Spring 5-1-2013

Influence of Genetic Variation on Birth Defects in Caenorhabditis elegans

Daniel Robert Kepple

Influence of Genetic Variation on Birth Defects in *Caenorhabditis elegans*

A Capstone Project Submitted in Partial Fulfillment of the Requirements of the Renée Crown University Honors Program at Syracuse University

Daniel Robert Kepple
Candidate for B.S. Degree
and Renée Crown University Honors
August 2013

Honors Capstone Project in Biological and Medical Physics

Capstone Project Advisor: _____
Assistant Professor of Physics at Syracuse University
Martin Forstner

Capstone Project Reader: _____
Associate Professor of Applied Physics at Harvard University Daniel Needleman

Honors Director: _____
Stephen Kuusisto, Director

Date: August 1st 2013

Abstract

In my Renee Crown Honors Capstone project, I studied how genetic variation influences birth defects that cause death in *C. elegans* embryos. I performed high-throughput hatching assay experiments of recombinant inbred advance intercross lines of *C. elegans*. These lines are genetically distinct from each other. I found significant variation in birth defects causing embryo death in these recombinant inbred advanced intercross lines. My results give evidence that gene interaction may play a significant role in causing birth defects resulting in death. My data also provides a starting point for studies making statistical arguments linking these birth defects to specific genetic interactions.

Table of Contents

Introduction

The majority of birth defects still have unknown etiologies [1]. Traditional genetics, in which genes are knocked down one at a time, is limited to observing the results of single-gene manipulations. These studies cannot readily show the genetic interactions that might lead to birth defects.

Gene interaction involves networks of causal relationships. These relationships can be determined by examining genetic perturbations[2]. The best kind of experiment to understand and discover these relationships is that which features a system that is perturbed by many random factors[3]. Randomized perturbations also include intrinsic redundancy that is not possible in single gene approaches[4]. So, for experiments aiming to determine causal networks of gene interaction, multifactorial, randomized perturbation approaches are highly advantageous. In biology, these kinds of perturbations are not hard to produce. Genetic variation throughout a population results in such a system[4].

To perform a multifactorial, randomized perturbation study, genetically characterized populations are essential. Typically these populations are generated through advanced intercrossing of two inbred lines, both with homozygous genotypes. This produces lines with highly distinct genomes, having randomly distributed genes from each of the parents[5]. These lines can then be genotyped

with molecular markers. Roles of different genes in the genotype-phenotype causal network are understood using statistical arguments. Those genetic variants linked to phenotypes are called "quantitative trait loci" (QTL)[6].

Caenorhabditis elegans is particularly well suited as a model organism for quantitative trait loci mapping. Inbreeding in *C. elegans* populations is rapid and easy due to androdieocious mating and self fertility. Desirable lines can be cryoperserved, preventing genetic drift[7]. Furthermore, using recombinant inbred advanced intercrossing methods on *C. elegans*, one can achieve a population with a large genetic map[5]. Coupled with increasingly high-density polymorphic markers on the *C. elegans* chromosome, high-precision quantitative trait loci mappings are possible[7].

In this study, I performed high-throughput experiments to test for variation in birth defects resulting in embryo death in genetically distinct lines of *C. elegans*. Further studies of this nature would permit use of statistical arguments linking these kinds of birth defects to interaction of specific genes.

Methods

Population

The *C. elegans* lines used for this study were selected from a group of 200 lines purchased from Matt Rockman's laboratory at NYU. To produce these lines, two genetically distinct parents, Hawaiian CB4856 and laboratory N2, were crossed. The resulting heterozygous lines were randomly mated for several generations, resulting in advanced intercross lines (AIL). Finally, generations of inbreeding in the AIL produced genetically distinct, homozygous lines, known as recombinant inbred advanced intercross lines (RIAIL)[8].

C. elegans Maintenance

The RIAIL used in this study were kept cryopreserved at -80C in 1.8ml cryotubes as a working stock until the time of the experiment. *C. elegans* lines can be preserved in this way for many years[9]. RIAIL were thawed in groups of five or ten by Katherine Kartheiser at Harvard University. They were transferred from their cryotubes to an agar plate by a platinum spatula sterilized in fire. The agar plates were seeded with an *E. coli* OP50 food lawn. Thawed *C. elegans* were stored at 21C. Once a line had repopulated, it was used for the experiment.

Due to variable fitness across the RIAIL, some lines repopulated much more slowly after thawing than others. Lines which reproduced slowly often became contaminated before they were ready for the experiment. To decontaminate these lines, a weak bleach solution was used to wash the contaminated *C. elegans*. This bleach will dissolve *C. elegans* but not their eggs. *C. elegans* that hatched after the bleach evaporated would travel to the opposite end of the plate where a food lawn was seeded. From here, they could be transferred to another agar plate and allowed to reproduce.

Food lawns and Agar Plates

I worked with two different sizes of *E. coli* OP50 food lawns. One lawn was very large, almost covering the entire agar plate. The other lawn was roughly 2cm in diameter and centrally located on the agar plate. Large lawn plates enabled *C. elegans* to reach high populations of about one thousand. Small lawn plates encourage *C. elegans* to stay in a small area which is ideal for imaging.

Both types of agar plates were kept at room temperature overnight to dry out the plates slightly. These dryer plates have no apparent consequence for growth of *C. elegans* populations. Dryer plates do, however, act as an adhesive to *C. elegans* eggs. This keeps *C. elegans* eggs from washing away in a light wash with an M9 buffer.

Synchronization

The standard method for synchronizing the age of a *C. elegans* population is to centrifuge the population with bleach[10]. This bleach will dissolve *C. elegans* organisms but not their eggs, leaving a synchronized population. Bleaching and centrifuging eggs may cause them stress and change their likelihood to hatch successfully. For my experiment, I developed a different, simple method for synchronization which is less harmful to *C. elegans* eggs but no less effective in synchronizing large populations at once.

Once a thawed line repopulated, and the majority of its *C. elegans* population developed to its L4 stage, I washed its agar medium by pipette with an M9 buffer. *C. elegans* on the plate suspend into the M9 buffer wash. I then collected the M9 buffer and suspended *C. elegans* from the plate and pipetted them into a plastic falcon tube. After about thirty seconds to a minute, the *C. elegans* sinks to the bottom of the tube. I then pipetted the M9 dense with *C. elegans* and transferred them directly onto the food lawn of a small-lawn agar plate. These *C. elegans* are left to lay eggs at 21C. Five hours later, I would wash the plate with M9. Because the plates were dried overnight, *C. elegans* eggs stick slightly to the agar, whereas the L4 stage *C. elegans* wash away. Once I pipette and discard the *C. elegans* and M9, only eggs remain on the agar plate. Using a tongue depressing stick sterilized by fire, I would cut the central region containing eggs out from the rest of the agar. I would then transfer this region face-down onto a large agar plate,

producing a synchronized population which I would store at 21C.

Hatching Assay

I used M9 to wash plates with synchronized L4 populations. I transferred suspended *C. elegans* in the M9 by pipette to a falcon tube. After thirty seconds to a minute, I pipetted *C. elegans* from the bottom of the falcon tube. Those with populations greater than a few hundred were pipetted directly to the food lawn of a small-lawn plate. I transferred smaller populations to a large-lawn plate in order to grow more.

After five hours, I washed the small-lawn plates with M9. *C. elegans* suspended in the M9 were disposed of. If there were any *C. elegans* remaining on the plate, I washed it again. Sufficiently washing the plate removes everything but eggs. Because *C. elegans* do not leave the small-food lawn, eggs on the plate are dense and centrally located.

After about an hour, any remaining M9 evaporates and the plates are imaged. Twenty-four hours later, I wash the plates again to remove hatched *C. elegans*. I image the washed plates again. Because embryogenesis in *C. elegans* only takes about fourteen hours, any eggs remaining at this time are dead.

Imaging

Images were taken with a dissecting microscope camera. Using an improvised stage constructed with microscope slides, motion of the agar plate was limited to only x and y movements relative to the camera's view. Further, the improvised stage labeled each region of the agar plate. This setup allowed for the same regions to be imaged before and after hatching. It also prevented the same region and eggs from being imaged twice during the same time period.

After-hatching images were taken only in lines where the number of dead eggs was too high to be easily counted. In most cases, there were only a few eggs per image-region and I counted them rather than imaging them.

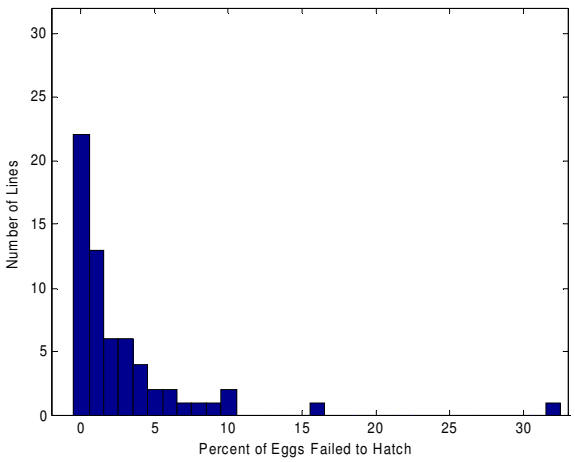
Image Analysis

Before-hatching images of a RIAIL provide the total number of eggs produced by that line. That RIAIL's after-hatching images provide the total number of eggs that died during embryogenesis. The quotient of these two measurements yields the percentage of eggs that died during embryogenesis in a RIAIL. Because each RIAIL had thousands of eggs imaged before hatching, the software program ImageJ was used to count the eggs.

Results

Due to time constraints, I was only able to test 62 of the RIAIL. Of these 62 lines, 38 had 2% or less of their eggs die before hatching. The average percentage of eggs that died during embryogensis in the remaining 24 lines was 6.4%. Two notable outliers were lines that had 16% and 32.5% embryo-death rates. The histogram in Figure 1 summarizes my results.

Figure 1



Redundancy was limited due to time constraints as well. However, four tests on laboratory N2 lines were performed to establish a baseline. All four tests were within .1% of 99.5% hatch success.

Discussion

My results show significant differences in hatch success rate among the RIAIL. This implies that genetic variation has an important role in determining the success of embryogenesis. Further, because these lines were derived from wild type and laboratory *C. elegans*, my results suggest that gene interactions or complex traits are behind the variation in rate of successful embryogenesis rather than single gene mutations.

I suspect that, if this hatching assay were performed on more of the RIAIL, there would be even more variation because of my experimental bias in selecting lines to test. Because of time constraints, I was only able to use lines which were able to quickly reproduce and create large populations. It is quite likely that slow reproduction in of some of the lines is caused by a poor hatching rate.

Because multiple tests could not be performed on each line, it is hard to say what variation in hatch success rate is due to variations in environmental conditions. The high precision of the four tests on N2 suggests that variations in environmental conditions are not very influential. However, it is highly likely that different RIAIL will have different susceptibilities to environmental conditions.

A flaw in my experimental method is that I did not consider failure in fertilization.

An embryo that has failed to fertilize would appear identical to one that has failed in embryogenesis. However, I suspect that failure in fertilization would be roughly consistent among the lines. If this is true, then fertilization would not result in the hatch success rate variation found in this study.

Further Work

This study succeeded in creating a viable protocol for studying hatch success rate in genetically distinct lines of *C. elegans*. It also showed that there is sufficient variation in hatch success rate to warrant further study. However, this study fell short of actually being able to specify any genetic basis that may result in failure in embryogenesis. In order to make the statistical arguments necessary to show causal relationships, one would have to test many more RIAIL. Further, it would be necessary to repeat the experiment multiple times for each RIAIL in order to see both the precision of this experiment and the importance of variations in environmental conditions.

Statistical arguments connecting failure in embryogenesis to genetic interaction are interesting for a few reasons. Firstly, they may help us predict these kinds of failures in humans. Secondly, they might suggest ways to prevent failure in embryogenesis. Thirdly, it might help us to infer the significance of cellular processes in embryogenesis and development. After specifying a genetic basis for a particular birth defect, we can compare that basis to genes known to be important for different cellular processes. If one finds significant overlap, the researcher could then argue that these processes are crucial for successful embryogenesis.

Acknowledgments

I would like to express my deep gratitude to Professor Daniel Needleman for his guidance and mentorship throughout this project. I would also like to show my appreciation for the guidance of Dr. Reza Farhadifar. I would like to thank Katherine Kartheiser for her help in performing experiments for this project. I would like to thank Professor Martin Forstner for his support of this project.

I am very grateful for the generous support of the Renee Crown Honors program. This project was only possible because of their Crown-Wise grant. I would also like to thank the Research Experience for Undergraduates program at the Harvard School of Engineering and Applied Sciences, the National Science Foundation, the Summer Research Opportunities at Harvard program, and the Leadership Alliance Consortium for their financial contributions.

References

[1]Malformations due to presumed spontaneous mutations in newborn infants. Nelson K, Holmes LB.Department of Newborn Medicine, Brigham and Women's Hospital, Boston, MA. *N. Engl. J. Med.* 320(1):19-23. (1989)

[2]Friedman, N., Linial, M., Nachman, I. & Pe'er, D. Using Bayesian networks to analyze expression data. *J. Comput. Biol.* 7, 601–620 (2000).

[3]Fisher R.A. The arrangement of field experiments. *J. Ministry Agric.* 33:503-511 (1926).

[4]Jansen, R. C. Studying complex biological systems using multifactorial perturbation. *Nature Rev. Genet.* 4, 145–151 (2003).

[5]Rockman MV, Kruglyak L. Recombinational Landscape and Population Genomics of *Caenorhabditis elegans*. *PLoS Genet* 5(3): e1000419. doi:10.1371/journal.pgen.1000419 (2009).

[6]Mackay, Trudy F C. The genetic architecture of quantitative traits. *Annual Review of Genetics* 35: 303-39 (2001).

[7]Gaerner, Bryn E; Phillips, Patrick C. Caenorhabditis elegans as a platform for

molecular quantitative genetics and the systems biology of natural variation. *Genetics Research* 92. 5-6 : 331-348. (2010).

[8]Breeding Designs for Recombinant Inbred Advanced Intercross Lines. Matthew V. Rockman and Leonid Kruglyak.

[9]Eisenmann, D. M., Wnt signaling (June 25, 2005), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.7.1, http://www.wormbook.org.

[10]Sulston, J. and Hodgkin, J. (1988). In: *The Nematode Caenorhabditis elegans*, W.B. Wood, ed. (New York: Cold Spring Harbor Laboratory Press), p. 587.

Capstone Summary:

Description of project

In this project I performed experiments to determine the percentage of embryos that have fatal birth defects in different *C. elegans* lines. *C. elegans* is a small nematode model organism. The different lines of *C. elegans* used in my experiment were bred so that they are highly distinct genetically. Because environmental conditions were controlled in my experiments, any dramatic variation in the percentage of embryos with fatal birth defects is likely caused by genetic variation across the lines.

Nontechnical Description of Methods

The *C. elegans* lines were purchased from Matt Rockman's lab in NYU. Each line has been bred so that it has a random assortment of genes of laboratory *C. elegans* and *C. elegans* found in Hawaii. The laboratory *C. elegans* was originally found in the UK, but is now distinct from the wild UK *C. elegans* due to years of isolation in the laboratory. The laboratory *C. elegans* and the Hawaiian *C. elegans* have been separated for far longer has a very different set of genes.

My experiment looked at several thousand eggs of about 60 of the *C. elegans*

lines and monitored whether or not they hatched. If an egg does not hatch after one day, it is dead. I took a picture of the eggs just after they were laid, and then one day later. The number eggs remaining after one day divided by the number of eggs laid yields the percentage of embryos with fatal birth defects.

Significance

My data showed that there was significant variation between the percentage of birth defects in genetically distinct *C. elegans* lines. Because the lines contain genes from either laboratory and Hawaiian *C. elegans*, which have very low percentages of fatal birth defects, it is unlikely that this variation comes from single gene mutations. Instead, it likely comes from interaction of multiple genes.

In addition to showing that genetic interaction may be responsible for significant numbers of birth defects, my study also provides a method and initial measurements for further studies on this topic. Further studies would permit statistical arguments linking birth defects to interaction of specific genes. These arguments could be used to help predict birth defects in humans. They may also suggest ways to prevent birth defects. Finally, identifying genetic bases for birth defects could implicate the importance of different biological functions in cell division and development. That is, if the genetic basis for a certain birth defect contains genes that are involved in a certain cellular process, one could argue that this cellular process is crucial for successful development.