Gene Regulatory Network Analysis of Transcriptomics Data Suggests Common and Diverse Regulators Across HD Model Systems and Perturbations

Jonathan R. Greene[1], Jian Chen[2], Jeffrey S. Aaronson[2], Vahri Beaumont[2], and Jim Rosinski[2]

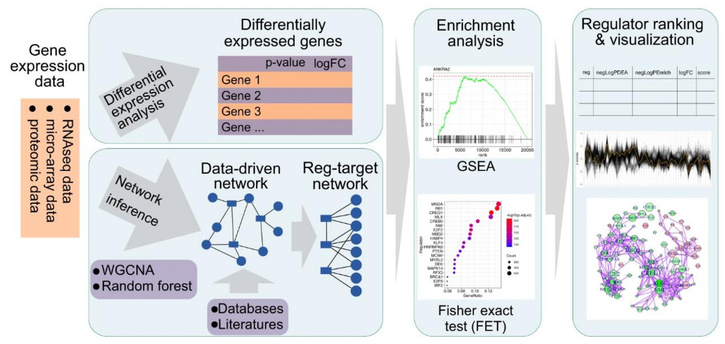
[1]Rancho BioSciences LLC, Rancho Santa Fe, CA; [2]CHDI Management/CHDI Foundation, Princeton, NJ

Abstract

Transcriptomic data are available across multiple Huntington's disease (HD) model systems. We have used the RegEnrich algorithm of Tao and Pandit (Tao et al, Nature Communications, 2022) to identify key gene expression regulators across a number of studies involving mouse HD models and perturbations from the literature and public data repositories. In addition to identifying potential "consensus" gene regulators that drive the progression of disease, the analysis suggests which of these regulators also play a role in rescue of transcriptional dysregulation under different perturbations.

RegEnrich Gene Regulatory Networks (GRNs)

The R RegEnrich package of Tao, et al combines network inference, differential gene expression analysis, and gene set enrichment analysis to identify key gene expression regulators in 'omics data sets such as bulk RNA-Seq. The overall analysis is described in this figure taken directly from their publication:



RegEnrich consists of four major steps: differential expression analysis, regulator-target network construction, enrichment analysis, and regulator ranking and visualization.

The key output from the analysis provides a list of gene expression regulators for a given differential expression contrast that are scored and ranked according to the following formula, again taken directly from their publication:

$$score = f(-\log(P_E)) + f(-\log(P_D))$$

where $f(x) = \frac{x - \min(x)}{\max(x) - \min(x)}$, and P_D is the vector of *p* values of regulators obtained from differential expression analysis, P_E is the vector of p-values of regulators obtained from the enrichment analysis.

GRNs from Allelic Series Striatum RNA-Seq

We used RNA-seq data from striatum samples of the HD model allelic series (Langfelder, et al, GEO: GSE65774) for *HTT* Q-lengths Q80, Q92, Q111, Q140, and Q175 at ages of 2, 6, and 10 months.

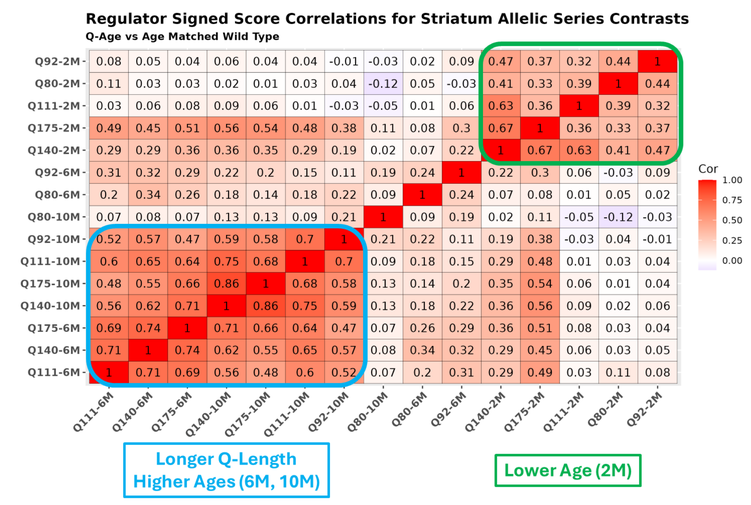
The following parameters were used to construct Gene Regulatory Networks (GRNs) for contrasts comparing each Q-length to age matched wildtype (WT) samples:

- Regulators for mouse were taken from TcoFbase (https://bio.liclab.net/TcoFbase/index.html)
- The Random Forest method was used to build the GRNs
- Fisher exact test (FET) was used for the enrichment step
- In addition to deriving scores for regulators, the scores were "signed" as positive or negative based on the direction of dysregulation for the given regulator

Striatal allelic series GRNs were clustered based on correlations between their signed RegEnrich scores. In the heatmap displaying these results, the following logical groupings can be inferred that generally align with combined age and Q-length in terms of disease progression:

- Lower left: 6 and 10 month contrasts for longer Q-lengths (Q111, Q140, and Q175) as well as the 10 month contrast for Q92.
- Upper right: 2 month samples for all Q-lengths, with those for Q140 and Q175 showing somewhat higher correlations with samples in the lower left.
- Q80 6 month and 10 month samples lie along with Q92 samples lie in a "transition zone" between these other two regions.

GRNs from Allelic Series Striatum RNA-Seq



Based on this logical grouping across samples we clustered the signed RegEnrich scores across individual regulators and were able to define clusters of regulators that follow progression of the disease based on combined age and Q-length:

- Early-Down
- Early-Up
- Late-Down Class 1
- Late-Down Class 2
- Late-Up Class 1
- Late-Up Class 2

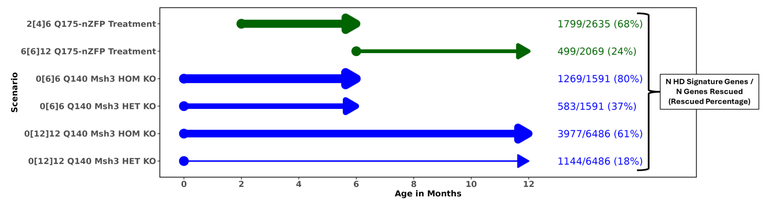
In addition, we can map a number of these regulators to recently described "**Phase C**" and "**Phase D**" DEGs from HD postmortem striatal samples reported by Handsaker, et al, 2025 as indicated in the tables below (Phase C = blue, Phase D, red).

GRNs in HD Perturbation Models

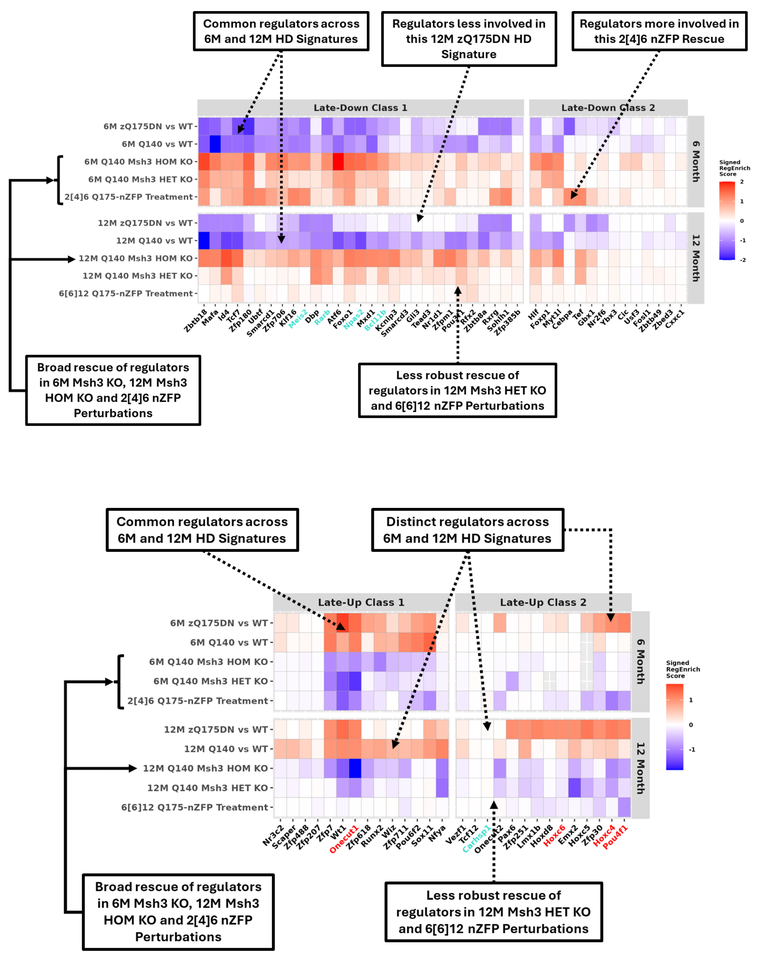
GRNs were constructed from RNA-Seq data for two HD model "perturbation" studies:

- zQ175DN mice treated with a CAG-directed and neuronally expressed zinc finger protein (nZFP) – GEO: GSE270727
  - Treatment at 2M, sampled at 6M, designated **2[4]6**
  - Treatment at 6M, sampled at 12M, designated **6[6]12**
- Q140 mice in which the Msh3 gene has been knocked out (Wang, et al, 2025)
  - Heterozygous and Homozygous Msh3 KO
  - Sampled at 6M and 12M, designated **0[6]6** and **0[12]12**

Rescue of transcriptome dysregulation is robust in the 2[4]6 nZFP treatments as well as in the Msh3 HOM KO perturbations as shown here with arrow start and end points designating perturbation start time and sampling time respectively. Rescue was determined as described in Marchionini, et al, 2022.



GRNs were constructed for underlying signature contrasts (eg Un-perturbed Q vs WT) and perturbation contrasts (eg Perturbed Q vs Un-perturbed Q) and signed RegEnrich scores for allelic series regulators were plotted as heatmaps. Note that opposing direction of perturbation and signature scores indicates rescue for a particular regulator.



Conclusions

1. GRNs were constructed from striatal bulk RNA-Seq data from the allelic series that cluster logically with respect to age and Q-length.
2. Individual regulators can also be clustered into groups that track with disease progression.
3. These regulator clusters can be used to suggest key regulators for disease progression and perturbations in other studies.

References

1. Tao, et al, RegEnrich gene regulator enrichment analysis reveals a key role of the ETS transcription factor family in interferon signaling, Nature Communications, 2022 (PMID: 35017649)
2. Zhang, et al, TcoFBase: a comprehensive database for decoding the regulatory transcription co-factors in human and mouse, Nucleic Acids Res, 2022 (PMID: 34718747)
3. Langfelder, et al, Integrated genomics and proteomics define huntingtin CAG length-dependent networks in mice, Nat Neurosci, 2016. (PMID: 26900923)
4. Handsaker, et al, Long somatic DNA-repeat expansion drives neurodegeneration in Huntington's disease, Cell, 2025 (PMID: 39824182).
5. Zeitler, et al, Allele-selective transcriptional repression of mutant HTT for the treatment of Huntington's disease, Nat Med, 2019. (PMID: 31263285)
6. Wang, et al, Distinct mismatch-repair complex genes set neuronal CAG-repeat expansion rate to drive selective pathogenesis in HD mice Cell, 2025 (PMID: 39938516).
7. Marchionini, et al, Benefits of global mutant huntingtin lowering diminish over time in a Huntington's disease mouse model, JCI Insight, 2022. (PMID: 36278490)