**ISB Spelman Collaboration**

Meeting Notes and Identification of Datasets 6/28/19

Goals: Transcription regulatory networks mediating response to kanamycin using the R package Trena developed at ISB. It makes use of methods for reconstructing transcriptional regulatory networks, especially in species for which genome-wide TF binding site information is available.

Specific goals:

* Regulation of WBC19 - It is of particular interest because we observed 45-50 fold increase in expression in roots of plants exposed to kanamycin. Note that most experimental conditions and RNA-seq data would show a 2 fold change at the most for WBC19. The dramatic induction of WBC19 by kanamycin suggests a very specific gene regulatory network under these conditions.

* Identification of an Fe-Citrate transporter that is inhibited by kanamycin. Our modeling work predicts the existence of an Fe-Ci transporter that is inhibited by kanamycin. We do not know if transcription, translation or activity is inhibited. Transcription inhibition is the easiest place to start, using Trena to identify a list of candidate transporters. The inhibition of the putative Fe-Citrate transporter would be very strong and counter mirror the increased expression of WBC19.

**ATAC-seq**

Root specific ATAC-seq data available from Tannenbaum et al., 2018 paper to build model

Noted that in Materials and Methods, they used plants grown on MS ⅛ strength vertically for 14 days.

**RNA-seq or MICROARRAY DATA SETS**

Apparently about 50 samples (read libraries) that are root specific, representing a variety of experimental conditions would be sufficient.

Supplementing existing data with our own RNA-seq data would probably be valuable. We will see how many samples etc could be sequenced at ISB. In general 10M reads (50 bp reads) per sample give good coverage. Per info on their website, the NextSeq at ISB generates 400 million single end reads up to 300bp or 800 million paired end reads for 1 flow cell. That would allow us to use 1 flow cell for about 40 samples. That is about 12 experimental conditions \* 3 replicates each = total of 36 samples.

**Possible experimental conditions:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| condition | genotype | Kanamycin (mg/l) | Citrate (mM) | Iron | Zinc |
| 1 | Control | 0 | 0 | 1 X | 1 X |
| 2 | Control | 0 | 100 | 1 X | 1 X |
| 3 | Control | 50 | 0 | 1 X | 1 X |
| 4 | Control | 50 | 100 | 1 X | 1 X |
| 5 | Control | 0 | 0 | 1 X | 0.5 X |
| 6 | Control | 0 | 100 | 1 X | 0.5 X |
| 7 | Control | 50 | 0 | 1 X | 0.5 X |
| 8 | Control | 50 | 100 | 1 X | 0.5 X |
| 9 | Control | 0 | 0 | 0.5 X | 1 X |
| 10 | Control | 0 | 100 | 0.5 X | 1 X |
| 11 | Control | 50 | 0 | 0.5 X | 1 X |
| 12 | Control | 50 | 100 | 0.5 X | 1 X |

**Some previous work done on transcription factors affecting metal homeostasis genes.**

There is quite a bit of transcriptome work done to understand the response to iron deficiency and some of the TFs involved have been characterized.

In this paper the focus is on the transcription factor FIT which is a major regulator of genes that respond to iron deficiency. They have their own microarray data with links to the **GEO datasets (included below).**

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5048462/>

This paper also references other iron deficiency related expression data.

FIT works by dimerizing with a number of bHLH transcription factors: bHLH38, bHLH39, bHLH100, and bHLH101, whose expression is dramatically induced by low Fe stress and repressed under Fe overload ([Yuan et al., 2008](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4558652/#bib64); [Wang et al., 2013b](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4558652/#bib60)).

Another transcription factors involved in regulating iron homeostasis is POPEYE (PYE). It is specifically induced in root pericycle under Fe-deficient conditions.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2929094/>

PYE can interact with IAA-LEUCINE RESISTANT3 (ILR3), which is also a [bHLH](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4558652/#def1) transcription factor and may mediate metal homeostasis by regulating transporter levels ([Rampey et al., 2006](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4558652/" \l "bib40))

Two other bHLH TFs, bHLH104 and ILR3 appears to regulate Fe homeostasis upstream of PYE and other bHLH transcription factors.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4558652/#bib20>

**Other Data sets of interest**

1) Your starting point was this data set from roots of plants grown under different zinc regimes:

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE77286>

2) Data set from article above on FIT

(<http://www.ncbi.nlm.nih.gov/geo/>) under the accession numbers [GSE65934](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65934) and [GSE80281](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80281)

They have data for six day-old seedlings and roots of six week-old plants using wild type, a *fit* knock-out mutant and a FIT over-expression line grown under iron-sufficient or iron-deficient conditions.

3) Data from the PYE paper mentioned above

(<http://www.ncbi.nlm.nih.gov/geo/>) under the accession numbers [GSE65934](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE65934) and [GSE80281](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80281)

Time series data from roots after exposure to iron deficiency.

4) There is a whole set of data produced by NASC

<http://affymetrix.arabidopsis.info/>

The most interesting would be:

[NASCARRAY270](https://uniofnottm-my.sharepoint.com/personal/sean_may_nottingham_ac_uk/Documents/Forms/All.aspx?slrid=b7c65a9e%2D00ab%2D5000%2D8344%2Db2ea8f46ee8e&FolderCTID=0x012000FA504B24ED2B7D49A915A47E836A87E3&id=%2Fpersonal%2Fsean%5Fmay%5Fnottingham%5Fac%5Fuk%2FDocuments%2FUoN%20Box%20Migration%2FNASC%2FNASCarrays%2FBy%5FExperiment%5FID%2FExp270)  has data from root RNA of A. thaliana (Columbia) seedlings hydroponically grown in complete nutrient medium (control) and Zn deficient seedlings grown in -Zn nutrient medium (experimental).

[NASCARRAY225](https://uniofnottm-my.sharepoint.com/personal/sean_may_nottingham_ac_uk/Documents/Forms/All.aspx?slrid=b7c65a9e%2D00ab%2D5000%2D8344%2Db2ea8f46ee8e&FolderCTID=0x012000FA504B24ED2B7D49A915A47E836A87E3&id=%2Fpersonal%2Fsean%5Fmay%5Fnottingham%5Fac%5Fuk%2FDocuments%2FUoN%20Box%20Migration%2FNASC%2FNASCarrays%2FBy%5FExperiment%5FID%2FExp225) has data from roots of wild type plants (ecotype Columbia) grown under iron-sufficient conditions and iron-deficient conditions. They also have data from roots of frd3 mutants grown with and without iron.

[NASCARRAY71](https://uniofnottm-my.sharepoint.com/personal/sean_may_nottingham_ac_uk/Documents/Forms/All.aspx?slrid=b7c65a9e%2D00ab%2D5000%2D8344%2Db2ea8f46ee8e&FolderCTID=0x012000FA504B24ED2B7D49A915A47E836A87E3&id=%2Fpersonal%2Fsean%5Fmay%5Fnottingham%5Fac%5Fuk%2FDocuments%2FUoN%20Box%20Migration%2FNASC%2FNASCarrays%2FBy%5FExperiment%5FID%2FExp71) This data set deals with mycotoxins. Since WBC19 is induced by an antibiotic I would be interesting to have another data set with a biologically active compound. Data comes from Arabidopsis seedlings (not just roots) either WT or root expansion mutants that are less sensitive to mycotoxins seedlings upon mycotoxin treatment or upon solvent treatment (control condition) for 2 or 24 hours.

[NASCARRAY35](https://uniofnottm-my.sharepoint.com/personal/sean_may_nottingham_ac_uk/Documents/Forms/All.aspx?slrid=b7c65a9e%2D00ab%2D5000%2D8344%2Db2ea8f46ee8e&FolderCTID=0x012000FA504B24ED2B7D49A915A47E836A87E3&id=%2Fpersonal%2Fsean%5Fmay%5Fnottingham%5Fac%5Fuk%2FDocuments%2FUoN%20Box%20Migration%2FNASC%2FNASCarrays%2FBy%5FExperiment%5FID%2FExp35) This Data set is from Arabidopsis roots treated with Arbuscular Mycorrhizae forming fungus and mock-inoculated control plants. Note that the interaction is not productive and does not lead to the establishment of symbiosis. Arabidopsis (Col-0) were grown in pot culture (1:1 sand/Terra-Green) at low concentrations of phosphate. Three week-old plants were inoculated with surface-sterilised spores of Gigaspora rosea. RNA was isolated 3 days post inoculation.

[NASCARRAY577](https://uniofnottm-my.sharepoint.com/personal/sean_may_nottingham_ac_uk/Documents/Forms/All.aspx?slrid=b7c65a9e%2D00ab%2D5000%2D8344%2Db2ea8f46ee8e&FolderCTID=0x012000FA504B24ED2B7D49A915A47E836A87E3&id=%2Fpersonal%2Fsean%5Fmay%5Fnottingham%5Fac%5Fuk%2FDocuments%2FUoN%20Box%20Migration%2FNASC%2FNASCarrays%2FBy%5FExperiment%5FID%2FExp577) Data is from 5 weeks old wt and wrky33 mutants (GABI\_324B11) plants that are highly susceptible to the fungus Botrytis cinerea either mock treated or spray infected with Botrytis cinerea.

[NASCARRAY151](https://uniofnottm-my.sharepoint.com/personal/sean_may_nottingham_ac_uk/Documents/Forms/All.aspx?slrid=b7c65a9e%2D00ab%2D5000%2D8344%2Db2ea8f46ee8e&FolderCTID=0x012000FA504B24ED2B7D49A915A47E836A87E3&id=%2Fpersonal%2Fsean%5Fmay%5Fnottingham%5Fac%5Fuk%2FDocuments%2FUoN%20Box%20Migration%2FNASC%2FNASCarrays%2FBy%5FExperiment%5FID%2FExp151) has data from roots at various stages of development.

In addition if more data is needed these series might be more interesting:

The NASCARRAY137-146 all deal with various stresses (cold, heat UV etc)

The NASCARRAY149-156 all deal with various developmental stages in different organs

**Misc Paul’s Question:**

*Since we lack information about chromosome looping, how do we select ATAC-seq open chromatin regions plausibly connected to any target gene, from among the many which are nearby, or perhaps even hundreds of kb away?*

I believe there is some info about chromosome looping through S/MARs and they have been ‘mapped’ using in silico prediction and/or in vivo.Those that are within genes seem to downregulate the genes; those that are upstream seem to enhance overall expressing within the entire loop.

<http://www.plantphysiol.org/content/135/2/715>

<https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.0020021>

<http://www.plantcell.org/content/26/1/102.short>

For now though we may not need to be concerned about these.