A *C. elegans* large-scale genome-wide association study reveals hundreds of quantitative trait loci underlying responses to biomedically relevant therapeutics

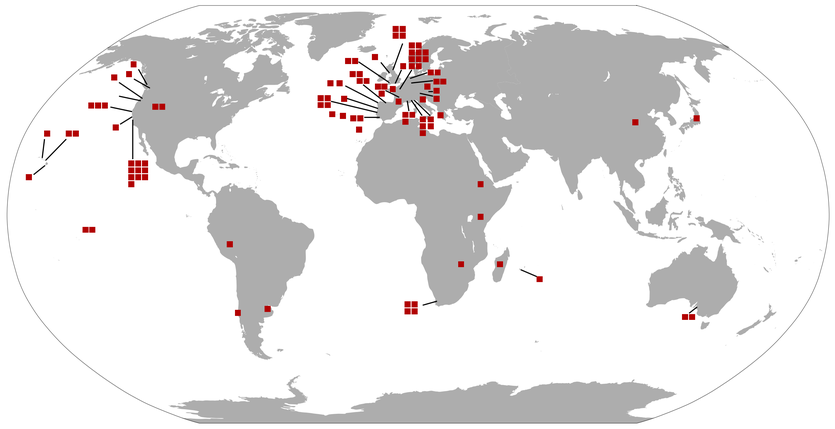
Stefan Zdraljevic[1,2], Sam K. Rosenberg[1], Robyn E. Tanny[1], Tyler C. Shimko[1], and Erik C. Andersen[1,2]

[1] Molecular Biosciences, Northwestern University, Evanston, IL, United States,
[2] Interdisciplinary Biological Sciences Program, Northwestern University, Evanston, IL, United States

Abstract

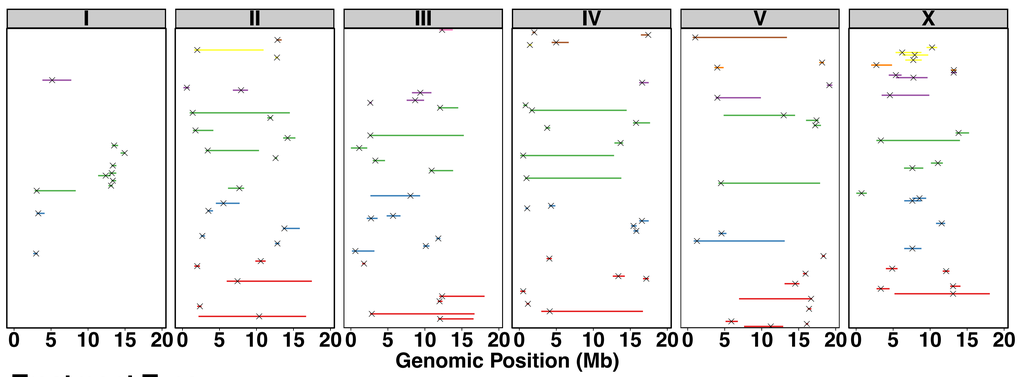
Individuals in a population vary in a wide range of traits, including susceptibilities to diseases and responses to therapeutics. Identifying the genetic determinants of such traits remains difficult in human studies because of prohibitive genotyping costs, inability to control environmental conditions, and the polygenic nature of many traits. *Caenorhabditis elegans* provides a powerful model to probe the genetic determinants of these traits. Importantly, the molecular and genetic toolkits available in *C. elegans* allow us to characterize how genetic variation alters molecular mechanisms. We have optimized a high-throughput and high-accuracy phenotyping pipeline capable of quantifying various fitness traits of 96 genetically distinct strains exposed to 24 different environmental conditions in one week. Altogether, we exposed 96 wild isolates and 359 recombinant inbred lines to 70 different environmental perturbations, including chemotherapeutics, neuroactive compounds, anthelmintics, heavy metals, pesticides, and various temperatures and bacterial food sources. This approach identified more than 500 unique quantitative trait loci (QTL). As a proof of concept for our approach, we recapitulated a previously identified QTL that explains variation in response to the anthelmintic abamectin. Additionally, we identified five novel QTL that explain variation in response to abamectin treatment. We are actively pursuing another QTL, on the right arm of chromosome II that explains variation in response to the topoisomerase II poison etoposide. A scan of genetic variants underlying this QTL identified the candidate gene *top-2*, which encodes for one of the two known topoisomerase II proteins in the *C. elegans* genome. Strains sensitive to etoposide contain a *top-2* gene with several predicted synonymous and non-synonymous variants when compared to the N2 genetic background. A *top-2* deletion in the N2 genetic background failed to complement the sensitive *top-2* variant phenotype, suggesting that *top-2* is the causal gene underlying this QTL. We hypothesize that the Q797M variant present in sensitive strains leads to more stable binding of etoposide to TOPOII and therefore increases its potency as a poison. Our identification of conserved drug susceptibilities between humans and *C. elegans* and our ability to probe the genetic determinants of these susceptibilities has introduced *C. elegans* as a powerful model for elucidating the mechanisms underlying complex traits.

Genetic Diversity Toolkit



Each red dot on the map above corresponds to the location where a wild *C. elegans* strain was isolated. To date we have 124 unique isotopes sequenced at approximately 84X coverage. This substantial level of coverage enables detection of genetic variation with high confidence. We use this genotypic variation in combination with phenotype data generated from our phenotyping platform to make phenotype-genotype associations.

Summary of Identified QTL



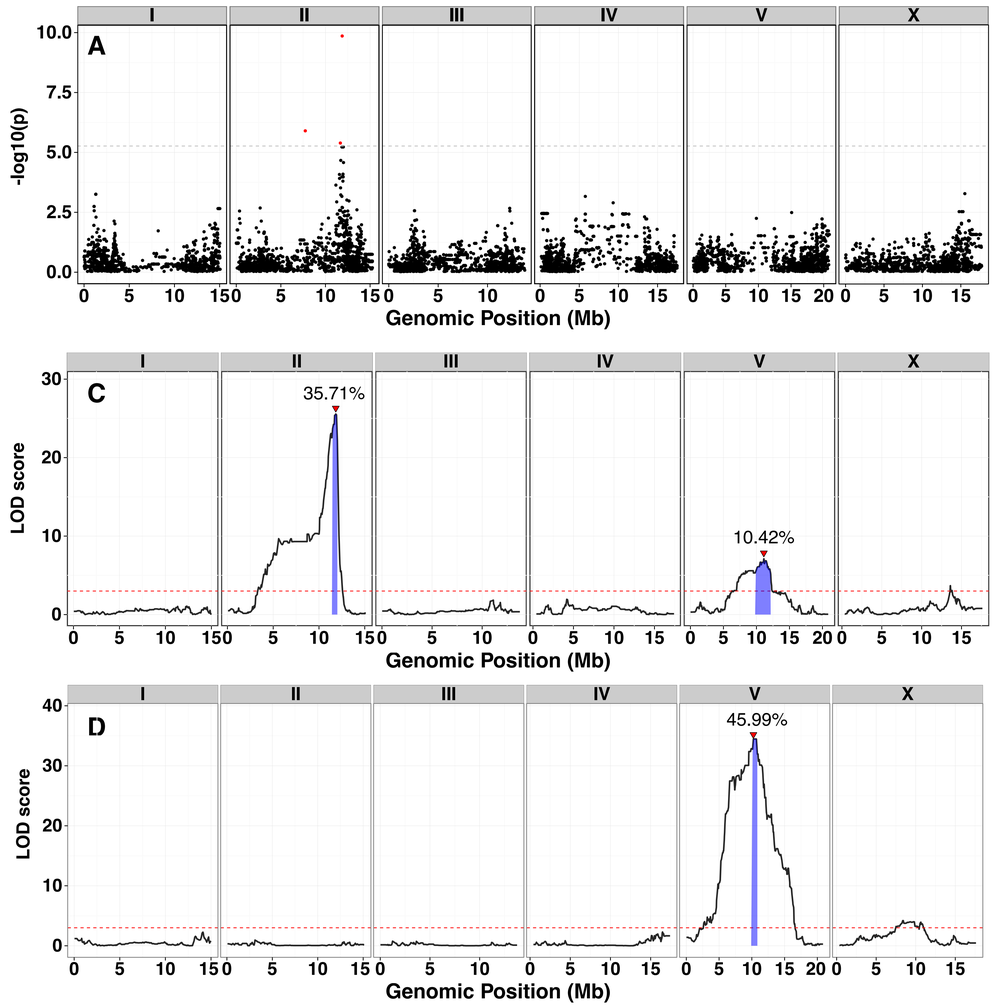
The above figure depicts all of the unique QTL identified in our screen. We quantified the phenotypes of 96 wild isolates in the presence of 70 environmental perturbations. Altogether, we identified 500 QTL. We manually verified that the QTL were not being driven by outliers. After manual curation of the phenotypes at each QTL, we ended up with 81 unique QTL, which are shown in the above figure.

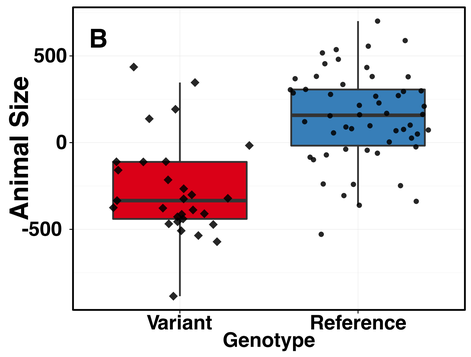
Funding

IBiS Travel Award

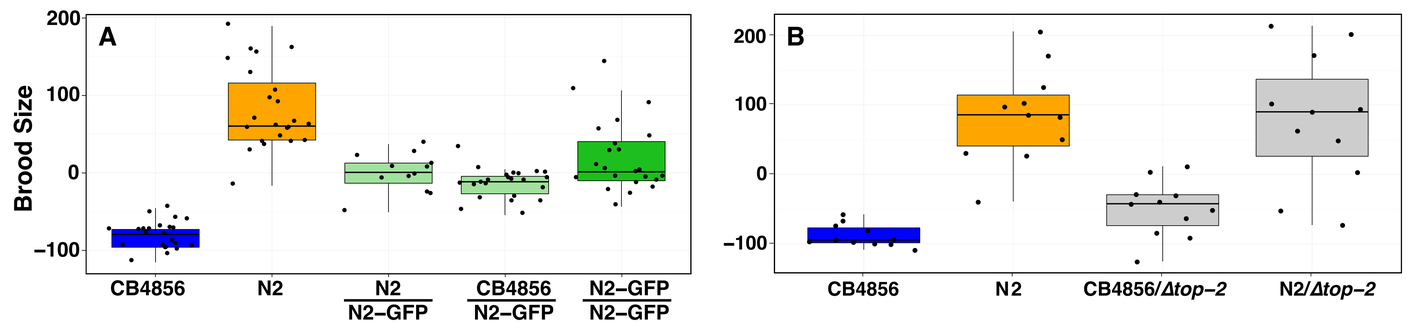
CMBD Training Grant

Variation in *top-2* Leads to Differential Etoposide Sensitivity





A) Manhattan plot showing that the right arm of chromosome II is associated with variation in animal size in response to the topoisomerase II inhibitor etoposide. **B)** Phenotypes of individual strains split by variants in the etoposide confidence interval. Five variants, which are in linkage disequilibrium with each other, in this confidence interval were most highly correlated with etoposide sensitivity and have the same phenotypic split (only one plot is shown). Four out of five are variants in the *top-2* gene, which encodes for the *C. elegans* topoisomerase II. The fourth variant is in the neighboring gene *npp-3,* which encodes for a nucleoporin protein essential for nuclear pore formation. **C)** Linkage mapping data generated by phenotyping 350 recombinant lines generated from N2 and CB4856 parental lines. We identified an additional QTL on chromosome V using this reagent set. **D)** Linkage results showing that phenotypic variation in response to a second topoisomerase II poison, amsacrine, does not map to the right arm of chromosome II. This suggests that the peak on chromosome II is specific to etoposide treatment and will be discussed below.

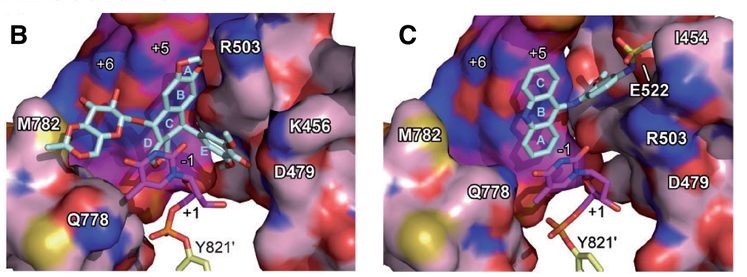


A) Resistance to etoposide is dominant to sensitivity. **B)** We showed that the N2 *Δtop-2* fails to complement the resistance of a natural allele in the CB4856 genetic background for body length reduction, indicating that *top-2* is the causative gene for this phenotypic difference. An N2 *npp-3* knockout strain complemented CB4856 sensitivity, suggesting that it is not the causal gene (not shown).

Future Directions

A

```
Human  TOPOIIα   VKVAQLAGSVAEMSSYHHGEMSLMMTIINLAQN
Human  TOPOIIß   VKVAQLAGSVAEMSAYHHGEQALMMTIVNLAQN
C.e.   TOPOII R  VKVAQLAGAVAEISAYHHGEQSLMGTIVNLAQD
C.e.   TOPOII S  VKVAQLAGAVAEISAYHHGEMSLMGTIVNLAQD
                 780.......790.......800.......810
```



A) Multiple sequence alignment of the two *C. elegans* TOPOII alleles and the human TOPOII alpha and beta isoforms. The variant we hypothesize is contributing to phenotypic variability in response to etoposide in *C. elegans* also varies between the two human isoforms (red). There is debate in the community whether this residue is important for etoposide binding. B-C) Electron density of the topoisomerase II poison binding pocket in hTOPOII beta showing etoposide (B) and amsacrine (C) bound [1]. Wu *et al.* hypothesize that the methionine reside in the hTOPOII alpha leads to more stable binding of etoposide in the binding pocket, while other groups believe this residue is not important for binding because it is in two conformations in the hTOPOII beta crystal structure. We hypothesize that this residue is contributing to etoposide binding and stability and have a physiologically relevant way to quantify this effect. Our hypothesis is supported by variation in response to amsacrine treatment not mapping to the *top-2* gene, suggesting to us that both N2 and CB4856 have normal topoisomerase II function. Additional support for this variant comes from the fact that the other highly correlated variants in the confidence interval are present in the hyper variable C-terminal domain of the protein. We have generated all the constructs to perform CRISPR/Cas9 mediated allele replacement to test our predictions.

Chyuan-Chuan Wu *et al. Nucleic Acids Research* 2013 doi:10.1093/nar/gkt828