responses assays of near-isogenic lines (NILs) and chromosome-substitution strains (CSSs), going further to delineate genetic causes than most experiments of this scale.

ABSTRACT

Phenotypic complexity results from the contributions of environmental factors and multiple genetic loci, interacting or acting independently. Studies of yeast and *Arabidopsis* found that the majority of natural variation across phenotypes is attributable to independent additive quantitative trait loci. Detected loci in these organisms explain most of the estimated heritable variation. By contrast, many heritable components underlying phenotypic variation in metazoan models remain undetected. Before the relative impacts of additive and interactive variance components on metazoan phenotypic variation can be dissected, high replication and precise phenotypic measurements are required to obtain sufficient statistical power to detect loci contributing to this missing heritability. Here, we used a panel of 296 recombinant inbred advanced intercross lines of *Caenorhabditis elegans* and a high-throughput fitness assay to detect loci underlying responses to 16 different toxins, including heavy metals, chemotherapeutic drugs, pesticides, and neuropharmaceuticals. Using linkage mapping, we identified 114 distinct genomic regions that underlie variation in responses to these toxins and predicted the relative contributions of additive loci and genetic interactions across various growth parameters. Additionally, we identified three genomic regions that impact responses to multiple classes of toxins. These quantitative trait loci hotspots could represent common factors impacting toxin responses. We went further to both generate near-isogenic lines and chromosome-substitution strains and then experimentally validate these QTL hotspots, implicating additive and interactive loci that underlie toxin-response variation. The discovery of these QTL hotspots indicate that pleiotropic loci that control responses to multiple conditions could underlie the means by which large regions of the genome were swept across the *C. elegans* species.

INTRODUCTION

Rapid advances in genome-sequencing technologies enabled the collection of high-quality genomic datasets for many species (Mardis 2017). These data, paired with a broad range of high-throughput phenotypic assays, made quantitative genetics a powerful tool in biology. Linkage mapping has been used to identify quantitative trait loci (QTL), leading to profound impacts on human health (Easton et al. 1993; Cowley 2006; Altshuler et al. 2008), agriculture and livestock (Rothschild et al. 2007; Johnsson et al. 2015; Leal-Bertioli et al. 2015; Shang et al. 2016), and basic biology (Mackay 2001; Peng et al. 2016; Andersen et al. 2015a). Despite the growing number of detected QTL across numerous traits, these QTL often do not explain the complete heritable component of trait variation (Rockman 2012). This missing heritability can be attributed to undetected small-effect additive loci and/or interactions between QTL (Bloom et al. 2015). Although some studies contend that epistatic effects among QTL might explain missing heritability (Nelson et al. 2013; Mackay 2015; Lachowiec et al. 2015; Malmberg et al. 2005; Zuk et al. 2012), others argue that missing heritability comprises small-effect additive loci that remain undetected in cases where statistical power is too low (Hill et al. 2008; Mäki-Tanila and Hill 2014; Ehrenreich 2017; Yang et al. 2010). Quantitative geneticists have leveraged large numbers of recombinant strains in both yeast and *Arabidopsis* to overcome power limitations and concluded that, when power is sufficient, small-effect additive components can be identified that account for nearly all of the heritability of a given trait (Bloom et al. 2015, 2013; Simon et al. 2008). We require a metazoan system with high statistical power to determine whether this predominantly additive-QTL model remains broadly applicable in animals.

One such tractable metazoan is *Caenorhabditis elegans.* The genetic variation among a panel of recombinant inbred advanced intercross lines (RIAILs) generated between the N2 and CB4856 strains of *C. elegans (Rockman and Kruglyak 2009; Andersen et al. 2015a)* has been leveraged in several linkage mapping analyses (McGrath et al. 2009; Bendesky et al. 2011; Zdraljevic et al. 2017a; Lee et al. 2017; Bendesky et al. 2012; Singh et al. 2016; Schmid et al. 2015; Balla et al. 2015). Additionally, a high-throughput phenotyping platform to rapidly and accurately measure animal fitness could provide the replication and precision required to detect small-effect additive loci and to determine the relative contributions of additive and/or epistatic loci to trait variation (Zamanian et al. 2018a; Zdraljevic et al. 2017b; Andersen et al. 2015b). Notably, the combination of this panel and phenotyping platform have facilitated linkage mappings of multiple distinct fitness parameters, resulting in the detection of a single QTL, in fact a single quantitative trait gene (QTG), that underlies several fitness-related traits (Andersen et al. 2014). This example of pleiotropy suggests that large-scale studies could reveal additional pleiotropic effects.

Such large-scale studies have implicated pleiotropic QTL that impact the expression of a broad range of genes (Keurentjes et al. 2007; Breitling et al. 2008; Rockman et al. 2010; Hasin-Brumshtein et al. 2016). Variation in the master regulators that are within these expression QTL hotspots have downstream effects on the transcription of many gene targets. Similarly, other QTL hotspots could impact multiple traits, such as responses to various conditions. In yeast, most chemical-response QTL are thought to be unique to one or a few conditions, whereas very few QTL have been found to have pleiotropic effects across many conditions (Ehrenreich et al. 2012; Singh et al. 2017; Knoch et al. 2017). Although QTL underlying responses to individual conditions have been identified across multiple animal models (Najarro et al. 2015; Marriage et al. 2014; Highfill et al. 2017; Bubier et al. 2014; Crusio et al. 2016), the existence of QTL hotspots that influence multiple condition responses has yet to be determined in metazoans.

Here, we performed a set of linkage-mapping experiments with a large panel of recombinant lines to identify QTL implicated in responses to 16 different toxins and found three QTL hotspots that underlie many of these responses. We demonstrated how high replication in the high-throughput fitness assay can enable the identification and validation of QTL, even in cases of small phenotypic effects. Additionally, we analyzed relative contributions of additive and epistatic genetic loci in various

toxin responses. Finally, we discovered evidence for interactions between loci of the N2 and CB4856 strains that impact several toxin responses and might suggest how large regions of the genome were swept toward fixation across the species.

MATERIALS AND METHODS

Strains: Animals were grown at 20ºC using OP50 bacteria spotted on modified nematode growth medium (NGMA), containing 1% agar and 0.7% agarose to prevent animals from burrowing. For each assay, strains were propagated for five generations after starvation to reduce transgenerational effects of starvation (Andersen et al. 2014)**.** Recombinant inbred advanced intercross lines (RIAILs) used for linkage mapping were constructed previously *(Andersen et al. 2015a)***.** The construction of near-isogenic lines (NILs) and chromosome substitution strains (CSSs) is detailed below, and all strains are listed in the **Supplementary Information**. Strains are available upon request.

High-throughput toxin response assay: We used a modified version of the high-throughput fitness assay described previously (Zdraljevic et al. 2017a). Populations of each strain were passaged for four generations, amplified, and bleach-synchronized. Approximately 25 embryos from each strain were then aliquoted to 96-well microtiter plates at a final volume of 50 µL of K medium (Boyd et al. 2012). Embryos hatched overnight and arrested in the L1 larval stage. The following day, arrested L1 animals were fed HB101 bacterial lysate (Pennsylvania State University Shared Fermentation Facility, State College, PA; (García-González et al. 2017)) at a final concentration of 5 mg/mL in K medium and were grown to the L4 larval stage for 48 hours at 20ºC with constant shaking. Three L4 larvae were then sorted using a large-particle flow cytometer (COPAS BIOSORT, Union Biometrica, Holliston, MA) into microtiter plates that contained HB101 lysate at 10 mg/mL, K medium, 50 µM kanamycin, and either diluent (1% DMSO or 1% water) or diluent and a toxin of interest. The sorted animals were then grown for 96 hours at 20ºC with constant shaking. During this time, the sorted animals matured to adulthood and laid embryos, yielding a population of parent and progeny in each microtiter well. Prior to the measurement of fitness parameters from the populations, animals were treated with sodium azide (50 mM in M9) to straighten their bodies for more accurate growth-response parameter measurements. Traits that were measured by the BIOSORT include brood size (n), animal length (time of flight, TOF), and optical density (extinction time, EXT).

Toxin-response trait calculations: Phenotypic measurements collected by the BIOSORT were processed using the R package *easysorter*, which was specifically developed for processing this type of data set (Shimko and Andersen 2014). Briefly, the function *read_data* imported raw phenotypic data then identified and eliminated bubbles. Next, the *remove_contamination* function discarded wells that contained bacterial or fungal contamination (determined by visual inspection) prior to analyzing population parameters. The *sumplate* function then calculated normalized measurements and summary statistics of the assayed traits for the population of animals in each well. The number of animals in each well was divided by the number of animals sorted into that well, yielding a normalized brood size (norm.n). Additionally, optical density (EXT) of each animal was divided by animal length (TOF), resulting in a normalized optical density (norm.EXT) for each animal in each well. The summary statistics calculated for each population include 10th, 25th, 50th, 75th, and 90th quantiles, mean, and median measurements of TOF, EXT, and norm.EXT as well as variance for TOF and EXT. In total, this analysis resulted in 24 phenotypic measurements for each condition tested. When strains were measured across multiple assay days, the *regress(assay=TRUE)* function was used to fit a linear model with the formula (*phenotype ~ assay*) to account for differences among assays. Next, outliers were removed by eliminating phenotypic values that were greater than two times the IQR plus the 75th quantile or less than two times the IQR minus the 25th quantile (unless at least 5% of the strains lied outside this range in the case of RIAIL assays). Finally, toxin-specific effects were

calculated using the *regress(assay=FALSE)* function from *easysorter*, which fits a linear model with the formula (*phenotype ~ control phenotype*) to generate residual phenotypic values that account for differences between populations that were present in control conditions. For this reason, strain phenotypes in control conditions can influence regressed toxin effects and trait categorizations (below).

Dose-response assays: For each toxin, a dose-response experiment was performed using quadruplicates of four genetically diverged strains (N2, CB4856, DL238, and JU258). Animals were assayed using the high-throughput fitness assay, and toxin-response trait calculations were performed as described above (**File S1**). The concentration of each toxin that provided a highly reproducible toxin-specific effect with variation between N2 and CB4856 was selected for linkage-mapping experiments. The chosen concentrations and diluents of each toxin are as follows: cadmium 100 µM in water, carmustine 250 µM in DMSO, chlorothalonil 250 µM in DMSO, chlorpyrifos 1 µM in DMSO, cisplatin 250 µM in water, copper 250 µM in water, diquat 250 µM in water, fluoxetine 250 µM in DMSO, FUdR 50 µM in water, irinotecan 125 µM in DMSO, mechlorethamine 200 µM in DMSO, paraquat 500 µM in water (50 µM was used for the CSS and NIL assays), silver 150 µM in water, topotecan 400 µM in water, tunicamycin 10 µM in DMSO, and vincristine 80 µM in water. Toxins assayed in this manuscript were purchased from Fluka (chlorothalonil, #36791-250MG; chlorpyrifos, #45395-250MG; diquat dibromide monohydrate, #45422-250MG-R), Sigma (vincristine sulfate salt, #V8879-25MG; cisplatin, #479306-1G; silver nitrate, #209139; carmustine, #1096724-75MG; topotecan hydrochloride, #1672257-350MG), Calbiochem (tunicamycin, #654380), Aldrich (mechlorethamine hydrochloride, #122564-5G, cadmium chloride #01906BX), Alfa Aesar (irinotecan hydrochloride trihydrate, #AAJ62370-MD), Bioworld (5-fluoro-2'-deoxyuridine, #50256011), Enzo Life Sciences (fluoxetine, #89160-860), Mallinckrodt (cupric sulfate, #4844KBCK), and Chem Service (paraquat, #ps-366).

Linkage mapping: A total of 296 RIAILs were assayed in the high-throughput fitness assay described previously in the presence of each toxin listed above as well as control conditions (water or DMSO, **File S2**). Linkage mapping was performed using the R package *linkagemapping* (www.github.com/AndersenLab/linkagemapping, **File S3**). In short, the genotypic data and residual phenotypic data were merged using the *merge_pheno* function and were scaled to have a mean of zero and a variance of one. Quantitative trait loci (QTL) were detected using the *fsearch* function, which calculated logarithm of odds (LOD) scores for each marker and each trait as $-n(ln(1-R^2)/2ln(10))$, where r is the Pearson correlation coefficient between RIAIL genotypes at the marker and trait values (Bloom et al. 2013). The phenotypic values of each RIAIL were then permuted randomly while maintaining correlation structure among phenotypes 1000 times to calculate a significance threshold based on a genome-wide error rate of 5%. The marker with the highest LOD score was then set as a cofactor and mapping repeated iteratively until no significant QTL were detected. Finally, the *annotate_lods* function was used to calculate the fraction of variation in RIAIL phenotypes explained by each QTL. 95% confidence intervals were defined by markers within a 1.5-LOD drop from the marker with the maximum LOD score. We additionally performed a two-dimensional genome scan using the function *scantwo()* in the *qtl* package (Broman et al. 2003) for all significantly mapped traits (**File S4**). Significant interactions were determined by permuting the phenotype data for two randomly chosen traits 1000 times and assigning each a genome-wide error rate of 5%. An interaction with a LOD score above 3.83 (the average lod.int threshold for the two permuted traits) is considered significant. Significant QTL within five cM of each other were considered linked rather than interacting.

Heritability estimates: Broad-sense heritability was estimated from dose-response data using the *lmer()* function in the *lme4* package (Bates et al. 2014) with the following model (*phenotype ~1 + (1|*

strain)) for each trait. Broad-sense heritability estimates could only be calculated for traits with replicates of more than two strains. A variance component model using the R package *regress* was used to estimate the fraction of phenotypic variation explained by additive genetic factors ('narrow-sense' heritability) (David Clifford And 2014, 2006; Bloom et al. 2015). The additive relatedness matrix was calculated as the correlation of marker genotypes between each pair of strains. In addition, a two-component variance component model was calculated with both an additive and pairwise-interaction effect (**File S5**). The pairwise-interaction relatedness matrix was calculated as the Hadamard product of the additive relatedness matrix.

Identification of distinct QTL: All QTL in which the confidence interval did not overlap the confidence interval of another QTL were considered unique. We looked at all pairs of overlapping QTL confidence intervals (within each toxin) to establish a threshold for distinct QTL among these overlapping regions. Because large-effect QTL have small confidence intervals, we assumed that QTL that explain a high proportion of phenotypic variance and have overlapping confidence intervals are less likely to be distinct from each other than are two low-effect overlapping QTL with large confidence intervals. Therefore, we plotted the lowest variance explained estimate within all pairs of overlapping QTL and set a cutoff at the 95th percentile of this distribution (**Figure S3A**). Pairs of overlapping QTL with a variance explained estimate above this cutoff were then plotted according to the correlation coefficient (calculated from the dose response data) between traits for which the overlapping QTL regions mapped (**Figure S3B**). From this distribution, we set a final threshold at the 5th percentile (correlation coefficient > 0.642), above which two overlapping QTL within the same toxin were not considered distinct (**File S6**). The trait with the most significant QTL (highest LOD score) was selected as the representative trait for each group of highly correlated, overlapping QTL (**File S7**).

Calculation of hotspots: We estimated cM distances from recombination events in the RIAIL panel to account for uneven amounts of genetic diversity across the genome. Each chromosome was divided into bins of 20 cM, resulting in a total of 82 bins across the genome. To determine if the 114 distinct QTL significantly clustered around particular genomic regions, we set a threshold for significant QTL hotspots based on the Bonferroni-corrected 95th percentile of a Poisson distribution with a mean of 1.39 QTL (total QTL/total bins).

Generation of near-isogenic lines (NIL): NILs were generated by crossing selected RIAILs to each parental genotype. For each NIL, eight crosses were performed followed by six generations of propagating isogenic lines to ensure homozygosity of the genome. For each cross, PCR amplicons for insertion-deletion (indel) variants on the left and right of the introgressed region were used to confirm progeny genotypes and select non-recombinants within the introgressed region. NILs were whole-genome sequenced as described below to confirm their genotype (**File S8**). Reagents used to generate NILs are detailed in the **Supplementary Information**. A statistical power simulation was used to determine the minimal number of technical replicates required to observe the predicted phenotypic effect of each QTL at 80% power. These calculations are listed in the **Supplementary Information**. The number of technical replicates tested per assay for any given toxin did not exceed 100 because of timing constraints.

Whole-genome sequence library prep and analysis: DNA was isolated from 100-300 µL of packed worms using Qiagen's Blood and Tissue kit (catalog # 69506). Following the ATL lysis step, 4 µl of 100 mg/mL RNAse was added to each sample and allowed to incubate for two minutes at room temperature. DNA concentration was determined using the Qubit dsDNA BR Assay Kit (catalog # Q32850). For each strain, a total of 0.75 ng of DNA was combined with 2.5 µL transposome (Illumina; kit # FC-121-1011) diluted 35x with 1x Tris Buffer (10x Tris Buffer: 100 mM Tris-HCl pH 8.0, 50 mM

$MgCl_2$) in a 10 µL final volume on ice. This reaction was incubated at 55°C for 10 minutes. The amplification reaction for each strain contained (final concentrations): 1x ExTaq Buffer, 0.2 mM dNTPs, 1 U ExTaq (Takara, catalog # RR001A), 0.2 µM primer 1, 0.2 µM primer 2, and 5 µL of tagmentation material from the previous step in a 25 µL total volume. Each strain had a unique pair of indexed primers. We first made a master mix containing buffer, water, dNTPs, and ExTaq then aliquoted the appropriate volume of this mix into each well. We added the specific primer sets to each well and finally the tagmentation reaction. The amplification reaction was incubated in a thermocycler with the following conditions: 72°C for 3 minutes (1x); 95°C for 30 seconds (1x); 95°C for 10 seconds, 62°C for 30 seconds, 72°C for 3 minutes (20x); 10°C on hold. We combined 8 µL from each amplification reaction to generate a pool of libraries. A portion of the libraries was electrophoresed on a 2% agarose gel. DNA was excised and gel purified using Qiagen's Gel Purification Kit (catalog # 28706). The libraries were sequenced on the Illumina HiSeq 2500 platform using a paired-end 100 bp reaction lane. Alignment, variant calling, and filtering were performed as described previously (Cook et al. 2016b). NIL and CSS genotypes were called using the VCF file and a Hidden Markov Model as described previously (Cook and Andersen 2017).

Generation of chromosome substitution strains (CSS): CSSs were generated by crossing N2 and CB4856 parental strains and mating cross progeny to each parental genotype. For each CSS, eight crosses were performed followed by six generations of propagating isogenic lines to ensure homozygosity of the genome. For each cross, PCR amplicons for indels on the left and right of the introgressed region were used to confirm progeny genotypes and select non-recombinants within the introgressed region. CSSs were whole-genome sequenced as described above to confirm their genotype (**File S8**). Reagents used to generate CSSs are detailed in the **Supplementary Information**. As described for NIL assays, power calculations were performed to determine the number of technical replicates required to observe the predicted phenotypic effect of the CSSs.

Categorization of CSS and NIL results: Toxin responses for NILs and CSSs were tested using the high-throughput fitness assay as described above (**File S9**). Complete pairwise statistical analyses of strains was performed for each trait tested in all CSS and NIL assays (Tukey honest significant difference (HSD) test, **File S10**). A *p-value* of $p < 0.05$ was used as a threshold for statistical significance. NIL recapitulation was defined by the significance and direction of effect of the NIL compared to the parental strains. Six categories were defined: 1) no parental difference, 2) recapitulation, 3) no QTL effect, 4) bidirectional interaction, 5) unidirectional interaction, and 6) miscellaneous (**Figure S5**). Traits for which N2 and CB4856 phenotypes were not statistically different comprise the ‘no parental difference’ category and were not further categorized. Traits in the ‘recapitulation’ category must satisfy the following criteria: significant difference between the parental strain phenotypes, significant difference between phenotypes of each NIL and the parent that shares its background genotype, and both NILs must display the expected direction of effect of the introgressed genotype. Traits with ‘no QTL effect’ displayed a significant parental phenotypic difference and the phenotype of each NIL was not statistically different from the phenotype of the parent sharing its background genotype. Traits that have a ‘bidirectional interaction’ must display a significant parental phenotypic difference, the phenotypes of both NILs must be significantly different from phenotypes of both parents, and the phenotypes of both NILs must be transgressive (lie beyond the phenotypic range of the parental strains). Lastly, traits with a ‘unidirectional interaction’ were categorized similarly to the bidirectional interaction, except only one NIL must display a transgressive phenotype and the other NIL either shows no QTL effect or recapitulation. Traits that did not fit these descriptions were categorized as ‘miscellaneous’.

Traits in the chromosome V hotspot were further categorized using the combined data from both the CSS and NIL assays. Seven categories were defined: 1) no parental difference, 2) recapitulation, 3) no QTL effect, 4) external inter-chromosomal interaction (uni- or bidirectional), 5)

internal inter-chromosomal interaction (uni- or bidirectional), 6) intra-chromosomal interaction (uni- or bidirectional), and 7) miscellaneous (**Figure S5**). 'No parental difference' was defined by traits in which the parental strains were either not significantly different from each other or did not have the same direction of effect in both the CSS and NIL assays. 'Recapitulation' and 'no QTL effect' traits were defined by traits that were classified as either recapitulating or no QTL effect, respectively, in both assays. Traits displaying an 'external inter-chromosomal interaction' show evidence for interaction in the CSS but no interaction (either recapitulating or no QTL effect) in the NIL. On the other hand, traits displaying an 'internal inter-chromosomal interaction' showed evidence of the same interaction for both the CSS and the NIL assays. Finally, traits displaying an 'intra-chromosomal interaction' showed evidence of an interaction in the NIL but not in the CSS assay. All other traits that did not fit these descriptions were categorized as 'miscellaneous' (**File S11**).

Statistical analysis: All statistical tests of phenotypic differences in the NIL and CSS assays were performed in R (version 3.3.1) using the *TukeyHSD* function (R Core Team 2017) on an ANOVA model with the formula (*phenotype ~ strain*). The *p-values* of individual pairwise strain comparisons were reported, and a *p-value* of $p < 0.05$ was deemed significant. The direction of effect of each NIL was determined by comparing the median phenotypic value of the NIL replicates to that of each parental strain. NILs whose phenotypes were significantly different from both parents and whose median lied outside of the range of the parental phenotype medians were considered hypersensitive or hyper-resistant. Comparing LOD scores and variance explained between traits with no parental effect and traits with a significant parental effect in the NIL assays was performed using a Wilcoxon rank sum test with continuity correction using the *wilcox.test()* function in R (R Core Team 2017).

Data availability: **File S1** contains results of the dose response assays for all toxins. **File S2** contains the residual phenotypic values for each RIAIL for each trait. **File S3** contains the annotated QTL and confidence intervals identified through linkage mapping. **File S4** contains the result of a two-factor genome scan for all traits with a significant QTL identified with linkage mapping. **File S5** contains the broad-sense heritability estimates as well as additive and interactive components of heritability for each trait. **File S6** includes QTL for all mapped traits, the representative trait for overlapping QTL within each toxin, and the correlation coefficient between the representative trait and each correlated trait calculated from the dose response data. **File S7** is a subset of **File S3** for only the 114 distinct QTL identified. **File S8** is a VCF file for all NILs and CSSs mentioned in this manuscript. **File S9** contains the residual phenotypic data for all strains, including parents, tested in the NIL and CSS assays. **File S10** contains the statistical significance for all pairwise combinations of strains tested for each trait. **File S11** contains the assay categorization for all traits tested with the NIL and CSS strains. The datasets and code for generating figures can be found at http://github/AndersenLab/QTLhotspot.

RESULTS

Identification of QTL underlying variation in responses to 16 diverse toxins

Using a high-throughput fitness assay (Materials and Methods), we tested variation in 24 fitness-related traits of four divergent strains in responses to different concentrations of 16 toxins comprising chemotherapeutics, heavy metals, pesticides, and neuropharmaceuticals (**Figure S1A, File S1**). The concentration of each toxin that minimized within-strain variation and maximized variation between N2 (the laboratory strain) and CB4856 (a wild isolate from Hawaii) was selected. The average broad-sense heritability of fecundity and body-size traits was 48% and 40%, respectively, indicating that a large proportion of the phenotypic variation observed within the population could be attributed to genetic factors (**File S5**). For the selected concentration of each toxin, we assayed 24 growth-response traits for a panel of 296 recombinant inbred advanced

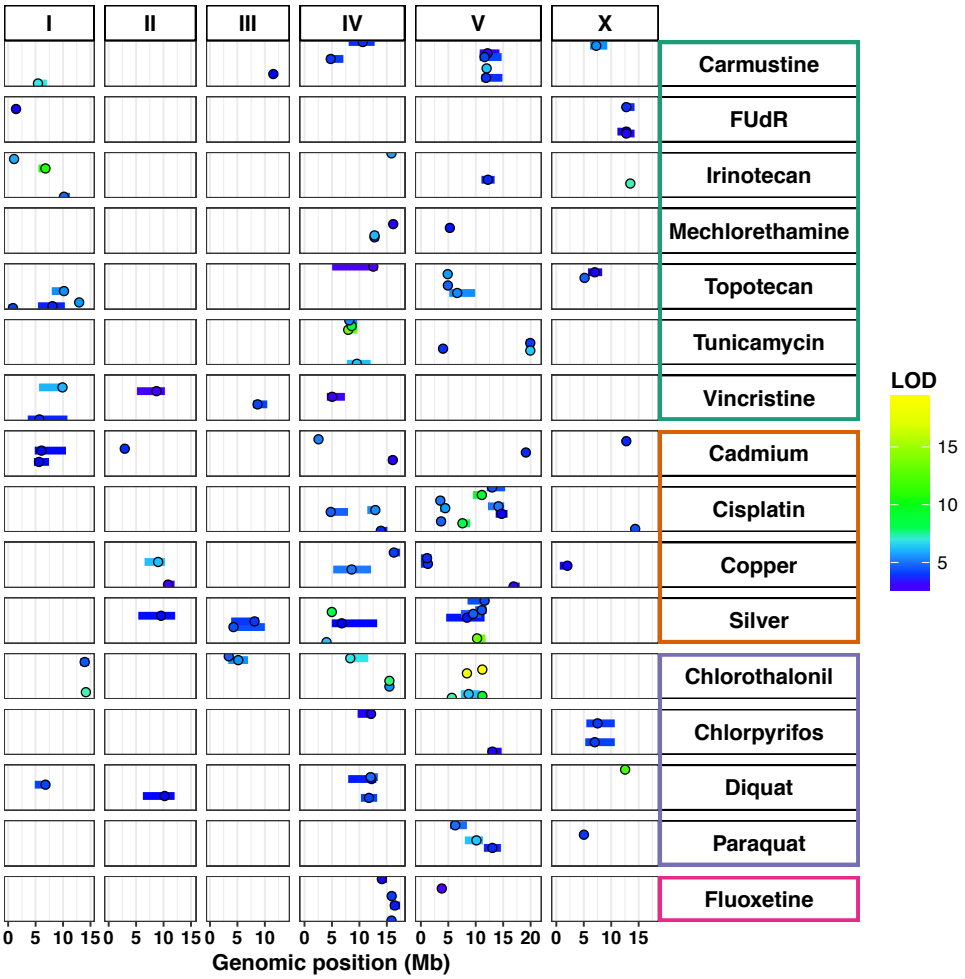

Figure 1 Diverse genetic architectures are implicated in responses to 16 toxins. Linkage mapping results for traits that represent 114 distinct QTL across 16 toxins, comprising chemotherapeutics (teal), heavy metals (orange), pesticides (purple), and neuropharmaceuticals (pink) are plotted. Genomic position (Mb) is shown along the x-axis, split by chromosome, and each trait that represents a distinct QTL is plotted along the y-axis. Each QTL is plotted as a point at the location of the most significant genetic marker and a line indicating the 95% confidence interval. QTL are colored by the logarithm of the odds (LOD) score, increasing in significance from blue to green to yellow.

intercross lines (RIAILs) (**File S2**) (Andersen et al. 2015a). We then used linkage mapping to identify quantitative trait loci (QTL) that underlie variation in responses across the RIAILs. We detected a total of 246 significant QTL (across 242 traits) from the 384 phenotypes tested (**Figure S2, File S3**).

Although each of the 24 traits tested for each toxin with our high-throughput fitness assay could be representative of a unique biological or developmental process (Andersen et al. 2015a), we observed that some traits were highly correlated (**Figure S1B**). For this reason, many QTL we identified could not be considered independent. Indeed, we identified 50 groups of correlated traits with overlapping QTL confidence intervals. We selected the most significant among correlated (Spearman's *rho* > 0.642) QTL with overlapping confidence intervals per toxin to more accurately identify distinct QTL (Materials and Methods, **Figure S3**). Using this method, we identified 64 traits either with (1) no QTL that overlapped with a different QTL controlling any other trait or (2) QTL that overlapped with another trait but that were not correlated. Linkage mappings of these 64 traits identified 114 distinct QTL across the 16 toxins assayed (**File S7, Figure 1**). We did not find a single toxin-response QTL shared robustly across all of the various traits and toxins tested nor across all traits within any one toxin. However, the majority of distinct QTL on chromosome I were detected in responses to chemotherapeutics. Additionally, each toxin had QTL that underlie trait variation on at least two different chromosomes, highlighting the diverse architectures implicated across traits, even within a single toxin. Despite the seemingly independent distributions of QTL, we found that the majority of the distinct QTL (63%) mapped to chromosomes IV and V.

Both additive and interactive QTL underlie toxin responses

For each of the traits that were impacted by the 114 distinct QTL identified using linkage mapping, we calculated the proportion of broad-sense heritability that could be attributed to additive genetic components (narrow-sense heritability) (**Figure 2A**) as well as the proportion of narrow-sense

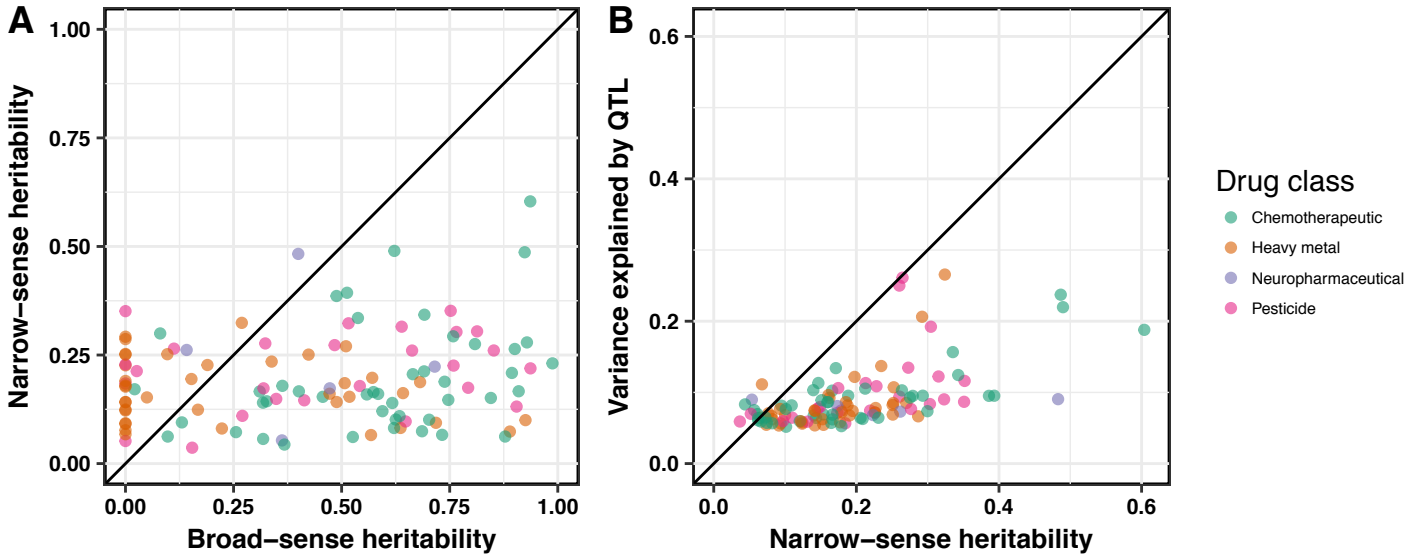

Figure 2 Additive genetic components identified by linkage mapping do not explain all heritable contributions to toxin-response variation. For traits representing the 114 distinct QTL, we compared (**A**) the broad-sense heritability (x-axis) calculated from the dose-response data versus the narrow-sense heritability (y-axis) estimated by linkage mapping and (**B**) the narrow-sense heritability (x-axis) versus the variance explained (VE) by all QTL detected by linkage mapping (y-axis). In both plots, each trait is plotted as a point with the color indicating drug class (chemotherapeutic, heavy metal, neurotherapeutic, or pesticide). The diagonal line represents y = x and is shown as a visual guide.

heritability that was explained by QTL detected through linkage mapping (**Figure 2B, File S5**). In many cases, additive genetic components could not explain the full amount of phenotypic variation predicted to be caused by genetic factors. These results suggest that other additive loci with small effect sizes impact toxin responses, but we failed to detect these QTL by our linkage mapping analyses, potentially because of insufficient statistical power. Alternatively, this missing heritability could be indicative of genetic interactions (Bloom et al. 2013).

To determine how much of the trait variance comes from additive or interacting genetic components, we fit a linear mixed-effect model to the RIAIL phenotype data for the traits controlled by the 114 distinct QTL. We observed a range of additive and epistatic components contributing to phenotypic variation across toxin classes (**Figure S2, Figure S4, File S5**). On average, FUdR, vincristine, and mechlorethamine have a larger percent of their phenotypic variance attributable to genetic interactions than additive effects (**Figure S4**). Alternatively, topotecan, diquat, and chlorpyrifos are primarily explained by additive models (**Figure S4**). To localize potential genetic interactions for these 114 traits, we scanned the genome for interactions between pairs of markers that might affect the phenotypic distribution of the RIAIL panel (Materials and Methods, **File S4**). We identified three significant interactions. This two-factor genome scan was unable to localize all epistatic components identified by the linear mixed-effect model in **Figure 2**, perhaps because of insufficient statistical power to identify small-effect interactions. Using the two-factor genome scan, we also identified 35 traits with significant evidence of two additive loci underlying phenotypic variation. In most of these cases, the mixed-effect model also suggested evidence of additive loci that remain undetected by linkage mapping (**Figure 2B**).

Three QTL hotspots underlie variation in responses to diverse toxins

The majority of toxin-response QTL cluster on chromosomes IV and V (**Figure 1**). We sought to determine if such QTL clustering could be expected by chance or if this clustering is indicative of toxin-response QTL hotspots. To account for the higher rate of recombination, and thus more genetic diversity, on the chromosome arms (Rockman and Kruglyak 2009), we divided the genome into 82 bins of 20 cM each and calculated the number of distinct QTL that mapped to each bin (**Figure 3**,

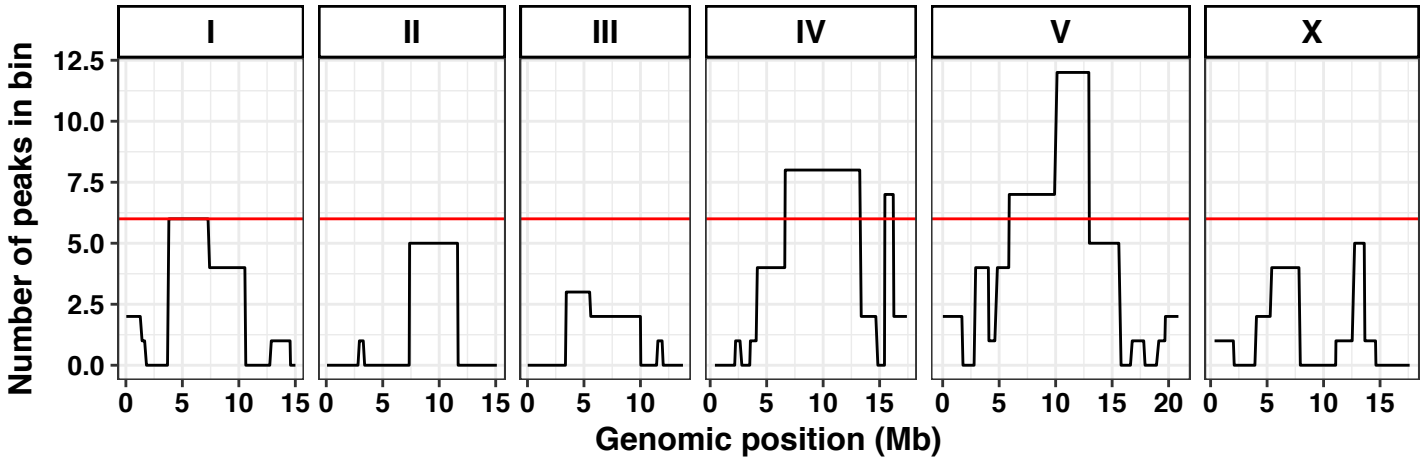

Figure 3 Three QTL hotspots impact condition responses. Each chromosome is divided into bins of 20 cM, resulting in a total of 82 bins across the genome. The x-axis shows the genomic position (Mb), and the y-axis shows the number of the 114 distinct QTL that lie within the corresponding bin. The red line indicates the Bonferroni-corrected 95th percentile of a Poisson distribution with a mean of 1.39 QTL (total QTL/total bins).

Materials and Methods). Three bins with more QTL than expected based on a Poisson distribution (Brem and Kruglyak 2005) were classified as hotspots. Importantly, these hotspots are not driven by correlated toxin-response phenotypes within a single toxin. Instead, hotspots comprise multiple distinct QTL across a variety of traits and toxins. The three hotspots are located on the center of chromosome IV, the right of chromosome IV, and the center of chromosome V. We sought to experimentally validate the predicted additive and epistatic effects on toxin responses for traits that mapped to the three hotspots.

Near-isogenic lines recapitulate some of the predicted QTL effects

To experimentally validate the QTL identified from linkage mapping, we created near-isogenic lines (NILs) for the hotspots on chromosome IV center, IV right, and V center. Each NIL has a small genomic region introgressed from one parental strain into the genome of the opposite parental strain (Materials and Methods). These NILs were whole-genome sequenced and found to match the expected genotype in the hotspot region; however, additional breakpoints were observed (Material and Methods, **File S8**). Because some QTL effects might require high levels of replication to recapitulate expected phenotypic differences, we considered NIL phenotypes for all 242 traits that mapped in the linkage study. Specifically, we tested the NILs in our high-throughput fitness assay for all toxins with more than two traits that have a detected QTL in a hotspot region – a total of 109 traits across 10 toxins (**Figure S2, Table 1, File S9**). Of these 109 traits, 33 did not display a significant phenotypic difference between N2 and CB4856 in the NIL assay. On average, QTL underlying these traits explained less variance ($p < 0.01$) and had less significant LOD scores ($p < 0.001$) than the 77 traits with a significant parental difference (Materials and Methods). One trait (chlorothalonil.q75.TOF) maps to both hotspots on chromosome IV and is duplicated in this dataset.

For the 77 traits in which a significant parental difference was observed, we classified each trait based on the phenotypes of the NILs (**Figure S5, File S11**). We predict that if a single QTL in the introgressed region contributes to the parental phenotypic difference, then each NIL would have a phenotype significantly different than the parental strain with the same genetic background. Furthermore, we would expect each NIL to have a phenotype similar to the parental strain of its introgressed genomic region. NIL phenotypes were consistent with this model for 11 traits, comprising seven of the ten toxins and all three hotspots (**Figure S5**). To identify traits that exhibit no QTL effect, we looked for traits in which the phenotype of each NIL was not significantly different from the phenotype of the parental strain sharing its background genotype. This phenotype indicates that the introgressed NIL region was not affecting the toxin-response phenotype. Nine traits across six

conditions and two hotspots were identified as having no QTL effect, suggesting that either we lacked sufficient statistical power to detect the QTL effect or the real QTL is outside the introgressed region. The phenotypes of the NILs for the remaining 57 traits cannot be explained by a single QTL model. For many of these traits, we observed NIL phenotypes that are more sensitive or more resistant than both parental strains, suggesting that loci of opposite genotypes are acting additively or interacting in the NILs to create transgressive phenotypes (Dittrich-Reed and Fitzpatrick 2013). This finding is supported by the mixed-effects model, which suggests that both additive and interacting QTL remain undetected by linkage mapping (**Figure 2**). We further explored these 57 traits by characterizing them based on the patterns of the transgressive phenotypes we observed.

For 32 of these traits, only one NIL showed a transgressive phenotype (**Figure S2**). These traits comprise eight of the ten toxins and are distributed among all three hotspots. Some of these 32 traits seem to show synergy that increased the effect of the introgressed region (a predicted sensitive phenotype becomes a hypersensitive phenotype or a predicted resistant phenotype becomes a hyper-resistant phenotype, *e.g.* chlorothalonil.mean.EXT, **Figure S2**). Other traits displayed an antagonism that counteracted the effect of the introgressed region (a predicted sensitive phenotype becomes hyper-resistant or a predicted resistant phenotype becomes hypersensitive, *e.g.* diquat.mean.norm.EXT, **Figure S2**). Interestingly, in most cases (~80%), the transgressive phenotype was observed in the strain with the N2 genotype introgressed into the CB4856 background.

In addition to unidirectional transgressive phenotypes, we identified ten traits with suggested bidirectional transgressive phenotypes in which both NILs showed an extreme phenotype compared to the parental strains. These cases comprised three toxins in the hotspots on the center of chromosome IV and the center of chromosome V. Some traits were suggestive of purely antagonistic effects (*e.g.* chlorothalonil.q10.EXT, **Figure S2**) while others suggested an antagonistic effect in one NIL and a synergistic effect in the other (*e.g.* chlorothalonil.q25.norm.EXT, **Figure S2**). We identified no cases of bidirectional synergistic effects. The remaining 15 traits of the 77 with a parental difference did not fall into any of the above categories and were classified as miscellaneous. We do not address these traits any further.

As presented above, toxin responses for 50 groups of correlated traits shared similar mapping architectures. The traits comprising each of these groups might share a mechanism by which QTL underlie toxin response variation, and we would expect them to share a common categorization in the NIL assay. We tested toxin responses for 22 of these 50 correlation groups and investigated the NIL assay-defined classifications for all traits in these 22 groups (**Figure S6**). For seven of the 22 correlated and overlapping QTL groups, all of the traits within the grouping were classified into the same category, supporting our hypothesis that the traits within each of these groups share a common mechanism by which QTL underlie toxin responses. Additionally, traits for 11 of the 22 correlation clusters were split between two categories; ten of these clusters include ‘miscellaneous’ or ‘no parental effect’ categories, suggesting that insufficient power or strict significance thresholds might have prevented some of these traits from being categorized into the same group as their correlated counterparts (*e.g.* chlorothalonil.norm.n and chlorothalonil.n). Interestingly, all 13 tunicamycin-response traits mapped to the same location on chromosome IV and were correlated with each other, thereby placing all of these traits in the same correlation cluster. Nine of these 13 traits displayed unidirectional transgressive phenotypes and the remaining four traits displayed bidirectional transgressive phenotypes. Our strict significance thresholds for categorization might have caused some bidirectional transgressive phenotypes to be classified as unidirectional (*e.g.* tunicamycin.q25.EXT).

Chromosome-substitution strains localize QTL underlying transgressive phenotypes

Because we found evidence of loci where opposite genotypes at each locus cause transgressive phenotypes, we attempted to further characterize these loci (**Figure S5**). To define each set of loci as either intra-chromosomal or inter-chromosomal, we built reciprocal chromosome-

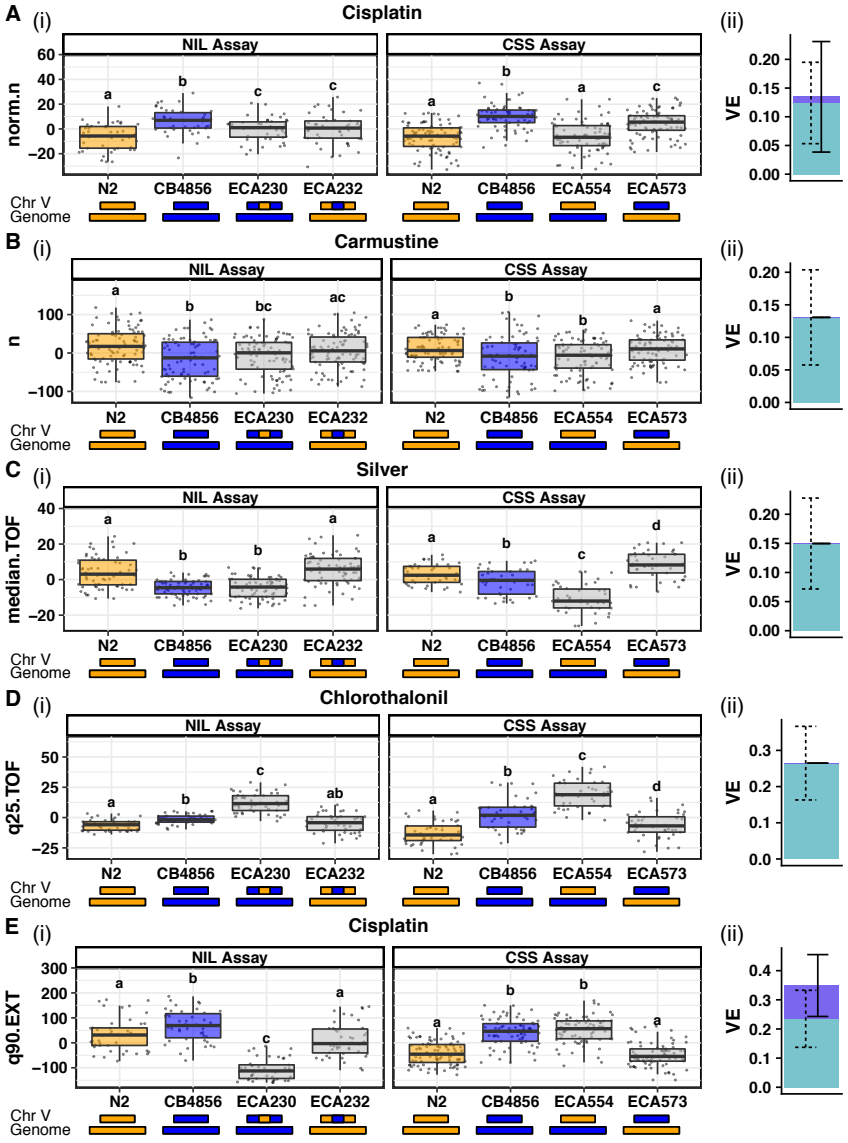

Figure 4 Results from chromosome-substitution strain (CSS) and near-isogenic line (NIL) assays are categorized based on potential variance components implicated in toxin responses. A trait representing each of five categories of variance components (**A**) Recapitulation (cisplatin norm.n), (**B**) No QTL effect (carmustine n) (**C**) Interchromosomal external bidirectional loci (silver median.TOF), (**D**) Interchromosomal internal unidirectional loci (chlorothalonil q25.TOF), and (**E**) Intra-chromosomal unidirectional lci (cisplatin q90.EXT)] is reported. In each case, we show results from (i) the CSS assay (left) and NIL assay (right) plotted as Tukey box plots. The y-axis indicates residual phenotypic values for the given trait. Pairwise comparisons are marked with letters. Different letters (a-d) above each Tukey box plot represent significant differences (Tukey HSD, $p < 0.05$) while the same letter represents non-significant differences among the strains (Tukey HSD, $p > 0.05$). For example, in (A), N2 is significantly different from CB4856, ECA230, and ECA232; CB4856 is significantly different from ECA230 and ECA232; ECA230 and ECA232 are not significantly different from each other. The genotype of each strain on the x-axis is modeled by the colored rectangles beneath the plots (N2 genotypes are orange, CB4856 genotypes are blue). (ii) A stacked bar plot shows the the proportion of phenotypic variance attributable to additive (light blue with dashed error bars) and interactive (dark blue with solid error bars) genetic factors, based on a linear mixed model.

substitution strains (CSSs) for the hotspot on chromosome V that had the entire fifth chromosome introgressed from one parental strain into the genome of the opposite parental strain (Materials and Methods). The hotspot on chromosome V was chosen to isolate the effects of one hotspot and avoid complications arising from traits whose confidence intervals might lie within both of the hotspots on chromosome IV. The CSSs were whole-genome sequenced and found to have the expected genotype at all markers (Materials and Methods, **File S8**), except for the chromosome I incompatibility locus (Seidel et al. 2011, 2008). We tested the CSSs in our high-throughput fitness assay for each of the 56 traits across the five toxins tested with the chromosome V NILs (**Figure S2, Supplemental Information, File S9**).

For traits in which the parental phenotypic difference was significant and consistent across the NIL and CSS experiments, NIL and CSS phenotypes could be compared across assays. Thirteen traits across five toxins fit these criteria and were further analyzed (**File S11**). One trait (cisplatin.norm.n) displayed phenotypic recapitulation of the introgressed region in both the NIL and the CSS assays, suggesting a single QTL model (**Figure 4A**). Additionally, one trait (carmustine.n) exhibited a QTL effect in neither the NIL nor the CSS assay. To identify intra-chromosomal loci that underlie transgressive phenotypes in the remaining 11 traits, we searched for traits that display evidence of either a uni- or bidirectional transgressive phenotype in the NILs but not in the CSSs (**Figure 5A**). This result would suggest that two loci of

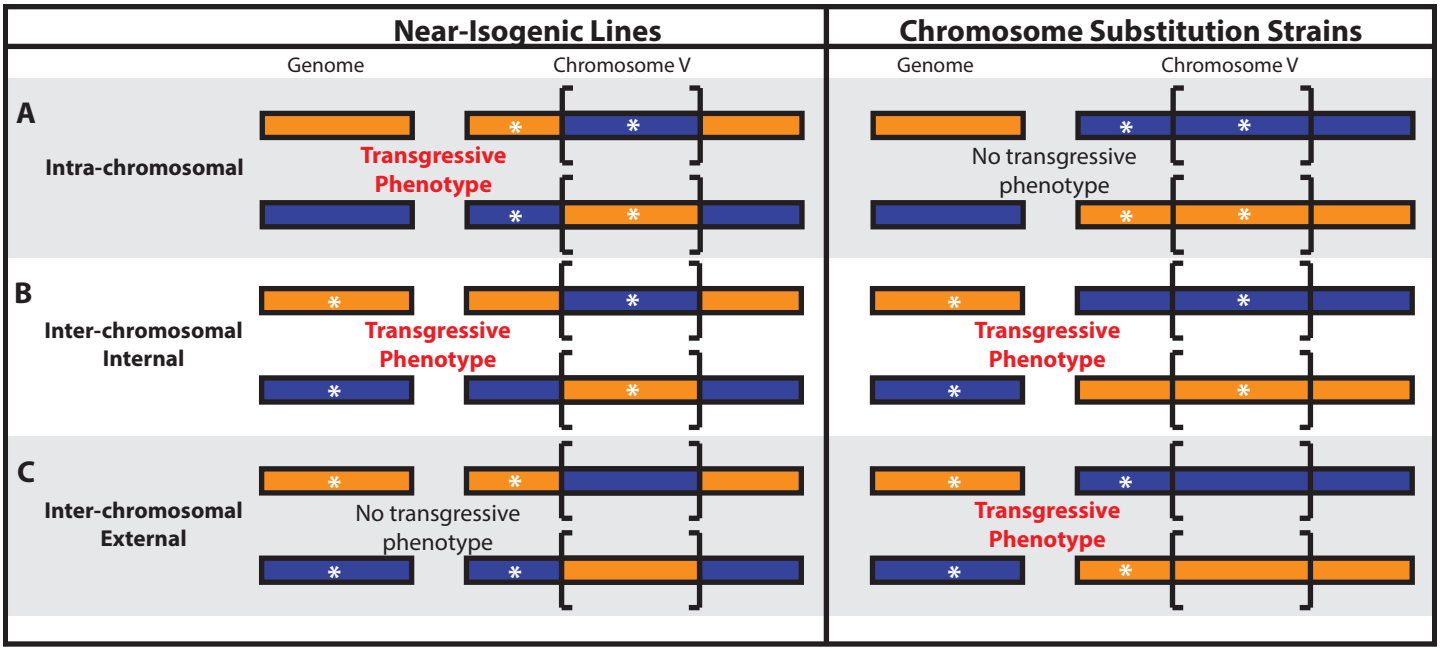

Figure 5 Model of potential additive or interacting loci that could cause hypersensitivity or hyper-resistance in introgressed strains. A model for potential locations of two loci is shown, according to toxin-response phenotypes of near-isogenic lines (NILs) and chromosome-substitution strains (CSSs). The NILs are represented on the left, and the CSSs are represented on the right. The strain genotype is indicated by colored rectangles. N2 is orange, and CB4856 is blue. Brackets indicate the genomic region that is introgressed in the NILs. White asterisks represent a potential location for additive or epistatic loci underlying transgressive phenotypes. Although bidirectional transgressive phenotype models are shown, each model could be bidirectional (both introgressed strains show transgressive phenotypes) or unidirectional (only one introgressed strain shows a transgressive phenotype). Models showing (**A**) intra-chromosomal effects between a locus within and a locus outside of the introgressed region in the NILs, (**B**) internal inter-chromosomal effects between a locus within the introgressed region in the NILs and a locus on another chromosome, and (**C**) external inter-chromosomal effects between a locus outside of the introgressed region in the NILs and a locus on another chromosome are drawn.

opposite genotypes on chromosome V, one within and one outside the region introgressed in the NILs, act additively or epistatically to cause transgressive phenotypes. We found three examples of such intra-chromosomal loci among two different chemotherapeutic compounds (chlorothalonil and cisplatin) (**Figure 4B**). Transgressive phenotypes controlled by inter-chromosomal loci are defined by two loci on separate chromosomes that act additively or epistatically. We further divided this class into two categories: traits for which the chromosome V locus is within the region introgressed in the NILs (internal, **Figure 5B**) and traits for which the chromosome V locus is outside the region introgressed in the NILs (external, **Figure 5C**). For an internal inter-chromosomal model, we expected both the CSSs and the NILs to display the same hypersensitivity or hyper-resistance, because both strains share the same genotype across the introgressed region in the NILs (**Figure 5B**). We identified one such trait with a unidirectional transgressive phenotype (chlorothalonil.q25.TOF) (**Figure 4C**). For an external inter-chromosomal model, we expect only the CSSs to display hypersensitivity or hyper-resistance, because both loci share the same genotype in the NILs (**Figure 5C**) and would therefore not result in a more extreme phenotype than both parents. We found one such trait with a bidirectional transgressive phenotype (silver.median.TOF) (**Figure 4D**). The remaining six traits could not be characterized beyond their NIL assay characterization based on the results of the CSS assay.

We revisited the two-factor genome scan results for each of these 13 empirically classified traits and compared the findings from these two independent methods used to identify multiple additive or epistatic QTL. One trait, silver.median.EXT, shows evidence of a significant secondary additive QTL on chromosome IV using the two-factor genome scan in addition to its detected QTL on chromosome V (**File S4**). Although this trait was placed into the 'miscellaneous' category based on

NIL and CSS phenotypes, further inspection of these phenotypes show suggestive evidence of an inter-chromosomal external effect (**Figure S2**). The strict criteria used to categorize results from the NIL and CSS assays might have led to potential type II errors, such as this example. No traits with significant interaction terms were identified by the two-factor genome scan. Although many other pairs of loci show suggestive evidence of additive or interacting effects (**File S4**), an increase in statistical power is required to definitively compare these suggestive findings to our empirically derived model. Overall, this study highlights the benefits of leveraging both experimental and computational strategies to further dissect genetic components that underlie quantitative traits.

DISCUSSION

Here, we show that three QTL hotspots underlie differences in responses to 16 diverse toxins. We further characterized these QTL using both modeling and empirical approaches. Through the use of near-isogenic lines and chromosome-substitution strains, we confirmed small-effect QTL and attempted to identify and localize genomic regions causing transgressive phenotypes. Finally, we used statistical analyses to computationally identify loci that might support some of our empirical findings. Although the number of biological replicates and recombinant strains in this study increased our power to detect QTL compared to previous animal studies, we are still too underpowered to definitively assess if missing heritability is composed of small additive effects or genetic interactions.

Pleiotropic regions underlie QTL shared between and among toxin classes

We assayed a large number of different fitness traits in a panel of RIAILs and used linkage mapping to identify 246 toxin-response QTL. Some of these QTL are unique to one particular trait, but others suggest the existence of pleiotropic QTL that overlap diverse toxin-response traits. In particular, three QTL hotspots across chromosomes IV and V were enriched for toxin-response QTL and were investigated further. Because the molecular mechanisms implicated in responses to each toxin differ drastically, the notion that a single gene in each hotspot is regulating the response to several toxins is unlikely. However, the possibility exists that a single gene involved in drug transport could underlie one or several of these hotspots. More likely, multiple genes in close proximity, each regulating a process controlling cellular proliferation and survival, might underlie these hotspots. Notably, two of the three QTL hotspots are in swept regions with lower genetic diversity at the species level (Andersen et al. 2012; Laricchia et al. 2017; Cook et al. 2016c, 2016a). The laboratory strain, N2, has experienced each of the selective sweeps, and CB4856 is one of a few strains that has not. Linkage mapping using a panel of RIAILs built between these two strains could identify QTL that underlie phenotypic differences between swept and non-swept strains. Moreover, identifying QTL in these swept regions that underlie variation in fitness-related traits might indicate selective pressures that could have led to these chromosomal sweeps. For example, N2 is more resistant than CB4856 to tunicamycin (**Figure S2**), an antibiotic and chemotherapeutic produced by the soil bacterium *Streptomyces clavuligerus (Price and Tsvetanova 2007).* This result might suggest that selective pressure toward responses to antibiotic compounds played a role in driving resistance-conferring alleles, such as those present in N2, to a high frequency. Alternatively, climate conditions could also impact local niche environments to sensitize toxin responses (Evans et al. 2017). However, we observed that N2 is more resistant than CB4856 in responses to the majority of conditions, which could indicate that alleles present in swept strains confer robustness in responses to many conditions. This result emphasizes the importance of genetic background when considering toxin effects (Zdraljevic and Andersen 2017).

In addition to the three QTL hotspots, pleiotropic QTL across toxins within certain classes are suggested by our linkage mapping results. We observed an enrichment of QTL from the chemotherapeutic class on chromosome I, which could be representative of QTL that underlie a common mechanism targeted by these toxins, such as DNA damage or cell-cycle control. However,

because many of these chemotherapeutics have distinct mechanisms of action and share these mechanisms with other toxin classes, this enrichment is likely caused by an overrepresentation of chemotherapeutics in our study. Direct comparisons of toxins with similar cellular mechanisms could provide more insights. Irinotecan and topotecan are both chemotherapeutics that cause DNA damage by inhibiting topoisomerase I (Pommier 2006) and share a QTL (along with FUdR) on the left arm of chromosome I. FUdR is involved in DNA damage as well, although it does not target topoisomerase I (2017). Taken together, these results might indicate a candidate gene on chromosome I with a role in DNA damage. By contrast, diquat and paraquat are both pesticides that act nearly identically by creating free-oxygen radicals (Dodge and Harris 1970), yet they share no QTL. The distinct genetic architectures underlying these highly similar compounds suggest that genetic variation implicated in responses to diquat and paraquat are specific to each compound and not representative of a general cellular redox mechanism. We have observed this phenomenon for benzimidazole compounds as well (Zamanian et al. 2018b).

A multi-faceted approach suggests that undetected loci impact toxin responses

To determine if we had sufficient power to experimentally validate even small effect sizes, we constructed NILs for the three hotspots and assayed them in responses to multiple toxins. For some of these toxins, NILs showed a significant phenotypic effect, suggesting that our assay had sufficient power to detect small phenotypic effects in at least some cases. We postulated that our inability to recapitulate other predicted phenotypic effects could be attributed to either insufficient power or additional additive or epistatic QTL that were undetected by linkage mapping. Particularly in cases where the NILs displayed transgressive phenotypes, undetected loci of opposite genotypes, acting additively or epistatically, likely caused these effects. Therefore, we investigated these interactions and found evidence for additional QTL that interact with the orginally detected loci. However, we must note that whole-genome sequence data revealed that three of our NILs had a portion of the genome carried over from the background of the starting RIAIL (**File S8**). Although we do not believe that these small regions are responsible for the unexpected phenotypes observed, this explanation could be a consideration for certain silver and chlorothalonil traits, as they have significant QTL in these identified regions. This example emphasizes the importance of whole-genome sequencing NILs to verify the expected genotypes before making conclusions about phenotypic effects of a targeted QTL.

We used the results from the NIL assays to classify each trait into a category that predicts a genetic model that might underlie NIL phenotypes. Phenotypic categorizations were consistent across correlated traits with overlapping QTL. This widespread consistency suggested that similar genetic architectures underlie phenotypes for these grouped traits, lending merit to our methods for identifying distinct QTL and classifying NIL phenotypes. Furthermore, this consistency highlights the reproducibility of our high throughput toxin response assay, because results from independent assays (trait correlations from dose response assays, linkage mappings from RIAIL assays, and phenotype classifications from NIL assays) often align to support the same conclusion gleaned from individual experiments.

The majority of cases of transgressive phenotypes occur when the N2 genotype is introgressed into the CB4856 genome. This trend might indicate allele-specific unidirectional incompatibilities between the two strains, and localizing these interactions could improve our understanding of the evolutionary processes driving such incompatibilities. However, identifying the loci that underlie these unidirectional transgressive phenotypes using a two-factor genomic scan is difficult, because only a small number of the RIAILs have the required allelic combinations to quantify such an effect. For example, cisplatin.q90.EXT fits a unidirectional intra-chromosomal model. The results of the NIL and CSS assays show that, although the CSSs seem to display no QTL effect, the NIL with the N2 genotype introgressed into the CB4856 genome displays strong hypersensitivity (**Figure 4E**). Using our mixed-model approach, we find that about one third of the broad-sense heritability cannot be explained by purely additive effects, potentially implicating undetected epistatic

loci. Additionally, only 59% of the narrow-sense heritability estimated by the mixed-effect model is explained by the detected QTL, suggesting undetected additive loci. A two-dimensional genome scan for multiple loci that underlie this trait provides suggestive evidence for a two-QTL model over a one-QTL model, with or without interaction between the loci. These two loci are located on the left of chromosome V (outside the NIL interval) and in the center of chromosome V (inside the NIL interval) and match our empirical evidence of two intra-chromosomal loci underlying the transgressive phenotype observed (**Figure 5A**). Because the transgressive phenotype is unidirectional, RIAILs without the allelic combination that causes extreme phenotypes could dilute our power to detect the loci. For this reason, combining both computational models and empirical investigation facilitates the detection of loci that control transgressive phenotypes.

Although we are statistically underpowered to identify all small-effect additive and interacting loci, the combination of three methods of searching for potential interactions suggests that not all fitness traits in *C. elegans* are composed of additive effects. Our two computational methods were used to identify additive and epistatic loci underlying many toxin responses, but their power was limited in cases of unidirectional transgressive phenotypes. Alternatively, the NIL and CSS phenotypic assays were able to identify unidirectional transgressive phenotypes, but they were restricted by their inability to distinguish between additive and epistatic loci. Constructing double CSS strains or multi-region NILs in which pairwise combinations of two genomic regions are introgressed within the opposite genotype could help to further define loci underlying transgressive phenotypes. However, each locus must be isolated to determine if the two loci act additively or epistatically. The results from the two-dimensional genome scan might provide insights into where to begin this approach. In cases where all three of our techniques suggested epistasis, we suspect that these QTL are not purely additive. Generating an even larger panel of recombinant strains and assaying a much larger number of biological replicates might allow us to further address the debate about how heritable loci contribute to trait variation in metazoans.

ACKNOWLEDGEMENTS

The authors would like to thank Bryn Gaertner, Samuel Rosenburg, and Robyn Tanny for assistance on mapping drug sensitivities and members of the Andersen lab for helpful comments on this manuscript. This work was supported by the following grants to E.C.A.: National Institutes of Health R01 subcontract to ECA (GM107227), the Chicago Biomedical Consortium with support from the Searle Funds at the Chicago Community Trust, and an American Cancer Society Research Scholar Award (127313- RSG-15-135-01-DD). S.C.B. was supported by the Biotechnology Training Grant (T32GM008449). K.S.E. was supported by the Cell and Molecular Basis of Disease training grant (T32GM008061). J.S.B. was supported by the Howard Hughes Medical Institute. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

LITERATURE CITED

Altshuler D, Daly MJ, Lander ES. 2008. Genetic mapping in human disease. *Science* **322**: 881–888.

Andersen EC, Bloom JS, Gerke JP, Kruglyak L. 2014. A variant in the neuropeptide receptor npr-1 is a major determinant of *Caenorhabditis elegans* growth and physiology. **10**: e1004156.

Andersen EC, Gerke JP, Shapiro JA, Crissman JR, Ghosh R, Bloom JS, Félix M-A, Kruglyak L. 2012. Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. *Nat Genet* **44**: 285–290.

Andersen EC, Shimko TC, Crissman JR, Ghosh R, Bloom JS, Seidel HS, Gerke JP, Kruglyak L. 2015a. A Powerful New Quantitative Genetics Platform, Combining Caenorhabditis elegans High-Throughput Fitness Assays with a Large Collection of Recombinant Strains. *G3* **5**: 911–920.

Andersen EC, Shimko TC, Crissman JR, Ghosh R, Bloom JS, Seidel HS, Gerke JP, Kruglyak L. 2015b. A Powerful New Quantitative Genetics Platform, Combining Caenorhabditis elegans High-Throughput Fitness Assays with a Large Collection of Recombinant Strains. *G3* **5**: 911–920.

Balla KM, Andersen EC, Kruglyak L, Troemel ER. 2015. A wild C. elegans strain has enhanced epithelial immunity to a natural microsporidian parasite. *PLoS Pathog* **11**: e1004583.

Bates D, Mächler M, Bolker B, Walker S. 2014. Fitting Linear Mixed-Effects Models using lme4. *arXiv [statCO]*. http://arxiv.org/abs/1406.5823.

Bendesky A, Pitts J, Rockman MV, Chen WC, Tan M-W, Kruglyak L, Bargmann CI. 2012. Long-range regulatory polymorphisms affecting a GABA receptor constitute a quantitative trait locus (QTL) for social behavior in Caenorhabditis elegans. *PLoS Genet* **8**: e1003157.

Bendesky A, Tsunozaki M, Rockman MV, Kruglyak L, Bargmann CI. 2011. Catecholamine receptor polymorphisms affect decision-making in C. elegans. *Nature* **472**: 313–318.

Bloom JS, Ehrenreich IM, Loo WT, Lite T-LV, Kruglyak L. 2013. Finding the sources of missing heritability in a yeast cross. *Nature* **494**: 1–6.

Bloom JS, Kotenko I, Sadhu MJ, Treusch S, Albert FW, Kruglyak L. 2015. Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. *Nat Commun* **6**: 8712.

Boyd WA, Smith MV, Freedman JH. 2012. Caenorhabditis elegans as a model in developmental toxicology. *Methods Mol Biol* **889**: 15–24.

Breitling R, Li Y, Tesson BM, Fu J, Wu C, Wiltshire T, Gerrits A, Bystrykh LV, de Haan G, Su AI, et al. 2008. Genetical genomics: spotlight on QTL hotspots. *PLoS Genet* **4**: e1000232.

Brem RB, Kruglyak L. 2005. The landscape of genetic complexity across 5,700 gene expression traits in yeast. *Proc Natl Acad Sci U S A* **102**: 1572–1577.

Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**: 889–890.

Bubier JA, Jay JJ, Baker CL, Bergeson SE, Ohno H, Metten P, Crabbe JC, Chesler EJ. 2014. Identification of a QTL in Mus musculus for alcohol preference, withdrawal, and Ap3m2 expression using integrative functional genomics and precision genetics. *Genetics* **197**: 1377–1393.

Cook DE, Andersen EC. 2017. VCF-kit: assorted utilities for the variant call format. *Bioinformatics* **33**: 1581–1582.

Cook DE, Zdraljevic S, Roberts JP, Andersen EC. 2016a. CeNDR, the Caenorhabditis elegans natural diversity resource. *Nucleic Acids Res*. http://dx.doi.org/10.1093/nar/gkw893.

Cook DE, Zdraljevic S, Tanny RE, Seo B, Riccardi DD, Noble LM, Rockman MV, Alkema MJ, Braendle C, Kammenga JE, et al. 2016b. The Genetic Basis of Natural Variation in

Caenorhabditis elegans Telomere Length. *Genetics* **204**: 371–383.

Cook DE, Zdraljevic S, Tanny RE, Seo B, Riccardi DD, Noble LM, Rockman MV, Alkema MJ, Braendle C, Kammenga JE, et al. 2016c. The Genetic Basis of Natural Variation in Caenorhabditis elegans Telomere Length. *Genetics* **204**: 371–383.

Cowley AW Jr. 2006. The genetic dissection of essential hypertension. *Nat Rev Genet* **7**: 829–840.

Crusio WE, Dhawan E, Chesler EJ, Delprato A. 2016. Analysis of morphine responses in mice reveals a QTL on Chromosome 7. *F1000Res* **5**: 2156.

David Clifford And. 2006. The regress function. *R News* **6**: 6–10.

David Clifford And. 2014. *The regress package*.

Dittrich-Reed DR, Fitzpatrick BM. 2013. Transgressive Hybrids as Hopeful Monsters. *Evol Biol* **40**: 310–315.

Dodge AD, Harris N. 1970. The mode of action of paraquat and diquat. *Biochem J* **118**: 43P–44P.

Easton DF, Bishop DT, Ford D, Crockford GP. 1993. Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* **52**: 678–701.

Ehrenreich IM. 2017. Epistasis: Searching for Interacting Genetic Variants Using Crosses. *G3* **7**: 1619–1622.

Ehrenreich IM, Bloom J, Torabi N, Wang X, Jia Y, Kruglyak L. 2012. Genetic architecture of highly complex chemical resistance traits across four yeast strains. *PLoS Genet* **8**: e1002570.

Evans KS, Zhao Y, Brady SC, Long L, McGrath PT, Andersen EC. 2017. Correlations of Genotype with Climate Parameters Suggest Caenorhabditis elegans Niche Adaptations. *G3* **7**: 289–298.

García-González AP, Ritter AD, Shrestha S, Andersen EC, Yilmaz LS, Walhout AJM. 2017. Bacterial Metabolism Affects the C. elegans Response to Cancer Chemotherapeutics. *Cell* **169**: 431–441.e8.

Hasin-Brumshtein Y, Khan AH, Hormozdiari F, Pan C, Parks BW, Petyuk VA, Piehowski PD, Brümmer A, Pellegrini M, Xiao X, et al. 2016. Hypothalamic transcriptomes of 99 mouse strains reveal trans eQTL hotspots, splicing QTLs and novel non-coding genes. *Elife* **5**. http://dx.doi.org/10.7554/eLife.15614.

Highfill CA, Tran JH, Nguyen SKT, Moldenhauer TR, Wang X, Macdonald SJ. 2017. Naturally Segregating Variation atUgt86DdContributes to Nicotine Resistance inDrosophila melanogaster. *Genetics* **207**: 311–325.

Hill WG, Goddard ME, Visscher PM. 2008. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet* **4**: e1000008.

Johnsson M, Jonsson KB, Andersson L, Jensen P, Wright D. 2015. Quantitative Trait Locus and Genetical Genomics Analysis Identifies Putatively Causal Genes for Fecundity and Brooding in the Chicken. *G3* **6**: 311–319.

Keurentjes JJB, Fu J, Terpstra IR, Garcia JM, van den Ackerveken G, Snoek LB, Peeters AJM, Vreugdenhil D, Koornneef M, Jansen RC. 2007. Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. *Proc Natl Acad Sci U S A* **104**: 1708–1713.

Knoch D, Riewe D, Meyer RC, Boudichevskaia A, Schmidt R, Altmann T. 2017. Genetic dissection of metabolite variation in Arabidopsis seeds: evidence for mQTL hotspots and a master regulatory locus of seed metabolism. *J Exp Bot* **68**: 1655–1667.

Lachowiec J, Shen X, Queitsch C, Carlborg Ö. 2015. A Genome-Wide Association Analysis Reveals Epistatic Cancellation of Additive Genetic Variance for Root Length in Arabidopsis thaliana. *PLoS Genet* **11**: e1005541.

Laricchia KM, Zdraljevic S, Cook DE, Andersen EC. 2017. Natural variation in the distribution and abundance of transposable elements across the Caenorhabditis elegans species. *Mol Biol Evol*. http://dx.doi.org/10.1093/molbev/msx155.

Leal-Bertioli SCM, Moretzsohn MC, Roberts PA, Ballén-Taborda C, Borba TCO, Valdisser PA, Vianello RP, Araújo ACG, Guimarães PM, Bertioli DJ. 2015. Genetic Mapping of Resistance to Meloidogyne arenaria in Arachis stenosperma: A New Source of Nematode Resistance for Peanut. *G3* **6**: 377–390.

Lee D, Yang H, Kim J, Brady S, Zdraljevic S, Zamanian M, Kim H, Paik Y-K, Kruglyak L, Andersen EC, et al. 2017. The genetic basis of natural variation in a phoretic behavior. *Nat Commun* **8**: 273.

Mackay TF. 2001. Quantitative trait loci in Drosophila. *Nat Rev Genet* **2**: 11–20.

Mackay TFC. 2015. Epistasis for quantitative traits in Drosophila. *Methods Mol Biol* **1253**: 47–70.

Mäki-Tanila A, Hill WG. 2014. Influence of gene interaction on complex trait variation with multilocus models. *Genetics* **198**: 355–367.

Malmberg RL, Held S, Waits A, Mauricio R. 2005. Epistasis for fitness-related quantitative traits in Arabidopsis thaliana grown in the field and in the greenhouse. *Genetics* **171**: 2013–2027.

Mardis ER. 2017. DNA sequencing technologies: 2006-2016. *Nat Protoc* **12**: 213–218.

Marriage TN, King EG, Long AD, Macdonald SJ. 2014. Fine-mapping nicotine resistance loci in Drosophila using a multiparent advanced generation inter-cross population. *Genetics* **198**: 45–57.

McGrath PT, Rockman MV, Zimmer M, Jang H, Macosko EZ, Kruglyak L, Bargmann CI. 2009. Quantitative mapping of a digenic behavioral trait implicates globin variation in C. elegans sensory behaviors. *Neuron* **61**: 692–699.

Najarro MA, Hackett JL, Smith BR, Highfill CA, King EG, Long AD, Macdonald SJ. 2015. Identifying Loci Contributing to Natural Variation in Xenobiotic Resistance in Drosophila. *PLoS Genet* **11**: e1005663.

Nelson RM, Pettersson ME, Carlborg Ö. 2013. A century after Fisher: time for a new paradigm in quantitative genetics. *Trends Genet* **29**: 669–676.

Peng W, Xu J, Zhang Y, Feng J, Dong C, Jiang L, Feng J, Chen B, Gong Y, Chen L, et al. 2016. An ultra-high density linkage map and QTL mapping for sex and growth-related traits of common carp (Cyprinus carpio). *Sci Rep* **6**: 26693.

Pommier Y. 2006. Topoisomerase I inhibitors: camptothecins and beyond. *Nat Rev Cancer* **6**: 789–802.

Price NPJ, Tsvetanova B. 2007. Biosynthesis of the tunicamycins: a review. *J Antibiot* **60**: 485–491.

R Core Team. 2017. R: A Language and Environment for Statistical Computing. https://www.R-project.org/.

Rockman MV. 2012. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**: 1–17.

Rockman MV, Kruglyak L. 2009. Recombinational landscape and population genomics of Caenorhabditis elegans. *PLoS Genet* **5**: e1000419.

Rockman MV, Skrovanek SS, Kruglyak L. 2010. Selection at linked sites shapes heritable phenotypic variation in C. elegans. *Science* **330**: 372–376.

Rothschild MF, Hu Z-L, Jiang Z. 2007. Advances in QTL mapping in pigs. *Int J Biol Sci* **3**: 192–197.

Schmid T, Snoek LB, Fröhli E, van der Bent ML, Kammenga J, Hajnal A. 2015. Systemic Regulation of RAS/MAPK Signaling by the Serotonin Metabolite 5-HIAA. *PLoS Genet* **11**: e1005236.

Seidel HS, Ailion M, Li J, van Oudenaarden A, Rockman MV, Kruglyak L. 2011. A novel sperm-delivered toxin causes late-stage embryo lethality and transmission ratio distortion in C. elegans. *PLoS Biol* **9**: e1001115.

Seidel HS, Rockman MV, Kruglyak L. 2008. Widespread genetic incompatibility in C. elegans maintained by balancing selection. *Science* **319**: 589–594.

Shang L, Ma L, Wang Y, Su Y, Wang X, Li Y, Abduweli A, Cai S, Liu F, Wang K, et al. 2016. Main Effect QTL with Dominance Determines Heterosis for Dynamic Plant Height in Upland Cotton. *G3* **6**: 3373–3379.

Shimko TC, Andersen EC. 2014. COPASutils: an R package for reading, processing, and visualizing data from COPAS large-particle flow cytometers. *PLoS One* **9**: e111090.

Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal F-X, Durand-Tardif M, Pelletier G, Camilleri C. 2008. Quantitative trait loci mapping in five new large recombinant inbred line populations of Arabidopsis thaliana genotyped with consensus single-nucleotide polymorphism markers. *Genetics* **178**: 2253–2264.

Singh KD, Roschitzki B, Snoek LB, Grossmann J, Zheng X, Elvin M, Kamkina P, Schrimpf SP, Poulin GB, Kammenga JE, et al. 2016. Natural Genetic Variation Influences Protein Abundances in C. elegans Developmental Signalling Pathways. *PLoS One* **11**: e0149418.

Singh UM, Yadav S, Dixit S, Ramayya PJ, Devi MN, Raman KA, Kumar A. 2017. QTL Hotspots for Early Vigor and Related Traits under Dry Direct-Seeded System in Rice (Oryza sativa L.). *Front Plant Sci* **8**: 286.

Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, et al. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* **42**: 565–569.

Zamanian M, Cook DE, Zdraljevic S, Brady SC, Lee D, Lee J, Andersen EC. 2018a. Discovery of genomic intervals that underlie nematode responses to benzimidazoles. *PLoS Negl Trop Dis* **12**: e0006368.

Zamanian M, Cook DE, Zdraljevic S, Brady SC, Lee D, Lee J, Andersen EC. 2018b. Discovery of genomic intervals that underlie nematode responses to benzimidazoles. *PLoS Negl Trop Dis* **12**: e0006368.

Zdraljevic S, Andersen EC. 2017. Natural diversity facilitates the discovery of conserved chemotherapeutic response mechanisms. *Curr Opin Genet Dev* **47**: 41–47.

Zdraljevic S, Strand C, Seidel HS, Cook DE, Doench JG, Andersen EC. 2017a. Natural variation in a single amino acid substitution underlies physiological responses to topoisomerase II poisons. *PLoS Genet* **13**: e1006891.

Zdraljevic S, Strand C, Seidel HS, Cook DE, Doench JG, Andersen EC. 2017b. Natural variation in a single amino acid substitution underlies physiological responses to topoisomerase II poisons. *PLoS Genet* **13**: e1006891.

Zuk O, Hechter E, Sunyaev SR, Lander ES. 2012. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci U S A* **109**: 1193–1198.

2017. NCI Thesaurus. https://ncit.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=NCI_Thesaurus&code=C504 (Accessed October 7, 2017).