Genetic Analysis of a Metazoan Pathway using Transcriptomic Phenotypes

enetic analysis of molecular pathways has traditionally been performed through epistatic analysis. Epistasis occurs when two genes interact, either directly (biochemical interaction of their gene products) or indirectly. If two genes interact, and the mutants of these genes have a quantifiable phenotype, the double mutant will have a phenotype that is not the sum of the phenotypes of the single mutants that make up its genotype. Epistatic analysis remains a cornerstone of genetics today (Phillips 2008).

Previous work in *S. cerevisiae* and *D. discoideum* using microarrays has shown that transcriptomes can be used to infer genetic relationships in simple eukaryotes (Hughes et al. 2000; Van Driessche et al. 2005). Additionally, eQTL studies in *C. elegans* and *Drosophila melanogaster* have established the usefulness of transcriptomic phenotypes for population genetics studies **???**. Developments in the area of transcriptomics have brought forward new protocols, such as RNA-Seq (Mortazavi et al. 2008), and have also made important progress towards cheaper sequencing (Metzker 2010), better and faster abundance quantification (Patro, Mount, and Kingsford 2014; Bray et al. 2016; Patro et al. 2016) and improved differential analysis of gene expression (Pimentel et al. 2016; Trapnell et al. 2013). As a result, RNA-Seq has been successfully used to identify genetic modules involved in a variety of processes, including T-cell regulation (Singer et al. 2016; Shalek et al. 2013), the *C. elegans* linker cell migration (Schwarz, Kato, and Sternberg 2012), or planarian stem cell maintenance (Van Wolfswinkel, Wagner, and Reddien 2014; Scimone et al. 2014). For the most part, the role of transcriptional profiling has been restricted to target gene identification. In cell culture, single-cell RNA-seq has seen significant progress towards using transcriptomes as phenotypes with which to test genetic interactions  (Adamson et al. 2016; Dixit et al. 2016). More recently, we have shown the first identification of a developmental state of *C. elegans* using whole-organism transcriptome profiling (Angeles-Albores, Leighton, et al. 2016).

To investigate the ability of transcriptomes to serve as quantitative phenotypes, we selected mutants in the *C. elegans* hypoxia pathway for transcriptome sequencing. Metazoans depend on the presence of oxygen in sufficient concentrations to support aerobic metabolism. Genetic pathways evolved to rapidly respond to any acute or chronic changes in oxygen levels at the cellular or organismal level. These oxygen sensitive pathways are involved in a broad range of human pathologies and they have been subject to investigation biochemical and genetic approaches (Semenza 2012). These approaches identified the Hypoxia Inducible Factors (HIFs) as an important group of oxygen responsive genes.

Hypoxia Inducible Factors are highly conserved in metazoans (Loenarz et al. 2011). A common mechanism for hypoxia-response induction is heterodimerization between a HIF and a HIF subunit. The heterodimer then initiates transcription of target genes (Jiang et al. 1996). The number and complexity of HIFs varies throughout metazoans, with humans having three HIF subunits and two HIF subunits, whereas in the roundworm *Caenorhabditis elegans* (*C. elegans*) there is a single HIF gene, *hif-1* (Jiang, Guo, and Powell-Coffman 2001) and a single HIF gene, *ahr-1* (Powell-Coffman, Bradfield, and Wood 1998). HIF target genes have been implicated in a wide variety of cellular and extracellular processes such as glycolysis, extracellular matrix modification, autophagy and immunity (Semenza et al. 1994; Bishop et al. 2004; Shen et al. 2005; Bellier et al. 2009; Semenza 2012).

Levels of HIF proteins tend to be tightly regulated. Under conditions of normoxia, HIF-1 exists in the cytoplasm and partakes in a futile cycle of continuous protein production and rapid degradation (Huang et al. 1996). HIF-1 is hydroxylated by three proline hydroxylases in humans (PHD1, PHD2 and PHD3) but is only hydroxylated by one proline hydroxylase (*egl-9*) in *C. elegans* (Kaelin and Ratcliffe 2008). HIF-1 hydroxylation increases its binding affinity to Von Hippel Lindau Tumor Suppressor 1 (VHL-1), which allows ubiquitination of HIF-1 leading to its subsequent degradation. In *C. elegans*, EGL-9 activity is inhibited by binding of CYSL-1, and CYSL-1 activity is in turn inhibited at the protein level by RHY-1, possibly by post-translational modifications to CYSL-1 (Ma et al. 2012).

Here, we show that transcriptomes contain strong, robust signals that can be used to infer relationships between genes in complex metazoans by reconstructing the hypoxia pathway in *C. elegans* using RNA-Seq. Furthermore, we show that the phenomenon of phenotypic epistasis, a hallmark of genetic interaction, holds at the molecular systems level. We also demonstrate that transcriptomes contain sufficient information, under certain circumstances, to order genes in a pathway using only single mutants. Finally, we were able to identify genes that appear to be downstream of *egl-9* and *vhl-1*, but are almost certainly not targets of *hif-1*. Using a single set of genome-wide measurements, we were able to observe and quantitatively assess significant fraction of the known transcriptional effects of *hif-1* in *C. elegans*. A complete, interactive version of the analysis is also available at <www.wormlabcaltech.github.io/mprsq>.

# Results

## The hypoxia pathway controls thousands of genes in *C. elegans*

We performed whole-organism RNA-seq of the hypoxia pathway at a moderate sequencing depth ( 7 million mapped reads for each individual replicate) under normoxic conditions, which allowed us to measure 13,598 isoforms across all replicates and genotypes, which constitutes over half of all isoforms in *C. elegans*. In spite of the low sequencing depth, transcriptome profiling of the hypoxia pathway revealed that this pathway controls thousands of genes in *C. elegans*. The *egl-9* transcriptome showed differential expression of 1,487 genes, similarly to the 1,816 genes differentially expressed in *rhy-1* mutants. The *vhl-1* transcriptome showed considerably fewer differentially expressed genes (605), possibly reflecting the known fact that it is a weaker controller of *hif-1* than *egl-9* (Shao, Zhang, and Powell-Coffman 2009). The *egl-9*;*vhl-1* double mutant transcriptome showed 1,989 differentially expressed genes. The *hif-1* mutant also showed a transcriptomic phenotype involving 481 genes. The *egl-9*;*hif-1* double mutant showed a similar number of genes with altered expression (364).

## Clustering visualizes epistatic relationships between genes

As a first step in our analysis, we analyzed our data using a general linear model (see [methods]) on logarithm-transformed counts. Genes that are significantly altered between wild-type and a given mutant have a genotype coefficient () that is statistically significantly different from 0. We refer to these coefficients through the greek letter. These coefficients are not identical to the average log-fold change per gene, although they are loosely related to this quantity. Larger magnitudes of correspond to larger perturbations. These coefficients can be used to study the RNA-Seq data in question.

Clustering is a well-known technique in bioinformatics that is used to identify relationships between high dimensional data points (Yeung, Medvedovic, and Bumgarner 2003). We wanted to make sure that clustering by differential expression yielded genetically relevant information. *hif-1* exhibits no obvious phenotypes under normoxic conditions, in contrast to *egl-9*, which exhibits an egg-laying (*egl*) phenotype in the same environment. *egl-9*; *hif-1* mutants suppress the *egl* phenotype. If transcriptomic phenotypes behave similarly to their macroscopic counterparts, *hif-1* should cluster with the *egl-9*; *hif-1* double mutant, whereas *egl-9* should cluster away from the *hif-1* mutant. Indeed, when blind, unsupervised clustering was performed on the data, three clusters emerged. *hif-1* and *egl-9*;*hif-1* clustered together, indicating suppression of the *egl-9* phenotype; whereas *egl-9*, *egl-9*;*vhl-1*, *vhl-1* and *rhy-1* all clustered separately. Finally, our negative control *fog-2* was in its own cluster (see Fig. [fig:dendrogram]). We conclude that expression data contains enough signal to cluster genes in a meaningful manner in complex metazoans.

![ Unsupervised aggregative clustering of various C. elegans mutants. Genes cluster in a manner that is biologically intuitive. Genes that inhibit hif-1 (i.e, egl-9, vhl-1, and rhy-1) cluster far from hif-1. hif-1 clusters with the suppressed egl-9; hif-1 double mutant. A a mutant fog-2 transcriptome, used as an outgroup, clusters farthest away. ](data:application/pdf;base64,)w

## Reconstruction of the hypoxia pathway from first genetic principles

Having shown that the signal in the mutants we selected was strong enough to cluster mutants using the regression coefficients, we set out to reconstruct the hypoxia pathway from first genetic principles. In general, to reconstruct a pathway, we must assess whether two genes act on the same phenotype (independence); then we must measure whether these genes act additively or epistatically on the measured phenotype; and if there is epistasis we must measure whether it is positive or negative, in order to assess whether the epistatic regulation is a genetic suppression or a synthetic interaction.

### Genes in the hypoxia mutant act on the same transcriptional phenotype

We observed that all the hypoxia mutants had significant overlap between their differentially expressed transcriptomes relative to a wild-type control (fraction of shared transcriptomes ranged from a minimum of 65 genes shared between *hif-1* and *egl-9*;*hif-1* to a maximum of 1,249 shared genes between *egl-9* and *egl-9*;*vhl-1*). For comparison, we also analyzed a previously published *fog-2* transcriptome (Angeles-Albores, Leighton, et al. 2016). *fog-2* is involved in masculinization of the *C. elegans* germline, which enables sperm formation, and has not been described to be involved in the hypoxia pathway. The hypoxia pathway transcriptomes and the *fog-2* transcriptome showed similar overlap as the hypoxia pathway to itself (123–618 genes). Given the similar overlaps between known interactors and an unknown transcriptome, we conclude that the *fog-2* mutant we studied acts on the same phenotype as mutants from the hypoxia pathway.

![ Strong transcriptional correlations can be identified between genes that share a positive regulatory connection. We took the egl-9 and the rhy-1 transcriptomes, identified differentially expressed genes common to both transcriptomes and ranked each gene according to its differential expression coefficient \beta. We then plotted the rank of each gene in rhy-1 versus the rank of the same gene in the egl-9 transcriptome. The result is an almost perfect correlation. Green, transparent large points mark inliers to the regression (blue line); red, opaque, small points mark outliers to the regression. The two furthest outliers are annotated as pseudogenes in WormBase. ](data:application/pdf;base64,)

Strong transcriptional correlations can be identified between genes that share a positive regulatory connection. We took the *egl-9* and the *rhy-1* transcriptomes, identified differentially expressed genes common to both transcriptomes and ranked each gene according to its differential expression coefficient . We then plotted the rank of each gene in *rhy-1* versus the rank of the same gene in the *egl-9* transcriptome. The result is an almost perfect correlation. Green, transparent large points mark inliers to the regression (blue line); red, opaque, small points mark outliers to the regression. The two furthest outliers are annotated as pseudogenes in WormBase.

Although overlapping transcriptomes may be enough to conclude that a set of mutants share a phenotype, we wanted to know whether we could draw out more information from looking at quantitative agreement between perturbations. To this end, we rank-transformed the regression coefficients for each transcriptome, and calculated lines of best fit using Bayesian regression with a Student-T distribution to mitigate noise from outliers (see Fig [fig:genetic\_interactions]). For transcriptomes associated with the hypoxia pathway, we found that these correlations tended to have values as high as 0.98 with a tight distribution around the line of best fit, whereas the correlations for mutants from the hypoxia pathway with the *fog-2* mutant were considerably weaker, with magnitudes between 0.6–0.85 and a considerably larger spread around the line of best fit. Although *hif-1* is known to be genetically repressed by *egl-9*, *rhy-1* and *vhl-1* (Epstein et al. 2001), all the correlations between these genes and *hif-1* were negative. The overlap between *hif-1* and all other genes was small, and each overlap involved different sets of genes, which suggests that we did not sequence deeply enough to identify the nature of these positive interactions. After we calculated the pairwise correlation between each transcriptome, we weighted the result of each regression by the number of differentially expressed isoforms shared by two transcriptomes and divided by the total number of differentially expressed isoforms present in the two transcriptomes, . The weighted regressions recapitulated a network with three ‘modules’: A control module, a responder module and an uncorrelated module (see Fig. [fig:heatmap]). We were able to identify a strong positive interaction between *egl-9* and *rhy-1*. The magnitude of this weighted correlation is derived from the fact that the transcriptomes for these genes consisted of 1,487 and 1,816 significantly altered genes respectively and the overlap between both genes was extensive, which makes the weighting factor considerably larger than other pairs. Likewise, the weak correlation between *hif-1* and *egl-9*, *vhl-1* and *rhy-1* is at least partially the result of its relatively weak transcriptomic phenotype relative to the other genes, particularly *egl-9* and *rhy-1*. The fine-grained nature of transcriptional phenotypes means that these weighted correlations between transcriptomes of single mutants are predictive of genetic interaction.

![ A: Heatmap showing pairwise regression values between all single mutants. B: Correlation network drawn from the diagram. Edge width is proportional to the logarithm of the magnitude of the weighted correlation between two nodes divided by absolute value of the weighted correlation value of smallest magnitude. Edges are also colored according to the heatmap in A. ](data:application/pdf;base64,)

**A**: Heatmap showing pairwise regression values between all single mutants. **B**: Correlation network drawn from the diagram. Edge width is proportional to the logarithm of the magnitude of the weighted correlation between two nodes divided by absolute value of the weighted correlation value of smallest magnitude. Edges are also colored according to the heatmap in **A**.

### A quality check of the transcriptomic data reveals excellent agreement with the literature

One way to establish whether genes are acting additively or epistatically to each other is to perform qPCR of a reporter gene in the single and double mutants. This approach was used to successfully map the relationships within the hypoxia pathway (see, for example (Shao, Zhang, and Powell-Coffman 2009; Shen, Shao, and Powell-Coffman 2006)). A commonly used reporter is *nhr-57*, which is known to exhibit large changes in expression upon induction of HIF-1(Shen, Shao, and Powell-Coffman 2006; Shen et al. 2005; Ackerman and Gems 2012; Park et al. 2012). Likewise, *rhy-1* and *egl-9* are both known to be up-regulated when HIF-1 becomes common in the cell (Powell-Coffman 2010).

Our dataset enables us to perform an equivalent computational experiment to qPCR by selectively looking at expression of a few genes at a time. Therefore, we queried the changes in expression of *rhy-1*, *egl-9*, *nhr-57* and *lam-3* as a negative control. In our dataset, this gene be upregulated in *egl-9*, *rhy-1* and *vhl-1*, but remains unchanged in *hif-1*. The *egl-9*;*vhl-1* had an expression level similar to *egl-9*; whereas the *egl-9*;*hif-1* mutant showed suppression of the reporter expression. All of these interactions reflect the literature.

![ Top: In silico qPCR. We extracted four genes (rhy-1, egl-9, nhr-57 and lam-3, shown on the x-axis) and plotted their regression coefficients, \beta, as measured for every genotype (represented by one of six colors) to study the epistatic relationships between each gene. Stars above a bar represent a regression coefficient statistically significantly different from 0, meaning that expression is altered relative to a wild-type control. Error bars show standard error of the mean value of \beta. nhr-57 is an expression reporter that has been used previously to identify hif-1 regulators (Shen, Shao, and Powell-Coffman 2006; Shao, Zhang, and Powell-Coffman 2009). The nhr-57 mRNA levels replicate what is observed in the literature. lam-3 is shown here as a negative control that should not be altered by mutations in this pathway. The increases in the levels of egl-9 and rhy-1 when repressors of hif-1 are knocked out are in agreement with previous literature (Powell-Coffman 2010). We measured modest increases in the levels of rhy-1 mRNA when hif-1 is knocked out. The mechanism behind this is unclear. Negative and positive feedback loops from hif-1 into its inhibiting genes could be a homeostatic mechanism. ](data:application/pdf;base64,)

**Top**: *In silico* qPCR. We extracted four genes (*rhy-1*, *egl-9*, *nhr-57* and *lam-3*, shown on the x-axis) and plotted their regression coefficients, , as measured for every genotype (represented by one of six colors) to study the epistatic relationships between each gene. Stars above a bar represent a regression coefficient statistically significantly different from 0, meaning that expression is altered relative to a wild-type control. Error bars show standard error of the mean value of . *nhr-57* is an expression reporter that has been used previously to identify *hif-1* regulators (Shen, Shao, and Powell-Coffman 2006; Shao, Zhang, and Powell-Coffman 2009). The *nhr-57* mRNA levels replicate what is observed in the literature. *lam-3* is shown here as a negative control that should not be altered by mutations in this pathway. The increases in the levels of *egl-9* and *rhy-1* when repressors of *hif-1* are knocked out are in agreement with previous literature (Powell-Coffman 2010). We measured modest increases in the levels of *rhy-1* mRNA when *hif-1* is knocked out. The mechanism behind this is unclear. Negative and positive feedback loops from *hif-1* into its inhibiting genes could be a homeostatic mechanism.

We also performed *in silico* qPCR of every gene under scrutiny to get a clearer idea of the relationships between them (see Fig. [fig:qpcr]). We observed changes in *rhy-1* expression consistent with previous literature (Shen, Shao, and Powell-Coffman 2006) when *hif-1* is activated. We also observed changes in *egl-9* expression when *egl-9* was mutated. *egl-9* is known as a hypoxia responsive gene (Powell-Coffman 2010) Although changes in *egl-9* expression were not statistically significant in *rhy-1* and *vhl-1* mutants, the mRNA levels of *egl-9* trended towards increased expression in these genotypes. As with *nhr-57*, the *egl-9* and *rhy-1* expression phenotypes were abrogated in the *egl-9*;*hif-1* mutant; whereas the *egl-9*;*vhl-1* mutant showed expression phenotypes identical to the *egl-9* mutant. Our dataset also shows that knockout of *hif-1* resulted in a modest increase in the levels of *rhy-1*. This suggests that *hif-1* is also a negative regulator of *rhy-1*, which constitutes a novel observation. Taken together, these results indicate that RNA-seq data is at least equivalent to qPCR for purposes of comparing gene expression of a reporter between genotypes. Using a single reporter we would have been able to reconstruct an important fraction of the genetic relationships between the genes in the hypoxia pathway.

### Genes in the hypoxia pathway exhibit genome-wide epistasis

As we have shown, it may be sufficient to extract the regression coefficients of a previously known reporter gene and study just that pattern in order rebuild a genetic pathway from RNA-seq data. However, we felt that by relying on a single gene, or even a handful of genes to rebuild the pathway was throwing out all of the valuable information present in our dataset. Therefore, we decided to explore a new epistatic metric—genome-wide epistasis.

Ideally, any measurement of genome-wide epistasis should conform to certain expectations. First, it should make use of the regression coefficients of as many genes as possible. Second, it should be summarizable in a single, well-defined number. Third, it should have an intuitive behavior, such that the special values of the statistic (maximum, minimum, zero) should have an unambiguous interpretation.

One way of defining genome-wide epistasis is to use linear regressions to describe the relationship between the change in expression for a set of genes caused by a single mutant and the change in expression in the same set of genes caused by a double mutant containing the single mutant. The set of genes to be studied can be defined as the set of differentially expressed genes common to both genotypes. Once the set is defined, the regression coefficient of each gene in the single mutant can be plotted against the difference between the regression coefficients of the double mutant and the single mutant. We reasoned that under ideal conditions, such a plot would have an intuitive explanation. If two genes are acting entirely independently of each other, the plot will show a line with slope equal to 0. This is because is acting on entirely different sets of genes, so the perturbation caused a gene is unchanged in the double mutant, . If two genes are acting only additively, then the plot will show a line with slope (and in fact, the slope should be equal to the slope between a plot of the single mutants and ). If two genes share a negative regulatory interaction, then epistasis will be reflected in the plot as a line with a negative slope that should approach , because the double mutant, should have regression coefficients near 0, such that the y-axis becomes equal to . If the two genes have a synthetic interaction, we would expect that the slope must be positive and it must be greater than the slope predicted by an additive model.

In our experiment, we studied two double mutants, *egl-9*;*hif-1* and *egl-9*;*vhl-1*. We wanted to understand how well the global epistasis agreed with the literature based on qPCR of single reporters. Therefore, we fit weighted linear regressions to each of the four possible combinations (*egl-9* vs. *egl-9*;*hif-1*; *hif-1* vs. *egl-9*;*hif-1*; *egl-9* vs. *egl-9*;*vhl-1*; and *vhl-1* vs. *egl-9*;*vhl-1*) to measure the slopes of the lines of best fit.

We observe that the *egl-9*;*vhl-1* mutant has an identical phenotype to the *egl-9* single mutant (slope = 0; see Table. [tab:double\_mutant\_comparison]). On the other hand, *vhl-1* has a positive slope, indicating that *egl-9* is additive to *vhl-1*. However, this positive slope has to be less than the slope that would be predicted by an additive model because the slope between *egl-9*;*vhl-1* and *egl-9* is not statistically different from zero. Partial additivity indicates that *egl-9* inhibits *hif-1* in *vhl-1*-dependent and independent manners, which has been documented in the literature (Shao, Zhang, and Powell-Coffman 2009).

Response Modeling of Double Mutants to Single Mutants

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Double Mutant | Single Mutant |  | SE | p-value |
| 1. *egl-9*;*vhl-1* | *egl-9* |  |  |  |
| 2. *egl-9*;*vhl-1* | *vhl-1* |  |  |  |
| 3. *egl-9*;*hif-1* | *egl-9* |  |  |  |
| 4. *egl-9*;*hif-1* | *hif-1* |  |  |  |

[tab:double\_mutant\_comparison]

![ The mutant egl-9 transcriptomic phenotype is suppressed by mutations in hif-1. The graph shows the \beta coefficients for egl-9 in the x-axis, and the change in \beta coefficient between the egl-9;hif-1 and egl-9 mutant. The dotted line is the regression line between the complete egl-9 and egl-9;hif-1 shared transcriptome. For clarity, only genes that were differentially expressed in the egl-9, rhy-1, vhl-1, hif-1 and egl-9;vhl-1 datasets are shown. These points constitute a very high-quality subset of the measured hypoxia response, as each isoform was identified as differentially expressed in 5 independent genotypes. The single outlier near (6, 0) is nog-1. It is probably downstream of egl-9, and is not likely a hif-1 target. ](data:application/pdf;base64,)

The mutant *egl-9* transcriptomic phenotype is suppressed by mutations in *hif-1*. The graph shows the coefficients for *egl-9* in the x-axis, and the change in coefficient between the *egl-9*;*hif-1* and *egl-9* mutant. The dotted line is the regression line between the complete *egl-9* and *egl-9*;*hif-1* shared transcriptome. For clarity, only genes that were differentially expressed in the *egl-9*, *rhy-1*, *vhl-1*, *hif-1* and *egl-9*;*vhl-1* datasets are shown. These points constitute a very high-quality subset of the measured hypoxia response, as each isoform was identified as differentially expressed in 5 independent genotypes. The single outlier near (6, 0) is *nog-1*. It is probably downstream of *egl-9*, and is not likely a *hif-1* target.

On the other hand, comparison of the *egl-9*;*hif-1* double mutant showed suppression of the *egl-9* transcriptomic phenotype. This suppression is expressed in various ways. First, the double mutant shows less statistically significantly differentially expressed genes than either single mutant. Secondly, the genes that are common to the *egl-9* and *egl-9*;*hif-1* transcriptomes show decreased expression in the *egl-9*;*hif-1* mutant than they do in *egl-9* on average (see Fig. [fig:egl9epistasis]). The slope coefficient for the line of best fit is . An interpretation for this slope value is that *hif-1* suppresses *at least* 85% of the *egl-9* phenotype. Meanwhile, the genes that are common to *hif-1* and *egl-9*;*hif-1* show no change in expression on average between these two mutants, which shows that *egl-9* and *hif-1* are acting in the same pathway.

Because of the feedback between *hif-1* and *egl-9*, we expected a small subset of genes to be differentially expressed in every hypoxia pathway mutant. Therefore, we searched for genes that were differentially expressed in all our hypoxia mutants (except the *hif-1*;*egl-9* mutant because it has the least number of differentially expressed genes), reasoning that these genes should constitute an extremely high-quality picture of the hypoxia response, and should filter out other pathways. We identified 53 genes that satisfied these conditions, of which 10 genes were up-regulated in every mutant, and 13 genes were down-regulated. These genes constitute a core response around the circuit in question, and their behavior should reflect the genetic relationships in our system the best. Although we performed the regressions using all the overlapped genes between the single and double mutants, when we plotted only these high-quality genes, we can see that they show beautiful agreement with the global regressions (see <www.wormlabcaltech.github.io/mprsq> for all interactive graphics).

## Transcriptomic decorrelation can be used to infer functional distance

We were interested in figuring out whether RNA-Seq could be used to identify functional interactions within a genetic pathway. Although there is no *a priori* reason why global gene expression should reflect functional interactions, the strength of the unweighted correlations between genes in the hypoxia pathway made us wonder how much information can be extracted from this dataset. Single genes are often regulated by multiple independent sources. The connection between two nodes can in theory be characterized by the strength of the edges connecting them (the thickness of the edge); the fraction of sources that regulate both nodes (the fraction of common inputs); and the fraction of genes that are regulated by both nodes (the fraction of common outputs). In other words we expected that expression profiles associated with a pathway would respond quantitatively to quantitative changes in activity of the pathway. Targeting a pathway at multiple points would lead to expression profile divergence as we compare nodes that are separated by more degrees of freedom, reflecting the flux in information between them.

![ Theoretically, transcriptomes can be used to order genes in a pathway under certain assumptions. Arrows in the diagrams above are intended to show the direction of flow, and do not indicate valence. A A linear pathway in which rhy-1 is the only gene controlling egl-9, which in turn controls hif-1 does not contain transcriptomes with enough information to infer the order between genes. B On the other hand, if rhy-1 and egl-9 have transcriptomic effects that are separable from hif-1, then the rhy-1 transcriptome should contain contributions from egl-9, hif-1 and egl-9- and hif-1-independent pathways. This pathway contains enough information to infer order. C If a pathway is branched in both upstream and downstream directions, observed transcriptomes will show even faster decorrelation. Nodes that are separated by many edges may begin to behave almost independently of each other with marginal transcriptomic overlap or correlation, reflecting the weak control distant nodes exert on each other. D The hypoxia pathway can be ordered according to functional distance. The rapid decay in correlation is probably due to a mixture of upstream and downstream branching that happens along this pathway. ](data:application/pdf;base64,)

Theoretically, transcriptomes can be used to order genes in a pathway under certain assumptions. Arrows in the diagrams above are intended to show the direction of flow, and do not indicate valence. **A** A linear pathway in which *rhy-1* is the only gene controlling *egl-9*, which in turn controls *hif-1* does not contain transcriptomes with enough information to infer the order between genes. **B** On the other hand, if *rhy-1* and *egl-9* have transcriptomic effects that are separable from *hif-1*, then the *rhy-1* transcriptome should contain contributions from *egl-9*, *hif-1* and *egl-9*- and *hif-1*-independent pathways. This pathway contains enough information to infer order. **C** If a pathway is branched in both upstream and downstream directions, observed transcriptomes will show even faster decorrelation. Nodes that are separated by many edges may begin to behave almost independently of each other with marginal transcriptomic overlap or correlation, reflecting the weak control distant nodes exert on each other. **D** The hypoxia pathway can be ordered according to functional distance. The rapid decay in correlation is probably due to a mixture of upstream and downstream branching that happens along this pathway.

We investigated the possibility that transcriptomic signals do in fact contain relevant information about the degrees of separation by weighting the robust bayesian regression of each pair of genes by . We plotted the weighted correlation of each gene pair, ordered by increasing functional distance (see Fig. [fig:decorrelation]). In every case, we see that the weighted correlation decreases monotonically due mainly, but not exclusively, to decreasing . We believe that this result is not due to random noise or insufficiently deep sequencing. Instead, we propose a framework in which every gene is regulated by multiple different molecular species, which induces progressive decorrelation. This decorrelation in turn has two consequences. First, decorrelation within a pathway implies that two nodes may be almost independent of each other if the functional distance between them is large. Second, it may be possible to use decorrelation dynamics to infer gene order in a pathway, as we have done with the hypoxia pathway[[1]](#footnote-36).

## The circuit topology of the hypoxia pathway explains patterns in the data

We noticed that while some of the rank-plots contained a clear positive correlation (see Fig. [fig:genetic\_interactions]), some of the other rank-plots showed a discernible cross-pattern (see Fig. [fig:xpattern]). In particular, this cross-pattern emerged between *vhl-1* and *rhy-1* or between *vhl-1* and *egl-9*, even though genetically *vhl-1*, *rhy-1* and *egl-9* are all inhibitors of *hif-1*. We reasoned that it could be possible that these cross-patterns reflected multiple interaction modes between genes Therefore, we hypothesized that patterns in the rank-plots contained valuable information for decoding more interactions in our circuit.

![ Top: A feedback loop can generate transcriptomes that are both correlated and anti-correlated. Bottom: hif-1 transcriptome correlated to the rhy-1 transcriptome. Green large points are inliers to the first regression. Red small points are outliers to the first regression. Only the red small points were used for the secondary regression. Blue lines are representative samples of the primary bootstrapped regression lines. Orange lines are representative samples of the secondary bootstrapped regression lines. ](data:application/pdf;base64,)

**Top**: A feedback loop can generate transcriptomes that are both correlated and anti-correlated. **Bottom**: *hif-1* transcriptome correlated to the *rhy-1* transcriptome. Green large points are inliers to the first regression. Red small points are outliers to the first regression. Only the red small points were used for the secondary regression. Blue lines are representative samples of the primary bootstrapped regression lines. Orange lines are representative samples of the secondary bootstrapped regression lines.

If the logic above is correct, then it should be possible to decouple transcriptomes in a logically consistent way. Currently, transcriptomes are decoupled via subtractive logic. In other words, to identify the *rhy-1*-specific transcriptome (the effects of *rhy-1* not dependent on *egl-9*), subtractive logic might suggest to find the overlap between the two transcriptomes. The genes that are differentially expressed but are not in the overlap would then be considered *rhy-1*-specific transcriptomes. Such a gene set would consider of almost 700 genes. However, this approach suffers from a number of drawbacks, principally that it does not take into account the relationship between the two genes in question. Moreover, these genes have no testable properties: i.e., a gene might not be in the overlap because it was not identified due to chance in one of the two transcriptomes. In aggreggate, there is no pattern that is present in these genes that can be used to identify them beyond overlapping the two transcriptomes.

*rhy-1* and *egl-9* share a well-defined relationship. *rhy-1* inhibits *cysl-1*, which in turn inhibits *egl-9* (Ma et al. 2012). Therefore, loss of *rhy-1* leads to inactivation of *egl-9*, which leads to increase in the cellular levels of HIF-1. HIF-1 in turn causes the mRNA levels of *rhy-1* and *egl-9* to increase, as they are involved in the *hif-1*-dependent hypoxia response. However, since *rhy-1* has been mutated, the observed transcriptome is *rhy-1* null; *egl-9* null; HIF-1 on. The situation is similar for a knockout of *egl-9*, except that *rhy-1* is not inactive, and therefore the observed transcriptome is the result of RHY-1 up; *egl-9* null; and HIF-1 on. From this pattern, we conclude that the *egl-9* and *rhy-1* transcriptomes should exhibit a cross-pattern: The positive arm of the cross is the result of the *egl-9* null; HIF-1 on dynamics; and the negative arm reflects the different direction of RHY-1 activity between transcriptomes. However, no negative arm is visible (with the exception of two outliers, which are annotated as pseudogenes in WormBase). Therefore, it is likely that a large portion of all the transcriptomic effects of RHY-1 in this dataset are downstream of *egl-9*.

Next, we wanted to know whether our dataset was able to capture *egl-9* *hif-1*-independent transcriptomic effects. We have observed that deletion of *hif-1* leads to a modest increase in the transcription of *rhy-1*, from which we concluded that EGL-9 would be more active in the *hif-1* mutant than in the wild-type. Therefore, we searched for genes that were regulated in opposite manner between the *hif-1* and *hif-1*;*egl-9* mutants, and that were regulated in the same direction between the *hif-1*;*egl-9* and *egl-9* (or *rhy-1*) mutants. We were only able to find a single gene, *clec-88*, which was down-regulated in *hif-1* mutants, but upregulated in every other mutant we studied. Although it may be the case that *egl-9* does not have a *hif-1*-independent transcriptomic phenotype, it is also possible that the change in HIF-1 dosage between a wild-type normoxic animal and a *hif-1* mutant is not sufficient to alter the activity of EGL-9 to a consistently detectable level given our read-depth. We leveraged this genetic logic to identify a main hypoxia response induced by removing inhibition on *hif-1* (260 genes). Although the hypoxic response is likely to involve between five and ten times more genes, this is a conservative estimate that minimizes false negative results, since these changes were identified in four independent genotypes with three replicates each. We also identified a *vhl-1*-specific response, resulting in 36 genes. We searched for candidates directly regulated by *hif-1*. Initially, we generated this list using the most stringent pattern matching, but this revealed only 2 genes (*R08E5.3* and *nit-1*). A relaxed set of conditions (target genes should go up in all mutants that induce HIF-1, and should not be up in *hif-1* mutants) identified 120 candidate genes.

### Enrichment analysis of the hypoxia response

In order to validate that our transcriptomes were correct, and to understand how functionalities may vary between them, we subjected each decoupled response to enrichment analysis using the WormBase Enrichment Suite (Angeles-Albores, N. Lee, et al. 2016).

![ GEA of genes associated with the main hypoxia response. A number of terms reflecting catabolism and bioenergetics are enriched. ](data:application/pdf;base64,)

GEA of genes associated with the main hypoxia response. A number of terms reflecting catabolism and bioenergetics are enriched.

Gene enrichment analysis (GEA) showed that the terms ‘oxoacid metabolic process’ (, 3.4 fold-change, 19 genes), ‘iron ion binding’ (, 5.5 fold-change, 10 genes), and ‘immune system process’ (, 3.4 fold-change, 17 genes) were enriched with the lowest q-values. GEA also showed enrichment of terms including ‘electron carrier activity’ (, 4.8 fold-change, 5 genes), ‘mitochondrion’ (, 2.5 fold-change, 20 genes) and ‘respiratory chain’ (, 4.6 fold-change, 4 genes) (see Fig. [fig:hyp\_gea]). Indeed, *hif-1* has been implicated in all of these biological and molecular functions (Luhachack et al. 2012; Ackerman and Gems 2012; Romney et al. 2011; Semenza 2011). Phenotype Enrichment Analysis (PEA) revealed that this gene list was enriched in two phenotypes: ‘oxygen response variant’ (, 5.8 fold-change, 7 genes) and ‘pleiotropic defects severe early embryo’ (, 4.4 fold-change, 9 genes). The overrepresented terms from PEA and GEA are biologically directly connected to the process we are studying, which suggests that we have correctly identified the main hypoxic response. As a final test to guarantee the quality of our data, we selected a set of 21 known reporters from the literature and asked whether these reporters were present in our list. We found known reporters, which constitutes a statistically significant result (). The small number of reporters found in this list probably reflects the conservative nature of our estimates. We also analyzed the list of predicted *hif-1* direct targets. Phenotype Enrichment Analysis revealed that this list was significantly enriched in ‘oxygen response variant’ (, 12.3 fold-change, 4 genes) and Tissue Enrichment Analysis (TEA) showed enrichment of the ‘coelomic system’ (, 2.7 fold-change, 16 genes). The *vhl-1*, *hif-1*-independent specific transcriptome was also submitted for enrichment analysis but no terms were significantly enriched.

## Identification of non-classical epistatic interactions

*hif-1* has traditionally been viewed as existing in a genetic OFF state under normoxic conditions. However, our dataset indicates that 481 genes show altered expression when it is removed in normoxic conditions. Moreover, we observed positive genome-wide expression correlations between *hif-1* expression levels and *egl-9*, *vhl-1* and *rhy-1* expression levels in spite of the negative regulatory relationships between these genes and *hif-1*. Such positive relationships could indicate a different relationship between these genes than has previously been reported. We wanted to explore whether these genome-wide positive correlations were substantiated by epistatic analyses.

To perform epistatic analyses, we first identified genes that exhibited violations of the canonical genetic model of the hypoxia pathway. To this end, we searched for genes that exhibited different behaviors between the *egl-9* and the *vhl-1* mutants, or between the *rhy-1* and *vhl-1* (we assume that all results from the *rhy-1* transcriptome reflect a complete loss of *egl-9* activity). We found 27 that satisfied this condition (see Fig. [fig:hif1oh]). Additionally, many of these genes exhibited new kinds of epistasis. Namely, *egl-9* was epistatic to *vhl-1*. Identification of a set of genes that have a consistent set of relationships with between themselves suggests that we have identified a new aspect of the hypoxia pathway.

![ Genes that have altered differential expression between egl-9 and vhl-1 also often exhibit the same epistatic patterns between egl-9 and hif-1, and egl-9 and vhl-1, which suggests they are the result of the same biological effect. ](data:application/pdf;base64,)

Genes that have altered differential expression between *egl-9* and *vhl-1* also often exhibit the same epistatic patterns between *egl-9* and *hif-1*, and *egl-9* and *vhl-1*, which suggests they are the result of the same biological effect.

In particular, we focused on three genes, *nlp-31*, *ftn-1* and *ftn-2*, which epistasis patterns that we felt reflected the population well. As a sanity check, we reviewed the literature, and found that *ftn-1* and *ftn-2* are both described in the literature as genes that are responsive to mutations in the hypoxia pathway. Moreover, these genes have been previously described to have aberrant behaviors previously (Ackerman and Gems 2012; Romney et al. 2011), specifically documenting the opposite effects of *egl-9* and *vhl-1*. Probably as a reflection of the oddity of these results, these studies showed that loss of *vhl-1* suppresses *ftn-1* and *ftn-2* using both RNAi and alleles, which allays concerns of strain-specific interference. Moreover, one of these studies showed that *vhl-1*;*hif-1* is epistatic to *hif-1* (Ackerman and Gems 2012), and that loss of *hif-1* is associated with increased expression of *ftn-1* and *ftn-2*. We observe that *hif-1* is epistatic to *egl-9*;*hif-1*, and that *egl-9* and *hif-1* promote *ftn-1* and *ftn-2* expression. This further validates the quality of our RNA-seq data and the analysis, and highlights the power of RNA-seq to identify novel interactions.

A qualitative epistatic analysis of *ftn-1* and *ftn-2* reveals that *egl-9* is epistatic to *hif-1*; that *vhl-1* has opposite effects to *egl-9*; and *vhl-1* is epistatic to *egl-9*. Epistatic analysis of *nlp-31* reveals similar relationships. *nlp-31* expression is decreased by loss of *hif-1*, and promoted by *egl-9*. However, *egl-9* is epistatic to *hif-1*. Like *ftn-1* and *ftn-2*, *vhl-1* has the opposite effect to *egl-9*, but is also epistatic to *egl-9*.

## Genome-wide effects of *hif-1*

The high quality of this dataset also provides us directly with a high-level overview of the transcriptional responses that lead to physiologic and metabolic changes in hypoxia. We wanted to better understand the transcriptional changes associated with bioenergetic pathways in *C. elegans*. To this end, we extracted from WormBase all genes associated with the tricarboxylic acid (TCA) cycle, the electron transport chain (ETC) and with the *C. elegans* energy reserve (glycogen metabolism, fatty acid metabolism, etc…). Previous research has described the effects of mitochondrial dysfunction in eliciting the hypoxia response (Lee, Hwang, and Kenyon 2010), but transcriptional feedback from *hif-1* into bioenergetic pathways has not been well described in *C. elegans*, although it has been extensively described in other organisms (see, for example (Semenza et al. 1994; Semenza 2012)).

### Bio-energetic pathways

Our data shows that most of the enzymes involved in the TCA cycle and in the ETC are down-regulated when HIF-1 is induced in agreement with the previous literature (Semenza 2012). However, *fum-1* and the mitochondrial complex II stood out as notable exceptions to the trend, as they were up-regulated in every single genotype that causes deployment of the hypoxia response. *fum-1* catalyses the reaction of fumarate into malate, and complex II catalyses the reaction of succinate into fumarate. Complex II has been identified as a source of reserve respiratory capacity in neonatal rat cardiomyocytes previously (Pfleger, He, and Abdellatif 2015). We found two energy reserve genes that were down-regulated by HIF-1. *aagr-1* and *aagr-2*, which are predicted to function in glycogen catabolism (Sikora et al. 2010) were both down-regulated in all the relevant mutants. Three distinct genes involved in energy reserve were up-regulated. These genes were *ogt-1*, an O-linled GlcNac Transferase; *T04A8.7*, an ortholog of human glucosidase acid beta (GBA); and *T22F3.3*, an ortholog of human glycogen phosphorylase isozymes.

### Protein synthesis and degradation

*hif-1* is also known to inhibit protein synthesis and translation in varied ways. For example, *hif-1* is known to control the translational machinery indirectly via inhibition of mTOR (Brugarolas et al. 2004). However, most reported effects of *hif-1* on the translation machinery are posttranslational, and no reports to date show decreases in transcription of the ribosomal machinery in *C. elegans*. We used the WormBase Enrichment Suite Gene Ontology dictionary **???** to extract 143 genes annotated as ‘structural constituents of the ribosome’ and we queried whether they were differentially expressed in our mutants. *egl-9*, *vhl-1*, *rhy-1* and *egl-9*;*vhl-1* mutants showed differential expression of 91 distinct ribosomal constituents (not all constituents were detected in all genotypes). For every one of these genotypes, these genes were always down-regulated. In contrast, the *hif-1* mutant showed up-regulation of a single ribosomal constituent.

Next, we wanted to know whether *hif-1* has any transcriptional effects on the proteasomal constituents, because no such effects of *hif-1* on the proteasome have been reported in *C. elegans*. Out of 40 WormBase annotated proteasomal constituents, we found 31 constituents that were differentially expressed in at least one of the four genotypes that induce a hypoxic response. Every gene we found was down-regulated in at least two out of the four genotypes we studied, although in each case the down-regulation was minor. It is impossible to distinguish whether these animals exhibit a decrease in proteasome expression is due to a lower requirement for degradation in these animals due to constitutively depressed translation rates or whether the decrease in expression is a direct result of HIF-1 stabilization.

# Discussion

## The *C. elegans* hypoxia pathway can be reconstructed entirely from RNA-seq data

We have presented the first genetic pathway reconstruction in a multicellular organism using whole-organism RNA-seq to measure transcriptomic phenotypes. We were able to reconstruct first-order and second-order interactions. We were able to infer order of action (*rhy-1* activates *egl-9*, *egl-9* and *vhl-1* inhibit *hif-1*), and we were able to infer from genome-wide epistatic measurements that *egl-9* exerts *vhl-1*-dependent and independent inhibition on *hif-1*.

In addition to reconstructing the pathway, our dataset afforded us the opportunity to observe a wide variety of physiologic changes that occur when the *hif-1*-dependent hypoxia response is activated. In particular, we observed down-regulation of most components of the TCA cycle and the mitochondrial electron transport chain. As an exception, *fum-1* and the mitochondrial complex II, which are involved in fumarate metabolism within these pathways were up-regulated. The mitochondrial complex II catalyses the reaction of succinate into fumarate. Complex II is known to be important for hypoxic survival in rat cardiomyocyte cells (Pfleger, He, and Abdellatif 2015). Complex II may play a similar role in *C. elegans*. The product of complex II activity is fumarate. In mouse embryonic fibroblasts, fumarate has been shown to antagonize HIF-1 prolyl hydroxylase domain (PHD) enzymes, which are orthologs of *egl-9*. Upregulation of complex II by HIF-1 during hypoxia may therefore result in increased intracellular levels of fumarate, which in turn could lead to artificially high levels of *hif-1* (if the inhibitory role of fumarate is conserved in *C. elegans*) even after hypoxic conditions have vanished in the absence of concurrent metabolic changes.

Under this framework, the up-regulation of *fum-1* agrees with intuition. By up-regulating *fum-1*, we speculate that *C. elegans* may be capable of harnessing reserve respiratory capacity via complex II, while rapidly degrading the excess fumarate that is generated. Degrading fumarate rapidly may allow *C. elegans* to maintain plasticity in the hypoxia pathway, keeping the pathway sensitive to oxygen levels.

## Non-classical epistasis in the hypoxia pathway

The observation of almost 30 genes that exhibit a specific pattern of non-classical epistasis reveals new aspects of the pathway. Some of these non-classical epistases had been observed previously, but no satisfactory mechanism has been proposed to explain this biology. (Romney et al. 2011) and (Ackerman and Gems 2012) suggest that *hif-1* integrates information on iron concentration in the cell to bind to the *ftn-1* promoter, but could not definitively establish a mechanism. In particular, it is unclear why deletion of *hif-1* induces *ftn-1* expression, but deletion of *egl-9* also causes induction of *ftn-1* expression, whereas *vhl-1* removes this inhibition. Moreover, (Luhachack et al. 2012) have previously reported that certain genes important for the *C. elegans* immune response against pathogens reflect similar expression patterns. Their interpretation was that *swan-1*, a binding partner to *egl-9* (Shao et al. 2010), is important for modulating HIF-1 activity somehow. The lack of a conclusive double mutant analysis in this work means the role of *swan-1* in modulation of HIF-1 activity remains to be demonstrated. At any rate, mechanisms that call for additional transcriptional modulators become more unlikely given our data the large number of genes with different biological functions that exhibit the same pattern.

![ A toy model showing that an interpretation where HIF-1-hydroxyl is biochemically active can potentially explain how genes that exhibit non-canonical epistasis are regulated. ](data:application/pdf;base64,)

A toy model showing that an interpretation where HIF-1-hydroxyl is biochemically active can potentially explain how genes that exhibit non-canonical epistasis are regulated.

One way to resolve this problem without invoking additional genes is to model HIF-1 as a protein with both activating and inhibiting states. In fact, HIF-1 already exists in two states in *C. elegans*: unmodified HIF-1 and HIF-1-hydroxyl. Under this model, HIF-1-hydroxyl would inhibit gene expression, whereas HIF-1 drives it. *vhl-1* stabilizes HIF-1-hydroxyl, which will cause inhibition of genes to which both forms of the protein bind; whereas *egl-9* selectively removes all HIF-1-hydroxyl, indirectly stimulating accumulation of HIF-1 and promoting gene activity. Whether deletion of *hif-1* is activating or inhibiting will depend on the relative contributions of each protein activity under normoxia (see Fig. [fig:hif1oh\_table]).

The possibility that HIF-1-hydroxyl has a function has not been previously considered in the existing literature, although experts have wondered about the possibility that HIF-1-hydroxyl may have transcriptional effects independent of HIF-1 (William Kaelin, pers. comm.). Here, we draw multiple, varied lines of circumstantial evidence to suggest that HIF-1 hydroxylation plays a role in the functionality of the hypoxia pathway. First, HIF-1-hydroxyl is challenging to study genetically because no mimetic mutations are available with which to study the pure hydroxylated HIF-1 species. Moreover, mutations in the Von-Hippel Landau gene stabilize the hydroxyl species, but also increase the quantity of HIF-1 by mass action. Since *hif-1* exists at low levels in cells under normoxic conditions, total HIF-1 protein (unmodified HIF-1 plus HIF-1-hydroxyl) is often tacitly assumed to be vanishingly rare.

Our data shows that there are hundreds of genes that change expression in response to loss of *hif-1* under normoxic conditions. This establishes that there is sufficient total HIF-1 protein to be biologically active. Previous literature showing substantial changes in mRNA expression of *ftn-1* support our claim that under normoxia *hif-1* is biologically relevant. Moreover, our analysis of the hypoxia pathway using transcriptomic phenotypes reveals that *hif-1* shares main positive correlations with *egl-9*, *rhy-1* and *vhl-1*, and that each of these genes also shows a set of genes negative rank-ordered expression correlations. These cross-patterns between all controllers of *hif-1* and *hif-1* can be most easily explained if HIF-1-hydroxyl is biologically active.

An additional argument in favor of the activity of HIF-1-hydroxyl is a homeostatic argument. At any point in time, the cell must measure the levels of multiple small molecules at once. Strictly speaking, the *hif-1*-dependent hypoxia response integrates information from O, -ketoglutarate (2-oxoglutarate) and iron concentrations in the cell. One way to encode this information is by encoding it only in the effective hydroxylation rate of HIF-1 by EGL-9. Then the dynamics in this system will evolve exclusively as a result of the total amount of HIF-1 in the cell. Such a system can be sensitive to fluctuations in the absolute concentration of HIF-1 (Goentoro et al. 2009). In the case of severe hypoxia, when the levels of HIF-1 are expected to rise enormously within the cell, simple information integration via *egl-9* would probably be sufficient encoding for a subset of protective genes.

For yet other set of genes that must change expression in response to the hypoxia pathway, it may not make as much sense to integrate metabolite information exclusively via EGL-9-dependent hydroxylation of HIF-1. In particular, genes that may increase survival in mild hypoxia may benefit from homeostatic regulation that is not susceptible to transient changes in protein copy number. Likewise, genes that are involved in iron or -ketoglutarate metabolism (such as *ftn-1*) may benefit from being able to sense, accurately, small and consistent deviations from basal concentrations of these metabolites. For these genes, the information may be better encoded by using HIF-1 and HIF-1-hydroxyl as an activator/repressor pair. Such paradoxical circuits are known to possess distinct advantages for controlling output in a manner that is robust to transient fluctuations in the levels of their components (Hart et al. 2012; Hart and Alon 2013).

Our RNA-seq data suggests that one of the targets that *hif-1* may target paradoxically is *rhy-1*. Although *rhy-1* does not exhibit non-classical epistasis, genotypes that included a loss-of-function *hif-1* caused increased *rhy-1* activity. We speculate that if *rhy-1* is controlled by both HIF-1 and HIF-1-hydroxyl, then this might mean that *hif-1* regulates the expression of its pathway (and therefore itself) in a manner that is robust to total HIF-1 levels.

## Looking forward

We have demonstrated the first complete reconstruction of a genetic pathway using whole-organism RNA-seq in a complex multicellular organism. Future work must rigorously demonstrate that macroscopically derived genetic rules of interaction hold genome-wide, and if not, how genetic rules deviate from expectations.

Ackerman, Daniel, and David Gems. 2012. “Insulin/IGF-1 and hypoxia signaling act in concert to regulate iron homeostasis in Caenorhabditis elegans.” *PLoS Genetics* 8 (3). doi:[10.1371/journal.pgen.1002498](https://doi.org/10.1371/journal.pgen.1002498).

Adamson, Britt, Thomas M. Norman, Marco Jost, Min Y. Cho, James K. Nuñez, Yuwen Chen, Jacqueline E. Villalta, et al. 2016. “A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response.” *Cell* 167 (7): 1867–1882.e21. doi:[10.1016/j.cell.2016.11.048](https://doi.org/10.1016/j.cell.2016.11.048).

Angeles-Albores, David, Daniel H W Leighton, Tiffany Tsou, Tiffany H Khaw, Igor Antoshechkin, and Paul Sternberg. 2016. “Transcriptomic Description of an Endogenous Female State in C. elegans.” *BioRxiv*, October. <http://biorxiv.org/content/early/2016/10/27/083113.abstract>.

Angeles-Albores, David, Raymond Y. N. Lee, Juancarlos Chan, and Paul W. Sternberg. 2016. “Tissue enrichment analysis for C. elegans genomics.” *BMC Bioinformatics* 17 (1). BMC Bioinformatics: 366. doi:[10.1186/s12859-016-1229-9](https://doi.org/10.1186/s12859-016-1229-9).

Bellier, Audrey, Chang Shi Chen, Cheng Yuan Kao, Hediye N. Cinar, and Raffi V. Aroian. 2009. “Hypoxia and the hypoxic response pathway protect against pore-forming toxins in C. elegans.” *PLoS Pathogens* 5 (12). doi:[10.1371/journal.ppat.1000689](https://doi.org/10.1371/journal.ppat.1000689).

Bishop, Tammie, Kah Weng Lau, Andrew C. R. Epstein, Stuart K. Kim, Min Jiang, Delia O’Rourke, Christopher W. Pugh, et al. 2004. “Genetic Analysis of Pathways Regulated by the von Hippel-Lindau Tumor Suppressor in Caenorhabditis elegans.” *PLoS Biology* 2 (10). doi:[10.1371/journal.pbio.0020289](https://doi.org/10.1371/journal.pbio.0020289).

Bray, Nicolas L., Harold J Pimentel, Páll Melsted, and Lior Pachter. 2016. “Near-optimal probabilistic RNA-seq quantification.” *Nature Biotechnology* 34 (5): 525–7. doi:[10.1038/nbt.3519](https://doi.org/10.1038/nbt.3519).

Brugarolas, James, Kui Lei, Rebecca L Hurley, Brendan D Manning, Jan H Reiling, Ernst Hafen, Lee A Witters, Leif W Ellisen, William G Kaelin, and William G Kaelin Jr. 2004. “Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1 / TSC2 tumor suppressor complex.” *Genes and Development* 18 (23): 1–12. doi:[10.1101/gad.1256804.(mTOR)](https://doi.org/10.1101/gad.1256804.(mTOR)).

Dixit, Atray, Oren Parnas, Biyu Li, Jenny Chen, Charles P. Fulco, Livnat Jerby-Arnon, Nemanja D. Marjanovic, et al. 2016. “Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens.” *Cell* 167 (7). Elsevier Inc.: 1853–1866.e17. doi:[10.1016/j.cell.2016.11.038](https://doi.org/10.1016/j.cell.2016.11.038).

Epstein, Andrew C. R., Jonathan M. Gleadle, Luke A. McNeill, Kirsty S. Hewitson, John O’Rourke, David R. Mole, Mridul Mukherji, et al. 2001. “C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation.” *Cell* 107 (1): 43–54. doi:[10.1016/S0092-8674(01)00507-4](https://doi.org/10.1016/S0092-8674(01)00507-4).

Goentoro, Lea, Oren Shoval, Marc W. Kirschner, and Uri Alon. 2009. “The Incoherent Feedforward Loop Can Provide Fold-Change Detection in Gene Regulation.” *Molecular Cell* 36 (5). Elsevier Ltd: 894–99. doi:[10.1016/j.molcel.2009.11.018](https://doi.org/10.1016/j.molcel.2009.11.018).

Hart, Yuval, and Uri Alon. 2013. “The Utility of Paradoxical Components in Biological Circuits.” *Molecular Cell* 49 (2). Elsevier: 213–21. doi:[10.1016/j.molcel.2013.01.004](https://doi.org/10.1016/j.molcel.2013.01.004).

Hart, Yuval, Yaron E Antebi, Avraham E Mayo, Nir Friedman, and Uri Alon. 2012. “Design principles of cell circuits with paradoxical components.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (21): 8346–51. doi:[10.1073/pnas.1117475109](https://doi.org/10.1073/pnas.1117475109).

Huang, L. Eric, Zoltan Arany, David M. Livingston, and H. Franklin Bunn. 1996. “Activation of hypoxia-inducible transcription factor depends primarily upon redox-sensitive stabilization of its alpha subunit.” *Journal of Biological Chemistry* 271 (50): 32253–9. doi:[10.1074/jbc.271.50.32253](https://doi.org/10.1074/jbc.271.50.32253).

Hughes, Timothy R., Matthew J. Marton, Allan R. Jones, Christopher J. Roberts, Roland Stoughton, Christopher D. Armour, Holly A Bennett, et al. 2000. “Functional Discovery via a Compendium of Expression Profiles.” *Cell* 102 (1): 109–26. doi:[10.1016/S0092-8674(00)00015-5](https://doi.org/10.1016/S0092-8674(00)00015-5).

Jiang, B H, E Rue, G L Wang, R Roe, and G L Semenza. 1996. “Dimerization, DNA binding, and transactivation properties of hypoxia-inducible factor 1.” *The Journal of Biological Chemistry* 271 (30): 17771–8. doi:[10.1074/jbc.271.30.17771](https://doi.org/10.1074/jbc.271.30.17771).

Jiang, Huaqi, Rong Guo, and Jo Anne Powell-Coffman. 2001. “The Caenorhabditis elegans hif-1 gene encodes a bHLH-PAS protein that is required for adaptation to hypoxia.” *Proceedings of the National Academy of Sciences of the United States of America* 98 (14): 7916–21. doi:[10.1073/pnas.141234698](https://doi.org/10.1073/pnas.141234698).

Kaelin, William G., and Peter J. Ratcliffe. 2008. “Oxygen Sensing by Metazoans: The Central Role of the HIF Hydroxylase Pathway.” *Molecular Cell* 30 (4): 393–402. doi:[10.1016/j.molcel.2008.04.009](https://doi.org/10.1016/j.molcel.2008.04.009).

Lee, Seung Jae, Ara B. Hwang, and Cynthia Kenyon. 2010. “Inhibition of respiration extends C. elegans life span via reactive oxygen species that increase HIF-1 activity.” *Current Biology* 20 (23). Elsevier Ltd: 2131–6. doi:[10.1016/j.cub.2010.10.057](https://doi.org/10.1016/j.cub.2010.10.057).

Loenarz, Christoph, Mathew L Coleman, Anna Boleininger, Bernd Schierwater, Peter W H Holland, Peter J Ratcliffe, and Christopher J Schofield. 2011. “The hypoxia-inducible transcription factor pathway regulates oxygen sensing in the simplest animal, Trichoplax adhaerens.” *EMBO Reports* 12 (1). Nature Publishing Group: 63–70. doi:[10.1038/embor.2010.170](https://doi.org/10.1038/embor.2010.170).

Luhachack, Lyly G., Orane Visvikis, Amanda C. Wollenberg, Adam Lacy-Hulbert, Lynda M. Stuart, and Javier E. Irazoqui. 2012. “EGL-9 Controls C. elegans Host Defense Specificity through Prolyl Hydroxylation-Dependent and -Independent HIF-1 Pathways.” *PLoS Pathogens* 8 (7): 48. doi:[10.1371/journal.ppat.1002798](https://doi.org/10.1371/journal.ppat.1002798).

Ma, Dengke K., Roman Vozdek, Nikhil Bhatla, and H. Robert Horvitz. 2012. “CYSL-1 Interacts with the O 2-Sensing Hydroxylase EGL-9 to Promote H 2S-Modulated Hypoxia-Induced Behavioral Plasticity in C. elegans.” *Neuron* 73 (5). Elsevier Inc.: 925–40. doi:[10.1016/j.neuron.2011.12.037](https://doi.org/10.1016/j.neuron.2011.12.037).

Metzker, Michael L. 2010. “Sequencing technologies - the next generation.” *Nature Reviews. Genetics* 11 (1). Nature Publishing Group: 31–46. doi:[10.1038/nrg2626](https://doi.org/10.1038/nrg2626).

Mortazavi, Ali, Brian A Williams, Kenneth McCue, Lorian Schaeffer, and Barbara Wold. 2008. “Mapping and quantifying mammalian transcriptomes by RNA-Seq.” *Nature Methods* 5 (7): 621–28. doi:[10.1038/nmeth.1226](https://doi.org/10.1038/nmeth.1226).

Park, Eun Chan, Piya Ghose, Zhiyong Shao, Qi Ye, Lijun Kang, Xz Shawn Xu, Jo Anne Powell-Coffman, and Christopher Rongo. 2012. “Hypoxia regulates glutamate receptor trafficking through an HIF-independent mechanism.” *The EMBO Journal* 31 (6): 1618–9. doi:[10.1038/emboj.2012.44](https://doi.org/10.1038/emboj.2012.44).

Patro, Rob, Geet Duggal, Michael I. Love, Rafael A. Irizarry, and Carl Kingsford. 2016. “Salmon provides accurate, fast, and bias-aware transcript expression estimates using dual-phase inference.” *BioRxiv*, 021592. doi:[10.1101/021592](https://doi.org/10.1101/021592).

Patro, Rob, Stephen M. Mount, and Carl Kingsford. 2014. “Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms.” *Nature Biotechnology* 32 (5). Nature Publishing Group: 462–64. doi:[10.1038/nbt.2862](https://doi.org/10.1038/nbt.2862).

Pfleger, J, M He, and M Abdellatif. 2015. “Mitochondrial complex II is a source of the reserve respiratory capacity that is regulated by metabolic sensors and promotes cell survival.” *Cell Death and Disease* 6 (7): 1–14. doi:[10.1038/cddis.2015.202](https://doi.org/10.1038/cddis.2015.202).

Phillips, Patrick C. 2008. “Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems.” *Nat Rev Genet* 9 (11): 855–67. doi:[10.1038/nrg2452](https://doi.org/10.1038/nrg2452).

Pimentel, Harold J, Nicolas L. Bray, Suzette Puente, Páll Melsted, and Lior Pachter. 2016. “Differential analysis of RNA-Seq incorporating quantification uncertainty.” *BioRxiv*, 058164. doi:[10.1101/058164](https://doi.org/10.1101/058164).

Powell-Coffman, Jo Anne. 2010. “Hypoxia signaling and resistance in C. elegans.” *Trends in Endocrinology and Metabolism* 21 (7). Elsevier Ltd: 435–40. doi:[10.1016/j.tem.2010.02.006](https://doi.org/10.1016/j.tem.2010.02.006).

Powell-Coffman, Jo Anne, Christopher A. Bradfield, and William B. Wood. 1998. “Caenorhabditis elegans Orthologs of the Aryl Hydrocarbon Receptor and Its Heterodimerization Partner the Aryl Hydrocarbon Receptor Nuclear Translocator.” *Proceedings of the National Academy of Sciences* 95 (6): 2844–9. doi:[10.1073/pnas.95.6.2844](https://doi.org/10.1073/pnas.95.6.2844).

Romney, Steven Joshua, Ben S. Newman, Colin Thacker, and Elizabeth A. Leibold. 2011. “HIF-1 regulates iron homeostasis in caenorhabditis elegans by activation and inhibition of genes involved in iron uptake and storage.” *PLoS Genetics* 7 (12). doi:[10.1371/journal.pgen.1002394](https://doi.org/10.1371/journal.pgen.1002394).

Schwarz, Erich M., Mihoko Kato, and Paul W. Sternberg. 2012. “Functional transcriptomics of a migrating cell in Caenorhabditis elegans.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (40): 16246–51. doi:[10.1073/pnas.1203045109](https://doi.org/10.1073/pnas.1203045109).

Scimone, M. Lucila, Kellie M. Kravarik, Sylvain W. Lapan, and Peter W. Reddien. 2014. “Neoblast specialization in regeneration of the planarian schmidtea mediterranea.” *Stem Cell Reports* 3 (2). The Authors: 339–52. doi:[10.1016/j.stemcr.2014.06.001](https://doi.org/10.1016/j.stemcr.2014.06.001).

Semenza, G L, P H Roth, H M Fang, and G L Wang. 1994. “Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1.” *The Journal of Biological Chemistry* 269 (38): 23757–63. <http://www.ncbi.nlm.nih.gov/pubmed/8089148>.

Semenza, Gregg L. 2011. “Hypoxia-inducible factor 1: Regulator of mitochondrial metabolism and mediator of ischemic preconditioning.” *Biochimica et Biophysica Acta - Molecular Cell Research* 1813 (7). Elsevier B.V.: 1263–8. doi:[10.1016/j.bbamcr.2010.08.006](https://doi.org/10.1016/j.bbamcr.2010.08.006).

———. 2012. “Hypoxia-inducible factors in physiology and medicine.” *Cell* 148 (3). Elsevier Inc.: 399–408. doi:[10.1016/j.cell.2012.01.021](https://doi.org/10.1016/j.cell.2012.01.021).

Shalek, Alex K., Rahul Satija, Xian Adiconis, Rona S. Gertner, Jellert T. Gaublomme, Raktima Raychowdhury, Schraga Schwartz, et al. 2013. “Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells.” *Nature* 498 (7453). Nature Publishing Group: 236–40. doi:[10.1038/nature12172](https://doi.org/10.1038/nature12172).

Shao, Zhiyong, Yi Zhang, and Jo Anne Powell-Coffman. 2009. “Two Distinct Roles for EGL-9 in the Regulation of HIF-1-mediated gene expression in Caenorhabditis elegans.” *Genetics* 183 (3): 821–29. doi:[10.1534/genetics.109.107284](https://doi.org/10.1534/genetics.109.107284).

Shao, Zhiyong, Yi Zhang, Qi Ye, Jenifer Neeta Saldanha, and Jo Anne Powell-Coffman. 2010. “C. elegans swan-1 binds to egl-9 and regulates hif-1- mediated resistance to the bacterial pathogen pseudomonas aeruginosa pao1.” *PLoS Pathogens* 6 (8): 91–92. doi:[10.1371/journal.ppat.1001075](https://doi.org/10.1371/journal.ppat.1001075).

Shen, Chuan, Daniel Nettleton, Min Jiang, Stuart K. Kim, and Jo Anne Powell-Coffman. 2005. “Roles of the HIF-1 hypoxia-inducible factor during hypoxia response in Caenorhabditis elegans.” *Journal of Biological Chemistry* 280 (21): 20580–8. doi:[10.1074/jbc.M501894200](https://doi.org/10.1074/jbc.M501894200).

Shen, Chuan, Zhiyong Shao, and Jo Anne Powell-Coffman. 2006. “The Caenorhabditis elegans rhy-1 Gene Inhibits HIF-1 Hypoxia-Inducible Factor Activity in a Negative Feedback Loop That Does Not Include vhl-1.” *Genetics* 174 (3): 1205–14. doi:[10.1534/genetics.106.063594](https://doi.org/10.1534/genetics.106.063594).

Sikora, Jakub, Jana Uřinovská, Filip Majer, Helena Poupětová, Jitka Hlavatá, Marta Kostrouchová, Jana Ledvinová, and Martin Hřebíček. 2010. “Bioinformatic and biochemical studies point to AAGR-1 as the ortholog of human acid ??-glucosidase in Caenorhabditis elegans.” *Molecular and Cellular Biochemistry* 341 (1-2). Springer US: 51–63. doi:[10.1007/s11010-010-0436-3](https://doi.org/10.1007/s11010-010-0436-3).

Singer, Meromit, Chao Wang, Le Cong, Nemanja D. Marjanovic, Monika S. Kowalczyk, Huiyan Zhang, Jackson Nyman, et al. 2016. “A Distinct Gene Module for Dysfunction Uncoupled from Activation in Tumor-Infiltrating T Cells.” *Cell* 166 (6). Elsevier Inc.: 1500–1511.e9. doi:[10.1016/J.CELL.2016.08.052](https://doi.org/10.1016/J.CELL.2016.08.052).

Trapnell, Cole, David G. Hendrickson, Martin Sauvageau, Loyal Goff, John L. Rinn, and Lior Pachter. 2013. “Differential analysis of gene regulation at transcript resolution with RNA-seq.” *Nature Biotechnology* 31 (1). Nature Publishing Group: 46–53. doi:[10.1038/nbt.2450](https://doi.org/10.1038/nbt.2450).

Van Driessche, Nancy, Janez Demsar, Ezgi O. Booth, Paul Hill, Peter Juvan, Blaz Zupan, Adam Kuspa, and Gad Shaulsky. 2005. “Epistasis analysis with global transcriptional phenotypes.” *Nature Genetics* 37 (5): 471–77. doi:[10.1038/ng1545](https://doi.org/10.1038/ng1545).

Van Wolfswinkel, Josien C., Daniel E. Wagner, and Peter W. Reddien. 2014. “Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment.” *Cell Stem Cell* 15 (3). Elsevier Inc.: 326–39. doi:[10.1016/j.stem.2014.06.007](https://doi.org/10.1016/j.stem.2014.06.007).

Yeung, Ka Yee, Mario Medvedovic, and Roger E. Bumgarner. 2003. “Clustering gene-expression data with repeated measurements.” *Genome Biology* 4 (5): R34. doi:[10.1186/gb-2003-4-5-r34](https://doi.org/10.1186/gb-2003-4-5-r34).

1. An important question is whether a looped circuit like the hypoxia pathway can be ordered in the way we have ordered it in Fig. [fig:decorrelation] since a loop does not technically have a beginning. One explanation is that we studied the hypoxia pathway under normoxic conditions, and therefore the control of *hif-1* over *rhy-1* and *egl-9* is weak, effectively turning the looped pathway into a linear one. Probably, under hypoxic conditions the pathway would effectively be reversed. [↑](#footnote-ref-36)