Synthetic lethal analysis of Caenorhabditis elegans posterior embryonic patterning genes identifies conserved genetic interactions

L Ryan Baugh*, Joanne C Wen*, Andrew A Hill†, Donna K Slonim†§, Eugene L Brown† and Craig P Hunter*

Addresses: *Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA. †Department of Genomics, Wyeth Research, Cambridge, MA 02140, USA. ¢Current address: Biology Division, California Institute of Technology, Pasadena, CA 91125. §Current address: Department of Computer Science, Tufts University, Medford, MA 02155.

Correspondence: Craig P Hunter. E-mail: hunter@mbc.harvard.edu

Published: 11 April 2005
Received: 21 October 2004
Revised: 24 December 2004
Accepted: 9 March 2005

© 2005 Baugh et al.; licensee BioMed Central Ltd
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Phenotypic robustness is evidenced when single-gene mutations do not result in an obvious phenotype. It has been suggested that such phenotypic stability results from 'buffering' activities of homologous genes as well as non-homologous genes acting in parallel pathways. One approach to characterizing mechanisms of phenotypic robustness is to identify genetic interactions, specifically, double mutants where buffering is compromised. To identify interactions among genes implicated in posterior patterning of the Caenorhabditis elegans embryo, we measured synthetic lethality following RNA interference of 22 genes in 15 mutant strains. A pair of homologous T-box transcription factors (tbx-8 and tbx-9) is found to interact in both C. elegans and C. briggsae, indicating that their compensatory function is conserved. Furthermore, a muscle module is defined by transitive interactions between the MyoD homolog hld-1, another basic helix-loop-helix transcription factor, hnd-1, and the MADS-box transcription factor unc-120. Genetic interactions within a homologous set of genes involved in vertebrate myogenesis indicate broad conservation of the muscle module and suggest that other genetic modules identified in C. elegans will be conserved.

Background

Forward and reverse genetic screens in flies and worms indicate that most genes are not essential to the development or viability of the organism [1-4]. Two primary explanations for such phenotypic robustness to mutation have been offered: homologous gene products may directly compensate for one another’s function, or indirect compensation of function may occur through regulatory networks via non-homologous genes acting in alternative pathways or feedback mechanisms [5,6]. In both Saccharomyces cerevisiae and C. elegans, genes with at least one homolog are less likely than unique genes to have a loss-of-function phenotype [7,8]. However, homology accounts for no more than two thirds of the observed phenotypic robustness to mutation in S. cerevisiae and even less in C. elegans, indicating a significant role of the regulatory network in genetic buffering [7,8]. By identifying gene disruptions that are viable in wild type but lethal in a specific mutant background, synthetic lethal screens can shed light on how regulatory networks buffer gene function [9]. Although genetic interactions are being mapped on a
genome-wide scale in yeast [9], no such efforts have been reported for an animal system. We report here the use of existing mutants and RNA interference (RNAi) to assemble a synthetic lethal matrix in C. elegans. Our aim was to build on prior knowledge by focusing on a characterized set of genes likely to have interactions among them.

The C blastomere is one of five somatic founder blastomeres in the C. elegans embryo. Each founder blastomere produces a characteristic set of cell types via an invariant lineage; the C blastomere predominantly gives rise to muscle and epidermal cells [10]. Maternal and zygotic activities of the homeodomain protein PAL-1 specify the identity and maintain the development of the C-blastomere lineage [11,12]. We have identified genes whose expression depends either directly or indirectly on PAL-1 function (‘PAL-1 targets’) and shown that they affect specification, differentiation, and morphogenesis of C-lineage cells. PAL-1 targets include many transcriptional factors and based on their temporal and spatial expression patterns in wild type, we have proposed a model for their regulatory relationships [13]. However, RNAi of most PAL-1 targets does not result in a detectable phenotype [13]. Because the regulatory network specified by PAL-1 appears to function autonomously while controlling multiple, discrete developmental processes (that is, specification and morphogenesis of muscle and epidermal cells) [11,13,14], we reasoned that the network is organized in a modular fashion and that PAL-1 targets may have overlapping or compensatory functions. We set out to measure synthetic lethality for as many pairs of PAL-1 targets as possible, and we included three additional genes encoding transcriptional regulators (lin-26, unc-62, and ceh-40) whose expression patterns and phenotypes suggest that they may have genetic interactions with PAL-1 or its targets [15,16].

Results and discussion
A synthetic lethal matrix
We began our analysis with 26 genes, 23 PAL-1 targets and three genes likely to have genetic interactions with PAL-1 targets. Mutations in 15 of these genes have been isolated, and appropriate strains were obtained from the Caenorhabditis Genetics Center (Table 1). To screen for synthetic lethal interactions, RNAi of the 13 nonessential genes was performed and embryonic lethality scored in each of the mutant strains as well as wild type to generate a synthetic lethal matrix (Figure 1 and Additional data file 1). However, RNAi of a given gene could not be used if it results in 100% lethality in wild type (pal-1, elt-1, unc-62, and lin-26), and so the matrix is not entirely symmetric. To reduce the penetrance of lethality following RNAi for these four genes we used a lower concentration of double-stranded RNA (dsRNA), but the results were variable (data not shown), and we decided that it was an unreliable approach. We also tried to assemble the matrix using exclusively RNAi, but we found that soaking worms in two different dsRNA molecules targeting two different genes at once to be unacceptably inefficient (Additional data file 2).

Given the effects of mutation alone without RNAi, and the effects of RNAi in wild type, it may not be obvious from the raw data which disruptions result in significantly more lethality when paired than when alone. Measurements of lethality were thus duplicated and a $t$-test was used to assign statistical significance to each of the interactions tested (Figure 1 and Additional data file 1). Seven genetic interactions were identified ($P < 0.001$) among six genes out of 195 total interactions tested. In addition to the genes shown in Figure 1, RNAi of the remaining genes for which no mutant was available was performed in all 16 strains. Because synthetic lethality was measured only a single time for each of these possible interactions, the data could not be subjected to statistical analysis. However, RNAi of these additional genes did not appear to elevate lethality in any of the genetic backgrounds tested (data not shown).

Conservation of a genetic interaction
The T-box family of proteins regulates diverse aspects of animal development, including cell fate specification, differentiation, and morphogenesis [17]. The C. elegans genome contains 21 of these animal-specific genes, compared to 32 in C. briggsae and eight in D. melanogaster [18,19], suggesting duplication or retention within nematodes. tbx-37 and tbx-38 have been shown to redundantly control embryonic patterning of the C. elegans AB lineage [20], and our systematic approach identified tbx-8 and tbx-9 as a synthetic lethal pair. tbx-8 and tbx-9 are expressed in nearly identical temporal and spatial patterns in the early embryo, with the strongest expression detected in dorsal epidermal cells and posterior expression depending on pal-1 function [13,21,22]. tbx-8(ok656), tbx-8(RNAi), and tbx-9(RNAi) each result in a low frequency of embryonic lethality (less than 5%) with 1-10% of hatching larvae displaying posterior morphological defects (Figure 2) [21,22]. However, 50% of tbx-8(ok656); tbx-9(RNAi) embryos fail to hatch and those that do display severe morphological defects and die as larvae (Figure 2). Phenotypic analysis demonstrates that tbx-8 and tbx-9 are required for intercalation of dorsal epidermal cells thus explaining the morphological defects observed following their disruption [22].

The overlapping function of tbx-8 and tbx-9 is not the residual result of a recent gene duplication event. The C. briggsae genome contains a pair of syntenic orthologs (by best reciprocal blastp match) [18], and as in C. elegans, the C. briggsae orthologs have overlapping function during morphogenesis (Table 2), indicating that the functional relationship between this pair of genes has been selectively maintained for over 100 million years. This observation suggests that the overlapping functions of other pairs of T-box genes, which are likely to be common throughout the Metazoa, may also be conserved.
A muscle differentiation module

It has been suggested that modules of genes comprise units of biological function [23]. Because they operate in concerted fashion, genes in a module are expected to have increased likelihood of being co-expressed and sharing physical and genetic interactions relative to a random set of genes. Clustering a large synthetic lethal matrix in *S. cerevisiae* identified multiple functional modules [9], and we identify a module in our matrix around the *hlh-1* gene (Figure 1). We detect five genetic interactions (*P* < 0.001) out of six tested among *hlh-1*, *hnd-1*, and *unc-120*. All three genes encode transcription factors expressed exclusively in the muscle progenitor cells of
Table 1

Genes included in the synthetic lethality matrix

<table>
<thead>
<tr>
<th>ORF</th>
<th>Gene</th>
<th>Identification</th>
<th>Genotype used in this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>F17A2.5</td>
<td>ceh-40</td>
<td>Homeodomain protein (Extradenticle ortholog)</td>
<td>ceh-40(ok740) X</td>
</tr>
<tr>
<td>K10B4.6</td>
<td>cwn-1</td>
<td>Putative Wnt ligand</td>
<td>cwn-1(ok546) II</td>
</tr>
<tr>
<td>D1081.2</td>
<td>unc-120</td>
<td>MAD5 domain TF</td>
<td>unc-120(stx364) I</td>
</tr>
<tr>
<td>B0304.1</td>
<td>hh-1</td>
<td>bHLH TF</td>
<td>hh-1(cc56) II</td>
</tr>
<tr>
<td>C44C10.8</td>
<td>hnd-1</td>
<td>Hand bHLH TF</td>
<td>hnd-1(q740) X</td>
</tr>
<tr>
<td>F35G12.6</td>
<td>mab-21</td>
<td>Highly conserved novel protein</td>
<td>mab-21(bx53/ll; him-5(e1490) V</td>
</tr>
<tr>
<td>F11C1.6</td>
<td>nhr-25</td>
<td>Nuclear hormone receptor</td>
<td>+/szT1 [lon-2(e678) I; nhr-25(mj2389)/szT1 X</td>
</tr>
<tr>
<td>Y75B8A.2</td>
<td>nab-1</td>
<td>Homeodomain TF (posterior Hox paralog)</td>
<td>nab-1(st223) dpy-18(e364) unc-25(stx156) III; eDp6 (III;4)</td>
</tr>
<tr>
<td>K02B9.4</td>
<td>ek-3</td>
<td>GATA TF</td>
<td>ek-3(gk121) X</td>
</tr>
<tr>
<td>T07C4.2</td>
<td>tbx-8</td>
<td>T-box TF (Brachyury)</td>
<td>tbx-8(ak656) III</td>
</tr>
<tr>
<td>M142.4</td>
<td>vab-7</td>
<td>Homeobox TF (even-skipped subfamily)</td>
<td>vab-7(e1562) III</td>
</tr>
<tr>
<td>T07C4.6</td>
<td>tbx-9</td>
<td>T-box TF (Brachyury)</td>
<td>Mutant not available</td>
</tr>
<tr>
<td>C55C2.1</td>
<td>Zinc-finger TF</td>
<td></td>
<td>Mutant not available</td>
</tr>
<tr>
<td>C38D4.6</td>
<td>pat-1</td>
<td>Homeobox TF (cad subfamily)</td>
<td>pat-1(cc224) dpy-17(e164) ncl-1(e1865) unc-36(e251) III; sDp3(III); lin-2(e1309) X</td>
</tr>
<tr>
<td>W09C2.1</td>
<td>ek-1</td>
<td>GATA TF</td>
<td>ek-1(zu180) unc-43(e408)/unc-24(e138) dpy-20(e1282) IV</td>
</tr>
<tr>
<td>F18A1.2</td>
<td>lin-26</td>
<td>Zinc-finger TF</td>
<td>ln-26(e156) II</td>
</tr>
<tr>
<td>T28F12.2</td>
<td>unc-62</td>
<td>Meis-class homeodomain (Homothorax ortholog)</td>
<td>unc-62(e444) V</td>
</tr>
<tr>
<td>C28C12.7</td>
<td>spp-10</td>
<td>Predicted prosaposin</td>
<td>Mutant not available</td>
</tr>
<tr>
<td>R02D3.1</td>
<td>Dehydrogenase</td>
<td></td>
<td>Mutant not available</td>
</tr>
<tr>
<td>C09D4.2</td>
<td>Uncharacterized</td>
<td></td>
<td>Uncharacterized</td>
</tr>
<tr>
<td>R07C1.11</td>
<td>Uncharacterized</td>
<td></td>
<td>Uncharacterized</td>
</tr>
<tr>
<td>T22B7.3</td>
<td>Uncharacterized</td>
<td></td>
<td>Uncharacterized</td>
</tr>
<tr>
<td>C46H11.2</td>
<td>Uncharacterized</td>
<td></td>
<td>Uncharacterized</td>
</tr>
<tr>
<td>ZK1307.1</td>
<td>Uncharacterized</td>
<td></td>
<td>Uncharacterized</td>
</tr>
<tr>
<td>C01G6.1</td>
<td>aqp-2</td>
<td>Aquaporin</td>
<td>Mutant not available</td>
</tr>
<tr>
<td>T27D12.1</td>
<td>Sodium/phosphate transporter</td>
<td>Mutant not available</td>
<td></td>
</tr>
</tbody>
</table>

Most of the genes were included in the matrix because they had been genetically determined to be PAL-1 targets (that is, their expression either directly or indirectly depends on PAL-1 function) and were therefore presumed to have genetic interactions among them. Three additional genes encoding transcriptional regulators (lin-26, ceh-40, and unc-62) were included as their phenotypes and expression patterns suggest that they may have genetic interactions with PAL-1 targets. TF, transcription factor.

the early embryo [13,24,25]. Disruption of function by mutation or RNAi of any one of the three genes results in a low frequency of embryonic lethality with embryos arresting paralyzed at the two-fold stage (Pat) (Figure 3). Embryos require muscle contraction to elongate and pass from the two-fold to three-fold embryonic stage, and the Pat phenotype results from a complete failure of muscle differentiation [26,27]. A large fraction of animals that do hatch after disruption of hh-1, unc-120, or hnd-1 show a less severe phenotype (uncoordinated, dumpy larvae) likely reflecting partial muscle function [26,27] (Figure 3). However, disruption of function of any two of the three genes significantly elevates the frequency of Pat embryos, with disruption of hh-1 and unc-120 resulting in 100% Pat embryos (Figure 3). The fact that hh-1, unc-120, and hnd-1 have identical individual and synthetic phenotypes, in addition to their identical early embryonic expression patterns, indicate that they comprise a muscle-differentiation module. Although fewer genes define this module than those identified in larger screens in yeast [9], the specificity of its defining phenotype indicates that hh-1, unc-120, and hnd-1 function in concert, and it is likely that future synthetic screens will identify additional members of the module.

If both RNAi and genetic disruptions of gene function are null, the synthetic lethal matrix should be symmetric. However, we can only test the pair of reciprocal interactions when mutations of both genes are available and RNAi of neither gene results in 100% lethality in wild type. Although there is some symmetry in the matrix around hh-1 (Figure 1), the interactions detected are not all reciprocal (Figure 3). The efficiency of RNAi varies from gene to gene and often fails to result in null phenotypes, suggesting a significant number of false negatives among the interactions tested and explaining the imperfect symmetry among the interactions detected. The pattern of interactions defining the muscle module suggests that hh-1 is the most essential of the three genes, or that its function is the most potent, and hnd-1 the least (Figure 3), a model that accounts for the reported lack of genetic interac-
tion between heterozygous *hlh-1* and homozygous *hnd-1* mutations [25].

The functional importance of the muscle module we identified is underscored by its conservation. Well characterized genetic interactions exist among the vertebrate homologs of *hlh-1* - the myogenic basic helix-loop-helix (bHLH) regulatory factors MyoD, myf5, myogenin, and MRF4 [28]. In addition, interactions between vertebrate myogenic bHLH genes and homologs of *unc-120* (the MEF2 group of MADS-box regulators) have been described [29]. Interactions among and between vertebrate myogenic bHLH and MADS-box regulators is likely to be the result of cooperative activation of common target genes as well as their ability to activate each other transcriptionally [28-30]. It is possible that *hlh-1*, *hnd-1*, and *unc-120*, like their vertebrate counterparts, display modular genetic interactions because they share common target genes, and they may also activate each other transcriptionally.

Conclusions

We show that the facility of RNAi in *C. elegans* can be used to screen a large number of genes for synthetic phenotypes. Using an unbiased experimental approach we identified both easy-to-predict genetic interactions between similar co-expressed genes and difficult-to-predict genetic interactions between non-homologous genes. We also used reverse genetics in *C. briggsae* to show that genetic interaction between a pair of co-expressed homologous transcription factors (*tbx-8* and *tbx-9*) is conserved, suggesting an adaptive role of overlapping function as opposed to the transient result of gene duplication. Furthermore, we identify a genetic module defined by reciprocal interactions among three transcription factors that is required for differentiation of muscle. A similar network among homologous genes is conserved in vertebrates, suggesting that a single master regulator cannot robustly control muscle development. Given the wealth of existing mutants and the ease of RNAi in *C. elegans*, our approach can be scaled up to identify the core set of metazoan developmental genetic interactions.
Materials and methods
DNA templates for in vitro transcription were prepared by amplification of 100-1000 base pair (bp) sequences from cDNA by PCR. PCR primers were designed to amplify unique sequences in order to ensure specificity of RNAi (Additional data file 3), and a minimal T7 promoter sequence was added to the 5' end of both primers. PCR DNA was purified on a DNA Clean and Concentrator-5 column (Zymo Research), and a high-yield in vitro transcription kit (Promega) was used to produce dsRNA. The dsRNA was reannealed by heating to 90°C and then cooling 6°C/min to 25°C before organic extraction with a 1:1 mixture of phenol and chloroform. Following ethanol precipitation, dsRNA was dissolved in Soaking Buffer [31] at high concentration, which was then measured by spectrophotometer and diluted to 5 mg/ml.

Eight to twelve larval stage 4 (L4) worms were soaked in 0.75 µl 5 mg/ml dsRNA for 24 h at 20°C and then transferred to a recovery plate with Escherichia coli strain OP50 for 12 h at 25°C and then transferred to a score plate with OP50. After 12 h at 25°C on the score plate, adults were removed, and after another 24 h at 25°C, embryos and larvae were counted under a dissecting microscope (with the exception of unc-120(st364) which is temperature-sensitive and was therefore scored at 15°C). For strains carrying viable alleles, embryonic lethality was scored for the homozygous progeny of homozygous parents. For strains carrying 100% lethal alleles, embryonic lethality was scored for the mixed progeny of either heterozygous or hemizygous parents. For RNAi in C. briggsae, 1 mg/ml dsRNA was injected into the germline of young wild-type adults (strain AF16), and the worms were rescued and scored in the same way as the soaked C. elegans.

For statistical analysis of the data, percent embryonic lethality was converted to survival (1 - % embryonic lethality). All survival values were normalized by the survival of N2 (wild type) in Soaking Buffer (no dsRNA). For each genetic interaction tested, the null hypothesis used is that the survival of the double disruption (mutation and RNAi) equals the product of survivals for each single disruption (mutation alone × RNAi alone). The pairs of duplicate survival measure-
ments for each single disruption are multiplied to generate four values for the expected survival of the double disruption, and a t-test is used to determine if the observed survival (n = 2) is significantly different from expected (n = 4). For example, if two genes function independently, and the survival following disruption of each alone is 90%, then our model expects 81% survival following double disruption since the 90% of individuals surviving one disruption have a 10% chance of dying from the other disruption. If survival is significantly different from the expected 81% then we conclude that the two genes interact.

For microscopy, embryos were mounted on 2% agar pads under coverslips and allowed to develop in a humid box for 12-15 h at 25°C. A Zeiss Axiophot with 100× differential interference contrast optics was used and digital images captured.

Additional data files
The following additional data are available with the online version of this paper. Additional data file 1 contains replicate embryonic lethality measurements and corresponding P-values for synthetic lethality. Additional data file 2 contains data showing that double RNAi by soaking worms in dsRNA for two genes at once is an inefficient means of synthetic genetic analysis. Additional data file 3 lists primer sequences used in this study to amplify coding sequences from cDNA for use as templates for dsRNA.

Acknowledgements
We thank the Kimble lab for providing strain JK3276, and the Caenorhabditis Genetics Center for providing the other strains used in this work. This work was funded by an NIH grant GM64429 to C.P.H.

References


