pie-1 and its effect on other genes

Jia Le, Lim CID: 00865029

Department of Biology, Imperial College London, South Kensington Campus, London, U.K.

Submitted: 10 Feb 2017

Contents

1 Introduction

pie-1 (pharynx and intestine excess) is a maternal effect gene in *Caenorhabditis elegans*. As its name suggests, worms carrying *pie-1* mutations produce embryos with extra pharyngeal and intestinal cells (Mello et al., 1992). The protein that it encodes, PIE-1, is an embryonic transcriptional repressor that is essential in regulating germ cell fate during the early stages of *C.elegans* development. PIE-1 is a cysteine-X_{8-10}-cysteine-X_5-cysteine-X_3-histidine (CCCH)-type zinc finger protein that is enriched in germline blastomeres and maintains germ cell fate by inhibiting transcriptional elongation, hence blocking genetic mechanisms responsible for somatic development. PIE-1 may also induce translation of other maternal RNAs such as *nos-2*, which are required for the maintenance of germ cell fate (Ghosh and Seydoux, 2008; Gilbert, 2000; Tenenhaus et al., 2001).

As one of the key transcription factors in germline cells, PIE-1 is likely to interact with many other maternal proteins and mRNA and directly or indirectly repress or activate the transcription of embryonic genes. Nonetheless, most of these interactions has not yet been studied (Gilbert, 2000). Hence, in this practical, likely candidates that interact with PIE-1 up to the first larval stage were identified. Thereafter, such interactions and experimental data that support them, if available, are further clarified in the report.

2 Methods and Results

71 possible candidate genes were identified using the list of gene interactions with *pie-1* available on the WormBase. Of which, 7 were promoter regions and were subsequently removed from the list. The expression patterns of the remaining 64 genes were isolated from the dataset obtained from a previous study by Levin et al. (2012). In Levin et al. (2012), the relative expression of 19051 genes in *C. elegans* at different time intervals between the 4-cell stage to the first larval stage were estimated using their respective mRNA levels in embryos collected at the corresponding developmental stage (Levin et al., 2012).

2.1 *pie-1* expression exponentially decreases during development.

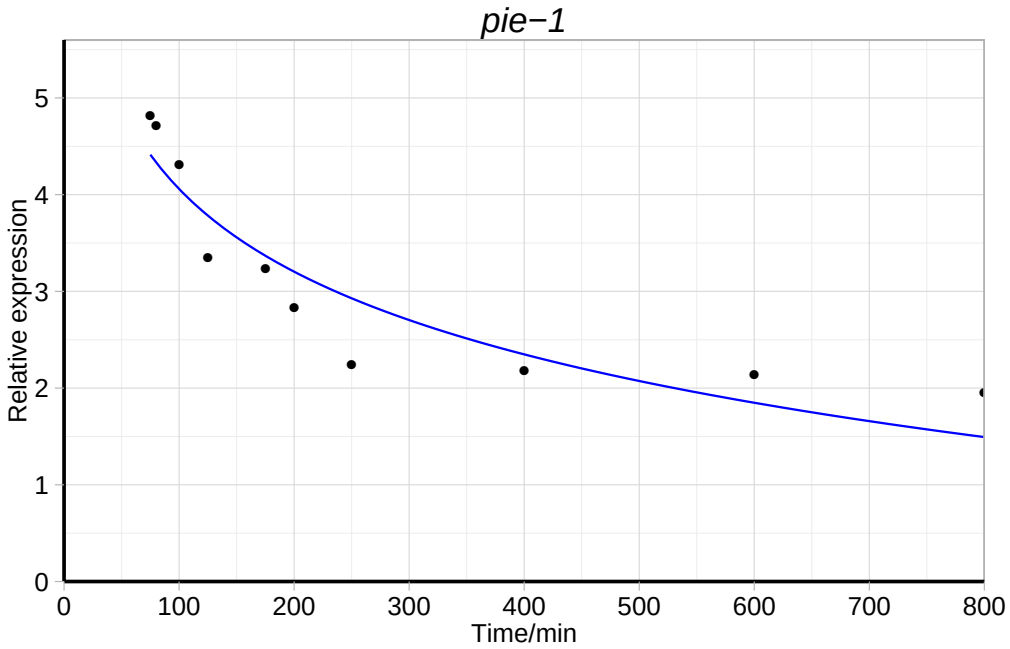

Figure 1: **Relative expression of *pie-1* during early *C. elegans* development.** Time on the x-axis indicates the amount of time since the 4-cell stage.

As expected, *pie-1* expression level decreases as *C. elegans* embryos proceed to the larval stage (Figure 1). An exponential decay model best fits the relative expression data of *pie-1* (adj. $r^2 = 0.84$, intercept = 9.74±0.94 , slope = -1.23±0.175, both $p<0.001$, res. df = 8).

2.2 Identifying the genes of interest

2.2.1 Heatmap and Hierarchal Cluster Analysis

A heatmap showing the predicted clusters of the 64 candidate genes was generated using Ward's method (ward.D2) (Figure 2).

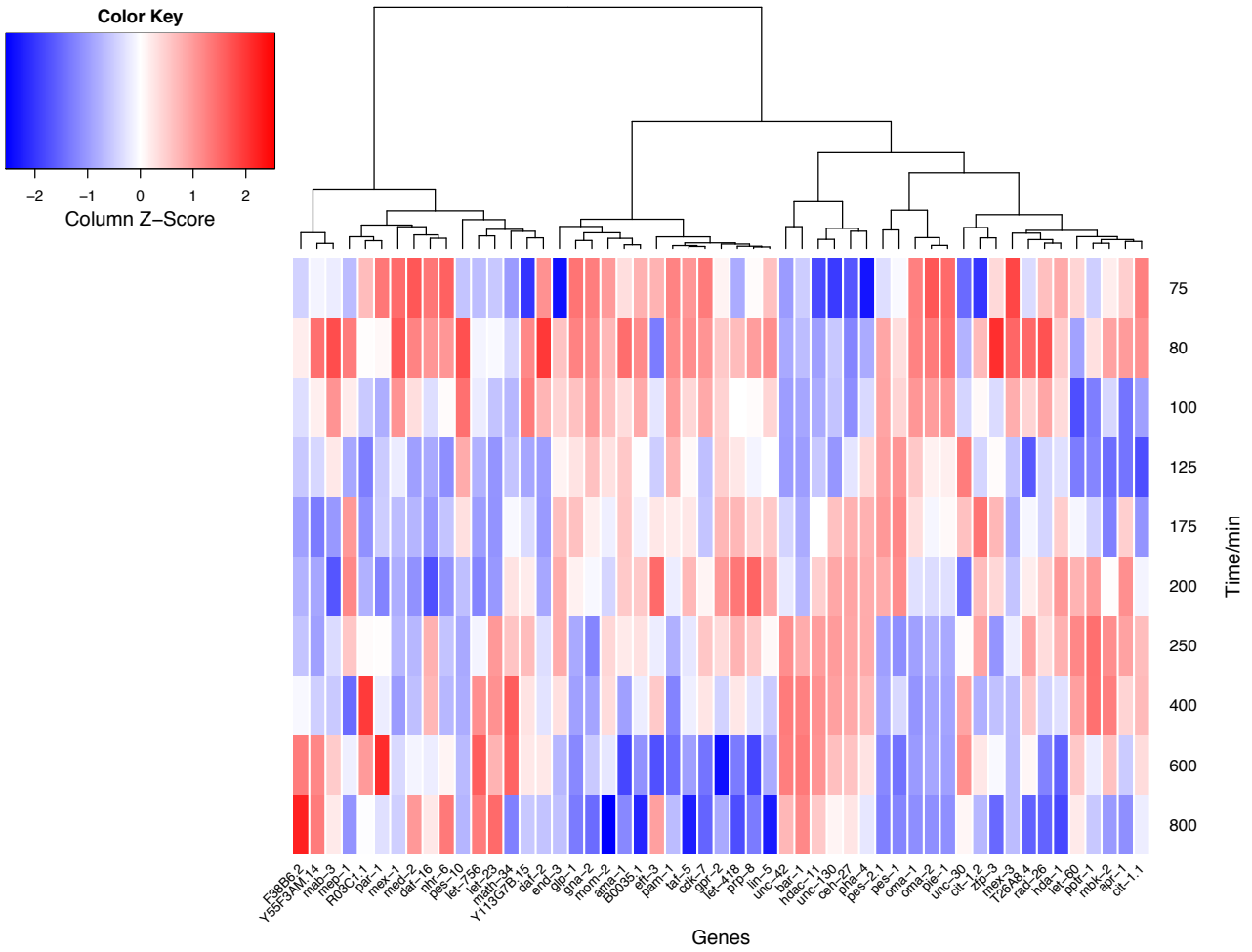

Figure 2: **Heatmap of all 64 candidate genes.**

The heatmap (Figure 2) does not show consistent expression patterns in clades and was hence not very informative, possibly due to the large number of gene expression patterns used in the analysis.

2.2.2 Network Construction

Networks were thus built to isolate a smaller group of genes, which have expression patterns that strongly correlate positively or negatively with that of *pie-1*.

Networks of all 64 genes were constructed using Pearson correlation values of above 0.8. As the community detection algorithm is only able to analyse positive correlation values, two different networks were created. The first network was built using only positive correlations (Figure 3) while the second network was constructed utilising only negative correlations (Figure 4).

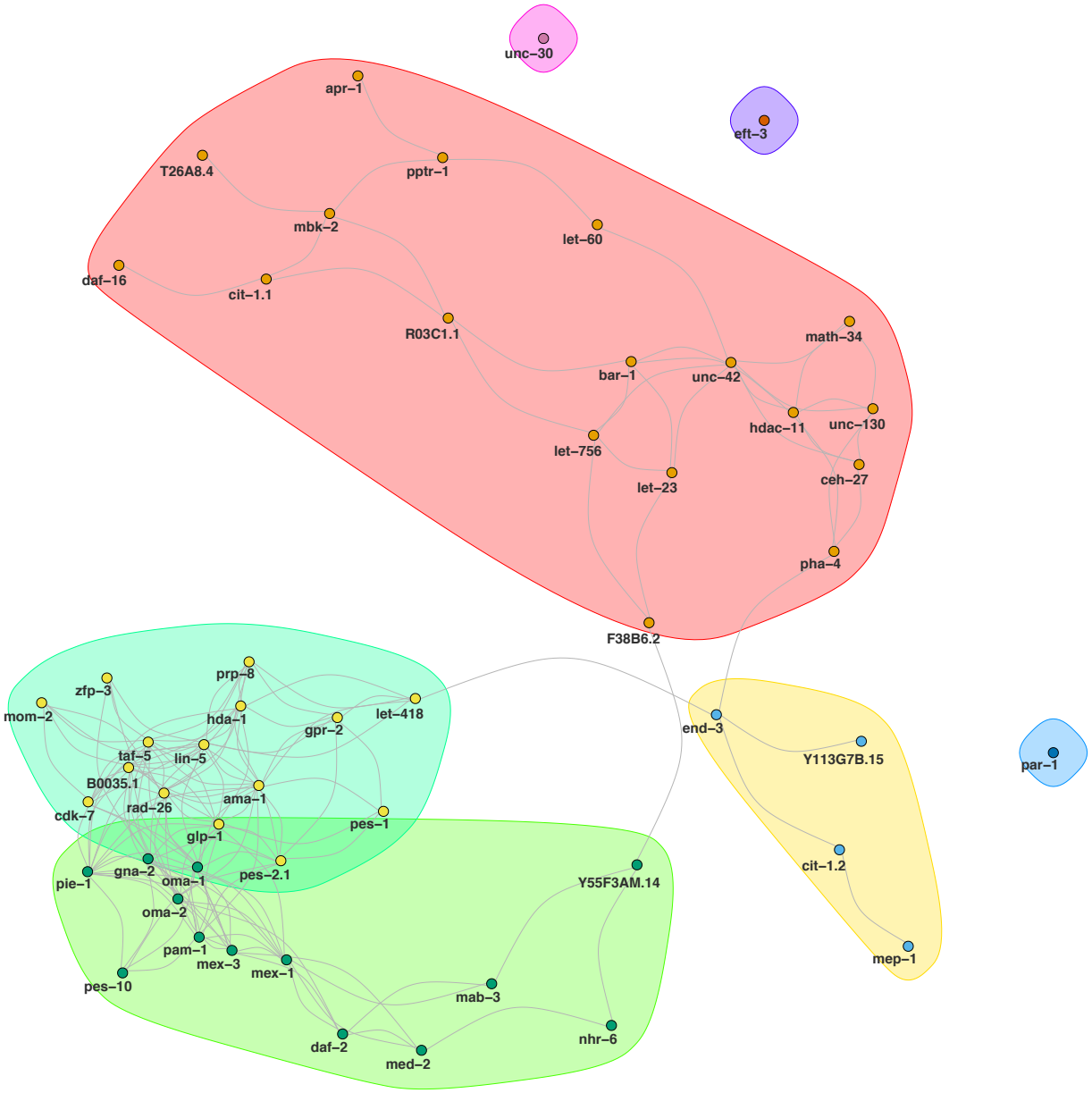

Figure 3: **Network constructed using only positive correlations between 64 candidate genes.**

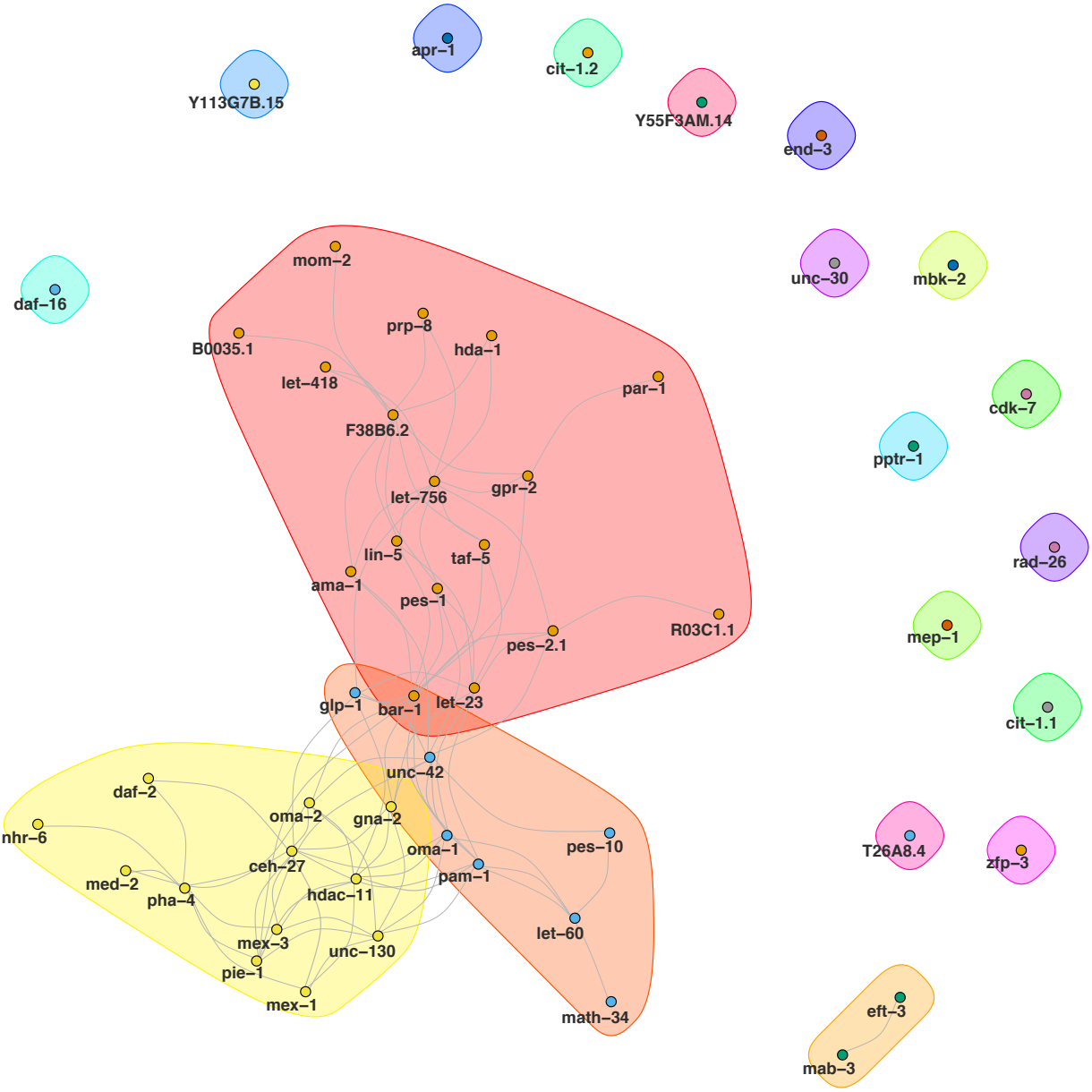

Figure 4: **Network constructed using only negative correlations between 64 candidate genes.**

2.3 Genes with expression patterns that correlate strongly with *pie-1*.

2.3.1 Postively correlated gene expressions

In the positive correlations network (Figure 3), 12 genes were found in the community containing *pie-1*. The gene expression data of these genes were used to generate another heatmap using Ward's method (ward.D2) (Figure 5).

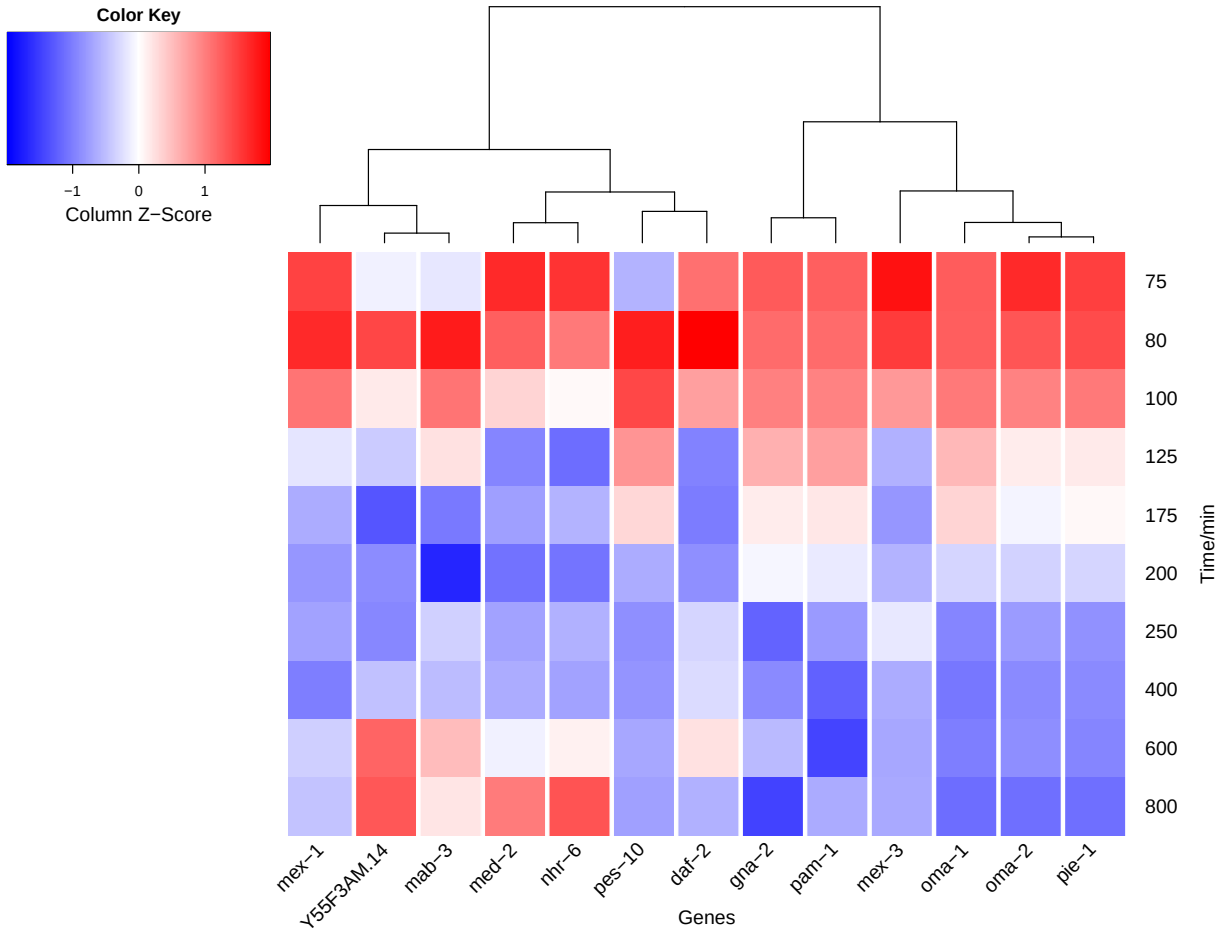

Figure 5: **Heatmap of gene expressions that correlate positively with *pie-1*.**

oma-1, *oma-2* and *mex-3* were subsequently selected for further analysis.

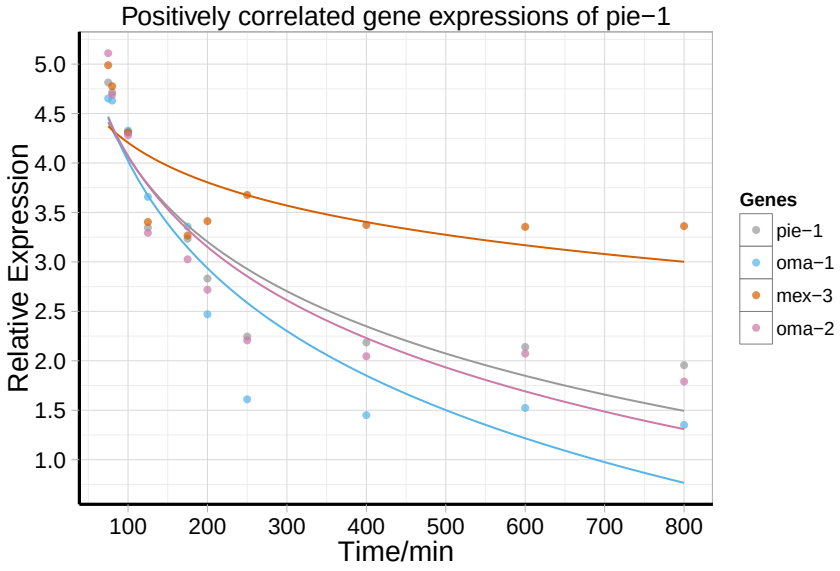

Figure 6: **Positively correlated gene expressions of *pie-1*.** Time on the x-axis indicates the amount of time since the 4-cell stage. Fitted lines were drawn using an exponential decay model.

All four genes decreases as embryos proceed with their development (Figure 6). *oma-1, oma-2* and *mex-3* expression data were best fitted with an exponential decay model (oma-1: adj. r^2 = 0.87, intercept = 11.23±1.08, slope = -1.57±0.20, both p<0.0001, res. df = 8)(oma-2: adj. r^2 = 0.83, intercept = 10.19±1.10 , slope = -1.33±0.20, both p<0.001, res. df = 8)(mex-3: adj. r^2 = 0.49, intercept = 6.87±1.00 , slope = -0.58±0.19, both p<0.05, res. df = 8).

2.3.2 Negatively correlated gene expressions

In the negative correlations network (Figure 4), 11 genes were found in the community containing *pie-1*. Their relative expression patterns were used to generate another heatmap using Ward's method (ward.D2) (Figure 7).

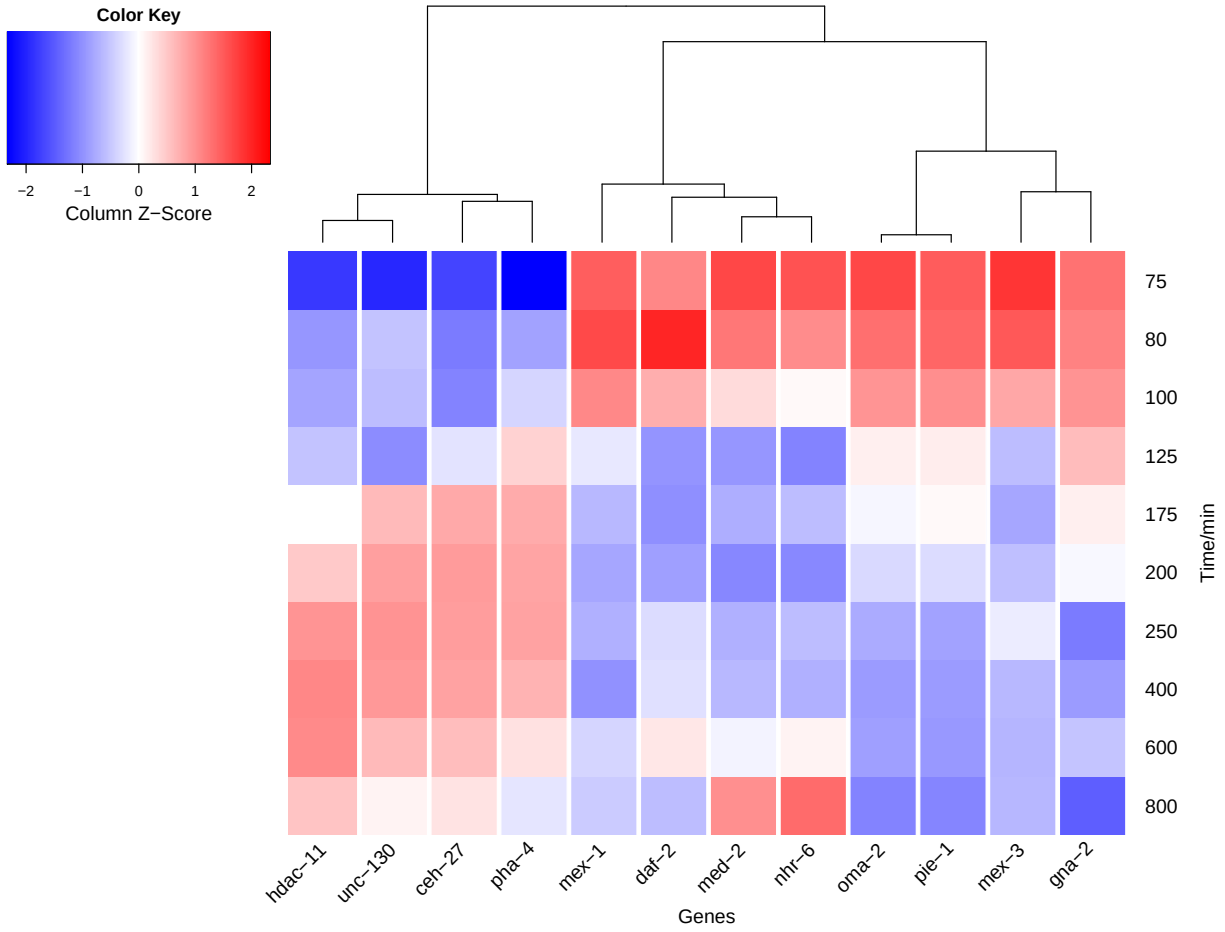

Figure 7: **Heatmap of gene expressions that correlate negatively with *pie-1*.**

hdac-11, unc-130, ceh-27 and *pha-4* have gene expression patterns that are the inverse of that of *pie-1* and are possible candidates of inhibition by PIE-1. Thus, these four genes were subsequently used for further analysis.

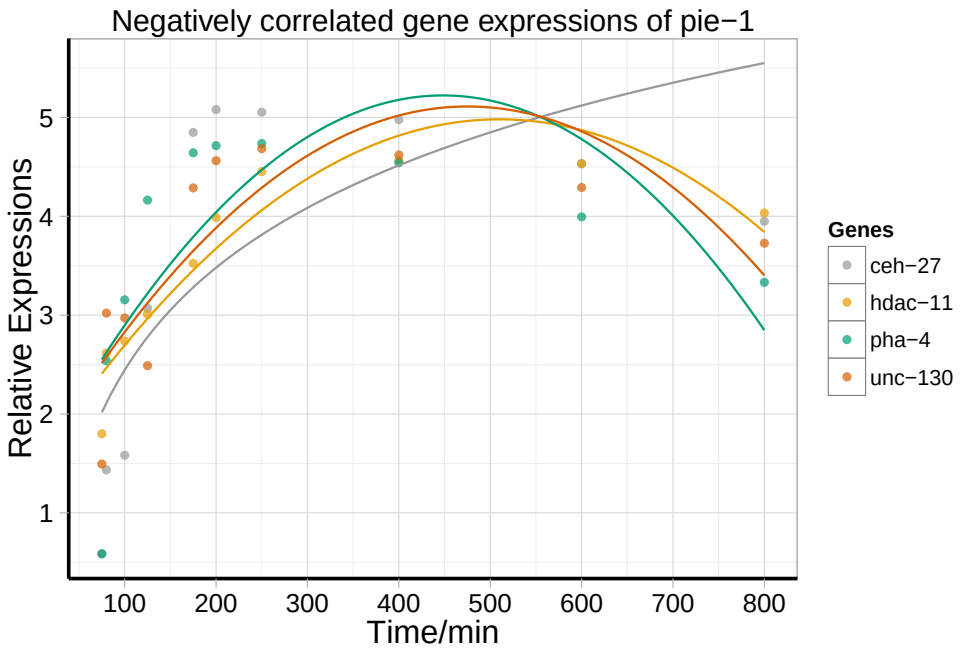

Figure 8: **Negatively correlated gene expressions of *pie-1*.** Time on the x-axis indicates the amount of time since the 4-cell stage.

All four genes have low expression levels at the 4-cell stage and have rapidly increasing expression which peaks at around 200 mins after the 4-cell stage, before expression levels start to fall. *ceh-27* expression levels were best modelled by an exponential decay curve while the expression data of the three other genes were best modelled by a quadratic function. Nonetheless, these models do not represent the data very well, as indicated by low adjusted R values and high p-values (Figure 8).

3 Discussion

3.1 Postively correlated genes

3.1.1 OMA-1 and OMA-2 indirectly activate PIE-1

Both *oma-1* and *oma-2* (oocyte maturation defective) encode CCCH-type zinc finger proteins and have redundant functions in oocyte maturation in *C. elegans* (Detwiler et al., 2001). OMA proteins repress transcription globally in the zygote and early germline blastomere by binding to TAF-4, an essential protein in the RNA polymerase II pre-initiation complex (Guven-Ozkan et al., 2008).

Another protein, ZIF-1, is a subunit of a E3 ligase that binds to PIE-1 and targets it for degra-

dation. OMA represses translation of *zif-1* by binding to the 5'untranslated region (UTR) of the *zif-1* mRNA, hence preventing the degradation of PIE-1. This indirectly allows transcriptional repression in later germline blastomeres by high levels of PIE-1 (Guven-Ozkan et al., 2010).

Furthermore, both OMA and PIE-1 are translated from germline-specific maternal RNAs and are associated with P granules, which are granules specific to germ cells (Shimada et al., 2006). As the ratio of germline cells to somatic cells exponentially decrease as the embryos develop, this may explain the exponential decay in gene expressions. Moreover, *oma-1, oma-2* and *pie-3* have similar gradient values and this could be a result of these genes being regulated together as they have such similar functions (Figure 6).

3.1.2 *mex-3*

mex-3 (muscle excess) encodes another maternally produced CCCH-type zinc finger protein. *mex-3* mutants give birth to embryos with excess muscle and hypodermal cells, resulting in a lethal phenotype (Pagano et al., 2009). Its protein, MEX-3, acts as a translational repressor by binding to MEX-3 recognition elements (MREs) found in the 3'UTR of several transcripts such as *pal-1*, hence maintaining the pluripotency of germline cells, where it accumulates (Pagano et al., 2009). However, it remains unknown if *pie-1* mRNA contains any MREs.

Similar to OMA, MEX-3 is a RNA-binding protein that is associated with P granules and act as a repressor of ZIF-1 (Draper et al., 1996; Oldenbroek et al., 2012), which may explain its similar trend of expression with that of *oma-1, oma-2* and *pie-1* (Figure 6).

3.2 Negatively correlated genes

HDAC-11, UNC-130, CEH-27 and PHA-4 were found to interact with PIE-1 in a recent study using enhanced yeast one-hybrid and two-hybrid assays (Reece-Hoyes et al., 2013). Moreover, community detection algorithm in the network with negative correlations placed the four genes that encode these proteins in the same community as PIE-1, further supporting the hypothesis that PIE-1 inhibits these proteins or genes. Nonetheless, there has been no further papers that have looked into the interaction between these four proteins and PIE-1.

3.2.1 *hdac-11* may regulate somatic neural differentiation.

Histone deacetylase-11 (HDAC-11), encoded by *hdac-11*, is an ortholog of the human HDAC11 protein. The human HDAC11 protein is enriched in oligodendrocytes and has been shown to be important in oligodendrocyte development (Liu et al., 2009). However, there has been no further research conducted regarding the *hdac-11* gene in *C. elegans*. Nonetheless, histone deacetylases play essential roles in suppressing and enhancing transcription across several locations in the genome by the modification of histones (Gregoretti et al., 2004). Hence, it is possible that a key germline transcription factor such as PIE-1, may inhibit either directly or indirectly the expression of *hdac-11*, thereby suppressing the expression of somatic neural differentiation genes.

3.2.2 *unc-130* is responsible for the formation of olfactory neurons.

unc-130 (uncoordinated-130) encodes a evolutionarily conserved Forkhead Box transcripition factor that can activate or repress transcription of other genes. This group of transcription factors play multiple complex roles in metazoan development, hence making it difficult to elucidate their functions. In *C.elegans*, UNC-130 participates in the formation of olfactory neurons as well as in the formation of mesoderm patterning (Kersey et al., 2016; Nash et al., 2000; Sarafi-Reinach and Sengupta, 2000). Similarly, PIE-1 may suppress theis gene expression to maintain pluripotency of germline cells.

3.2.3 *pha-4* participates in the organogenesis of the pharynx.

The *pha-4* gene is expressed during embryonic development, producing a FoxA transcription factor that acts as a organ identity factor. PHA-4 is crucial in the development of the pharynx in *C.elegans* (Mango et al., 1994). As PIE-1 maintains germline cell fate, it is likely that it participates in the inhibition of *pha-4* inhibition.

3.2.4 Functions of *ceh-27* remain unclear.

ceh-27 is crucial for embryogenesis and seems to be involved in the maintenance of hypodermal integrity. However, the functions of CEH-27 remains unclear as little research regarding the protein itself has been conducted.

4 Further Research

Many of the gene products discussed in this report are primarily accumulated in the germline blastomeres. Hence, it would be interesting to compare the transcriptomics of germline blastomeres to that of somatic blastomeres to elucidate other genes that are enriched in germline blastomeres. Such genes may display a highly correlated expression with *pie-1* and may not have been identified in this practical as expression levels were averaged throughout the embryonic cells.

Although statistical analysis of gene expressions is able to highlight significant correlations between gene expression patterns, additional supporting data will be needed before these gene interactions can be proven true. Moreover, further cell assays will be required to elucidate the molecular mechanisms behind such gene interactions.

Lastly, by researching on key embryonic transcription factors such as PIE-1 in *C. elegans*, we can gain a better understanding of the molecular mechanisms behind the maintenance of pluripotency and embryonic development. As there are many conserved pathways between the two species (Kaletta and Hengartner, 2006), using *C. elegans* as a model to understand human embryonic development is hence a useful tool.

References

Detwiler, M. R., Reuben, M., Li, X., Rogers, E. and Lin, R. (2001). Two Zinc Finger Proteins, OMA-1 and OMA-2, Are Redundantly Required for Oocyte Maturation in *C. elegans*. Developmental Cell *1*, 187–199.

Draper, B. W., Mello, C. C., Bowerman, B., Hardin, J. and Priess, J. R. (1996). MEX-3 is a KH domain protein that regulates blastomere identity in early *C. elegans* embryos. Cell *87*, 205–216.

Ghosh, D. and Seydoux, G. (2008). Inhibition of transcription by the *Caenorhabditis elegans* germline protein PIE-1: Genetic evidence for distinct mechanisms targeting initiation and elongation. Genetics *178*, 235–243.

Gilbert, S. F. (2000). Early Development of the Nematode *Caenorhabditis elegans*. Sinauer Associates.

Gregoretti, I. V., Lee, Y. M. and Goodson, H. V. (2004). Molecular evolution of the histone deacetylase family: Functional implications of phylogenetic analysis. Journal of Molecular Biology *338*, 17–31.

Guven-Ozkan, T., Nishi, Y., Robertson, S. M. and Lin, R. (2008). Global Transcriptional Repression in *C. elegans* Germline Precursors by Regulated Sequestration of TAF-4. Cell *135*, 149–160.

Guven-Ozkan, T., Robertson, S. M., Nishi, Y. and Lin, R. (2010). zif-1 translational repression defines a second, mutually exclusive OMA function in germline transcriptional repression. Development *137*, 3373–3382.

Kaletta, T. and Hengartner, M. O. (2006). Finding function in novel targets: *C. elegans* as a model organism. Nature reviews. Drug discovery *5*, 387–98.

Kersey, R. K., Brodigan, T. M., Fukushige, T. and Krause, M. W. (2016). Regulation of UNC-130/FOXD-mediated mesodermal patterning in *C. elegans*. Developmental Biology *416*, 300–311.

Levin, M., Hashimshony, T., Wagner, F. and Yanai, I. (2012). Developmental Milestones Punctuate Gene Expression in the *Caenorhabditis* Embryo. Developmental Cell *22*, 1101–1108.

Liu, H., Hu, Q., D'Ercole, A. J. and Ye, P. (2009). Histone deacetylase 11 regulates oligodendrocyte-specific gene expression and cell development in OL-1 oligodendroglia cells. GLIA *57*, 1–12.

Mango, S. E., Lambie, E. J. and Kimble, J. (1994). The pha-4 gene is required to generate the pharyngeal primordium of *Caenorhabditis elegans*. Development (Cambridge, England) *120*, 3019–3031.

Mello, C. C., Draper, B. W., Krause, M., Weintraub, H. and Priess, J. R. (1992). The pie-1 and mex-1 genes and maternal control of blastomere identity in early *C. elegans* embryos. Cell *70*, 163–176.

Nash, B., Colavita, A., Zheng, H., Roy, P. J. and Culotti, J. G. (2000). The forkhead transcription factor UNC-130 is required for the graded spatial expression of the UNC-129 TGF guidance factor in *C. elegans*. Genes and Development *14*, 2486–2500.

Oldenbroek, M., Robertson, S. M., Guven-Ozkan, T., Gore, S., Nishi, Y. and Lin, R. (2012). Multiple RNA-binding proteins function combinatorially to control the soma-restricted expression pattern of the E3 ligase subunit ZIF-1. Developmental Biology *363*, 388–398.

Pagano, J. M., Farley, B. M., Essien, K. I. and Ryder, S. P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. Proceedings of the National Academy of Sciences of the United States of America *106*, 20252–7.

Reece-Hoyes, J. S., Pons, C., Diallo, A., Mori, A., Shrestha, S., Kadreppa, S., Nelson, J., DiPrima, S., Dricot, A., Lajoie, B. R., Ribeiro, P. S. M., Weirauch, M. T., Hill, D. E., Hughes, T. R., Myers, C. L. and Walhout, A. J. M. (2013). Extensive rewiring and complex evolutionary dynamics in a *C.elegans* multiparameter transcription factor network. Molecular Cell *51*, 116–127.

Sarafi-Reinach, T. R. and Sengupta, P. (2000). The forkhead domain gene unc-130 generates chemosensory neuron diversity in *C. elegans*. Genes and Development *14*, 2472–2485.

Shimada, M., Yokosawa, H. and Kawahara, H. (2006). OMA-1 is a P granules-associated protein that is required for germline specification in *Caenorhabditis elegans* embryos. Genes to Cells *11*, 383–396.

Tenenhaus, C., Subramaniam, K., Dunn, M. A. and Seydoux, G. (2001). PIE-1 is a bifunctional protein that regulates maternal and zygotic gene expression in the embryonic germ line of *Caenorhabditis elegans*. Genes and Development *15*, 1031–1040.