

5 Single Cell Data Exploration

Bryan Teefy

12/20/2019

Load required libraries

```
library(URD)
library(Seurat)
```

Single Cell Data Exploration

We analyzed the homeostatic expression patterns of the 24 high mapping gene transcripts upregulated in response to hywi knockdown by interrogating single-cell expression data (Siebert, 2019). To interrogate expression, we used URD spline objects for ectoderm and endoderm. These objects contain expression data for genes that are expressed in at least 1% of the ectodermal or endodermal epithelial cells. We found epithelial expression for 16 of the 24 putative gene targets: 1) 13 gene transcripts are expressed in both the endodermal and ectodermal epithelial lineages, 2) one transcript is expressed only in the ectodermal epithelial lineage, and 3) two transcripts are expressed only in the endodermal epithelial lineage. No epithelial expression was found for 8 of the gene transcripts, which could indicate low expression in a homeostatic animal. Higher expression at the extremities in both the ectoderm and endoderm and lower expression in body regions would be consistent with Hywi-mediated repression. Putative targets could show high expression. The gene, t14391, shows such a pattern.

Load the Necessary Files

```
# Load the list of searchable high mapping genes for each lineage

endo_mappers <- read.table("objects/Endodermal_High_Mappers.txt", header = F)
colnames(endo_mappers) <- "ID"

ecto_mappers <- read.table("objects/Ectodermal_High_Mappers.txt", header = F)
colnames(ecto_mappers) <- "ID"

# Load the seurat object to load the searchable transcript IDs The seurat object
# (file Hydra_Seurat_Whole_Transcriptome) can be downloaded from Dryad:
# https://doi.org/10.5061/dryad.v5r6077. After download, place in the 'objects'
# folder

Hydra_Object <- readRDS("objects/Hydra_Seurat_Whole_Transcriptome.rds")

# Load single cell pseudotime objects (spline objects) and High Mapping Gene List
# The spline objects (file Hydra_URD_analysis_objects) can be downloaded from
# Dryad: https://doi.org/10.5061/dryad.v5r6077. After download, place in the
# 'objects' folder
```

```
endoderm.splines <- readRDS("objects/Splines-Endoderm.rds")
ectoderm.splines <- readRDS("objects/Splines-Ectoderm.rds")
```

Data Exploration

```
# Establish hFind function to return the searchable transcript IDs

hFind <- function(x) {
  return(Hydra_Object@data@Dimnames[[1]][grep(x, Hydra_Object@data@Dimnames[[1]],
    ignore.case = T)])
}

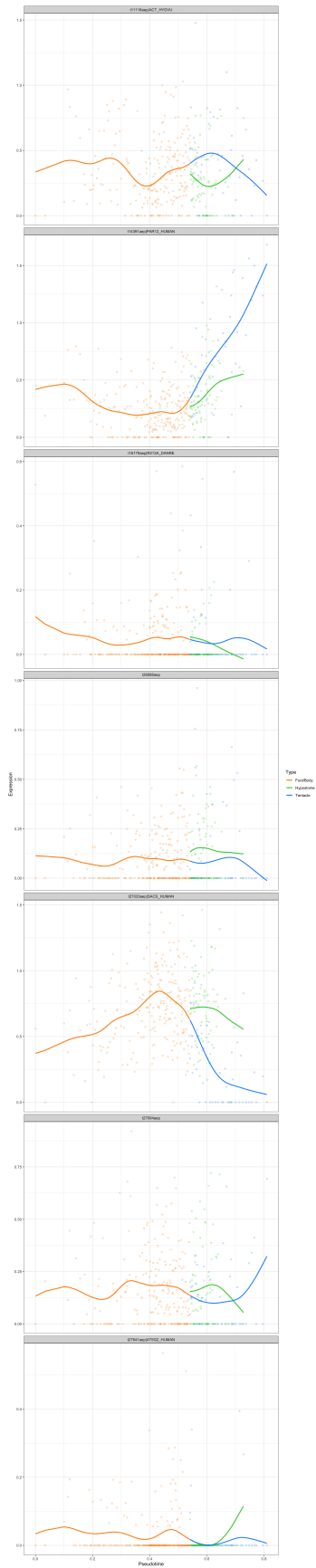
# Run hFind function to convert to searchable transcript IDs

endo_mappers$ID <- sapply(endo_mappers$ID, FUN = hFind)

ecto_mappers$ID <- sapply(ecto_mappers$ID, FUN = hFind)

# Generate Endodermal and Endodermal Plots

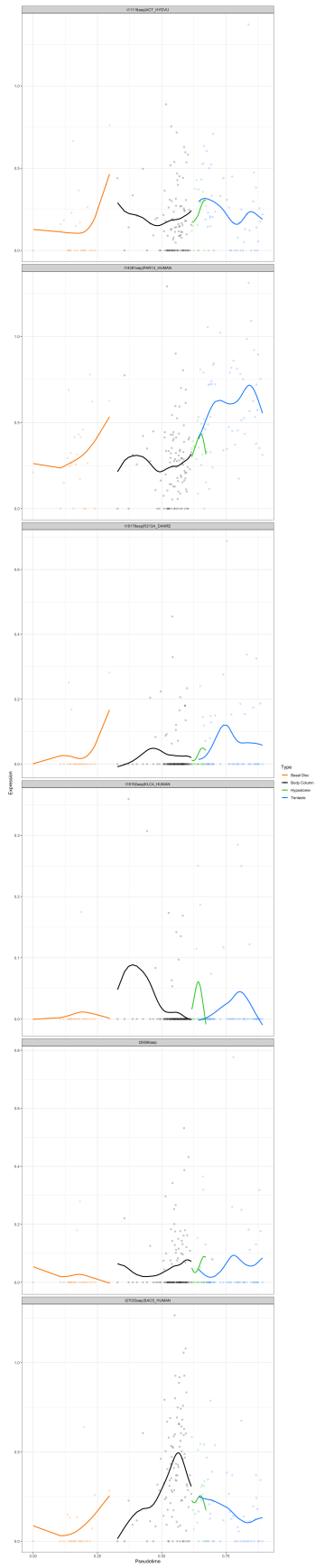
plotSmoothFitMultiCascade(endoderm.splines, genes = endo_mappers$ID[1:7], scaled = F,
  colors = c(`Foot/Body` = "#FF8C00", Tentacle = "#1E90FF", Hypostome = "#32CD32"),
  ncol = 1)
```



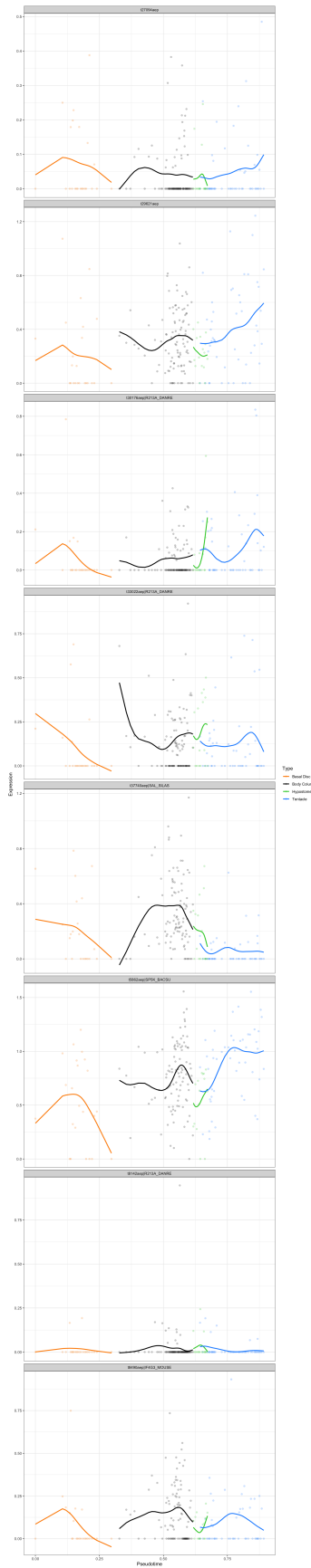
```
plotSmoothFitMultiCascade(endoderm.splines, genes = endo_mappers$ID[8:15], scaled = F,  
  colors = c(`Foot/Body` = "#FF8C00", Tentacle = "#1E90FF", Hypostome = "#32CD32"),  
  ncol = 1)
```



```
plotSmoothFitMultiCascade(ectoderm.splines, genes = ecto_mappers$ID[1:6], scaled = F,  
  colors = c(`Basal Disc` = "#FF8C00", `Body Column` = "#000000", Tentacle = "#1E90FF",  
    Hypostome = "#32CD32"), ncol = 1)
```



```
plotSmoothFitMultiCascade(ectoderm.splines, genes = ecto_mappers$ID[7:14], scaled = F,  
  colors = c(`Basal Disc` = "#FF8C00", `Body Column` = "#000000", Tentacle = "#1E90FF",  
    Hypostome = "#32CD32"), ncol = 1)
```

Software versions

This document was computed on Thu Jan 09 15:24:08 2020 with the following R package versions.

R version 3.5.3 (2019-03-11)

Platform: x86_64-apple-darwin15.6.0 (64-bit)

Running under: macOS Mojave 10.14.5

Matrix products: default

BLAS: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRblas.0.dylib

LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRlapack.dylib

locale:

[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:

[1] stats graphics grDevices utils datasets methods base

other attached packages:

[1] Seurat_2.3.4 cowplot_0.9.4 URD_1.0.3 Matrix_1.2-16 ggplot2_3.2.0

[6] knitr_1.22

loaded via a namespace (and not attached):

| | |
|-----------------------------|---------------------|
| [1] reticulate_1.11.1 | R.utils_2.8.0 |
| [3] tidyselect_0.2.5 | htmlwidgets_1.3 |
| [5] grid_3.5.3 | trimcluster_0.1-2.1 |
| [7] ranger_0.11.2 | BiocParallel_1.14.2 |
| [9] Rtsne_0.15 | munsell_0.5.0 |
| [11] destiny_2.10.2 | codetools_0.2-16 |
| [13] ica_1.0-2 | withr_2.1.2 |
| [15] colorspace_1.4-1 | Biobase_2.40.0 |
| [17] rstudioapi_0.9.0 | stats4_3.5.3 |
| [19] ROCR_1.0-7 | robustbase_0.93-3 |
| [21] dtw_1.20-1 | vcd_1.4-4 |
| [23] VIM_4.8.0 | TTR_0.23-4 |
| [25] gbRd_0.4-11 | labeling_0.3 |
| [27] Rdpack_0.10-1 | lars_1.2 |
| [29] GenomeInfoDbData_1.1.0 | polyclip_1.10-0 |
| [31] bit64_0.9-7 | farver_1.1.0 |
| [33] xfun_0.5 | ggthemes_4.2.0 |
| [35] diptest_0.75-7 | R6_2.4.0 |
| [37] GenomeInfoDb_1.16.0 | RcppEigen_0.3.3.5.0 |
| [39] hdf5r_1.0.1 | flexmix_2.3-15 |
| [41] bitops_1.0-6 | DelayedArray_0.6.6 |
| [43] assertthat_0.2.1 | SDMTools_1.1-221 |
| [45] scales_1.0.0 | ggraph_1.0.2 |
| [47] nnet_7.3-12 | gtable_0.3.0 |
| [49] npsurv_0.4-0 | rlang_0.4.0 |
| [51] scatterplot3d_0.3-41 | splines_3.5.3 |
| [53] lazyeval_0.2.2 | acepack_1.4.1 |
| [55] checkmate_1.9.1 | yaml_2.2.0 |
| [57] reshape2_1.4.3 | abind_1.4-5 |
| [59] backports_1.1.4 | Hmisc_4.2-0 |

| | | |
|-------|----------------------|-----------------------------|
| [61] | tools_3.5.3 | gplots_3.0.1.1 |
| [63] | RColorBrewer_1.1-2 | proxy_0.4-23 |
| [65] | BiocGenerics_0.26.0 | ggridges_0.5.1 |
| [67] | Rcpp_1.0.1 | plyr_1.8.4 |
| [69] | base64enc_0.1-3 | zlibbioc_1.26.0 |
| [71] | purrr_0.3.2 | RCurl_1.95-4.12 |
| [73] | rpart_4.1-13 | pbapply_1.4-0 |
| [75] | viridis_0.5.1 | S4Vectors_0.18.3 |
| [77] | zoo_1.8-4 | SummarizedExperiment_1.10.1 |
| [79] | haven_2.1.0 | ggrepel_0.8.1 |
| [81] | cluster_2.0.7-1 | magrittr_1.5 |
| [83] | data.table_1.12.2 | openxlsx_4.1.0.1 |
| [85] | gmodels_2.18.1 | lmtest_0.9-36 |
| [87] | RANN_2.6.1 | mvtnorm_1.0-10 |
| [89] | fitdistrplus_1.0-14 | matrixStats_0.54.0 |
| [91] | hms_0.4.2 | lsei_1.2-0 |
| [93] | evaluate_0.13 | smoother_1.1 |
| [95] | rio_0.5.16 | mclust_5.4.2 |
| [97] | readxl_1.3.1 | IRanges_2.14.12 |
| [99] | gridExtra_2.3 | compiler_3.5.3 |
| [101] | tibble_2.1.3 | KernSmooth_2.23-15 |
| [103] | crayon_1.3.4 | R.oo_1.22.0 |
| [105] | htmltools_0.3.6 | segmented_0.5-3.0 |
| [107] | Formula_1.2-3 | snow_0.4-3 |
| [109] | tidyr_0.8.3 | tweenr_1.0.1 |
| [111] | formatR_1.7 | MASS_7.3-51.1 |
| [113] | fpc_2.1-11.1 | boot_1.3-20 |
| [115] | car_3.0-3 | R.methodsS3_1.7.1 |
| [117] | gdata_2.18.0 | parallel_3.5.3 |
| [119] | metap_1.1 | igraph_1.2.4.1 |
| [121] | GenomicRanges_1.32.7 | forcats_0.4.0 |
| [123] | pkgconfig_2.0.2 | foreign_0.8-71 |
| [125] | laeken_0.5.0 | sp_1.3-1 |
| [127] | foreach_1.4.4 | XVector_0.20.0 |
| [129] | minpack.lm_1.2-1 | bibtex_0.4.2 |
| [131] | stringr_1.4.0 | digest_0.6.20 |
| [133] | tsne_0.1-3 | rmarkdown_1.12 |
| [135] | cellranger_1.1.0 | htmlTable_1.13.1 |
| [137] | curl_3.3 | kernlab_0.9-27 |
| [139] | gtools_3.8.1 | modeltools_0.2-22 |
| [141] | jsonlite_1.6 | nlme_3.1-137 |
| [143] | carData_3.0-2 | viridisLite_0.3.0 |
| [145] | pillar_1.4.2 | lattice_0.20-38 |
| [147] | httr_1.4.0 | DEoptimR_1.0-8 |
| [149] | survival_2.43-3 | glue_1.3.1 |
| [151] | xts_0.11-2 | zip_2.0.3 |
| [153] | png_0.1-7 | prabclus_2.2-7 |
| [155] | iterators_1.0.10 | bit_1.1-14 |
| [157] | ggforce_0.2.2 | class_7.3-15 |
| [159] | stringi_1.4.3 | mixtools_1.1.0 |
| [161] | doSNOW_1.0.16 | latticeExtra_0.6-28 |
| [163] | caTools_1.17.1.2 | dplyr_0.8.3 |
| [165] | irlba_2.3.3 | e1071_1.7-2 |
| [167] | ape_5.2 | |