UNIVERSITY OF CALIFORNIA, SAN DIEGO

Elucidation of time-varying gene regulatory networks controlled by REST during neural differentiation of hiPSCs

A Thesis submitted in partial satisfaction of the requirements for the degree
Master of Science

in

Bioengineering

by

Vivek Parthasarathy

Committee in charge:

Professor Shankar Subramaniam, Chair
Professor Marcos Intaglietta
Professor Shyni Varghese

2016

The Thesis of Vivek Parthasarathy is approved, and it is acceptable

in quality and form for publication on microfilm and electronically:

_____

_____

_____
Chair

University of California, San Diego

2016

DEDICATION

To Mother, for always telling me that I can

To Father, for guidance that helped me figure out how I could

To my high school teacher, Santhi Sathiapalan, for providing me with a solid foundation in the life sciences

To my guru, Shankar Subramaniam, for believing in me, ensuring that I actually did

And to God, for everything else

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

REST.....RE1-Silencing Transcription Factor
bHLH ..... Basic Helix-Loop-Helix
GRN ..... Gene Regulatory Network
hiPSC ..... Human Induced Pluripotent Stem Cell
ECM .....Extracellular Matrix
GO..... Gene Ontology
TF..... Transcription Factor
FA..... Functional Annotation
BMP .....Bone Morphogenic Protein

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my thesis advisor, Dr. Shankar Subramaniam, for his infinite patience and for providing me with invaluable guidance. I could not have asked for a better interdisciplinary introduction to Systems Biology research.

I would like to acknowledge Dr. Noel Buckley (Oxford University, UK) for collaborating with our lab and providing the datasets we analyzed.

I would also like to thank Dr. Shakti Gupta for his involvement in my project, and for the enriching discussions that helped me gain perspective.

To Dr. Mano Ram Maurya, Dr. Jun Min, Sindhu, Priya, Nassim, Shamim, Dr. Nirupama (Pam) Bhattacharya, Kenan, Carol and all other denizens of the Subramaniam lab for creating an amazing environment to learn and work in.

I would like to thank my committee members, Dr. Marcos Intaglietta and Dr. Shyni Varghese, for their time and consideration.

I would also like to thank my friends and family for their support. Chetan Srinath (Burge Lab, MIT), for all the conversations about research, neurons and life. Dr. Anjanaa Narayan, for being a pillar of support. My brother, Mukunth, for being around when it mattered.

Last, but not least, I would like to thank the Department of Bioengineering and UC San Diego for the journey thus far.

ABSTRACT OF THE THESIS

Elucidation of time-varying gene regulatory networks controlled by REST during neural differentiation of hiPSCs

by

Vivek Parthasarathy

Master of Science in Bioengineering

University of California, San Diego, 2016

Professor Shankar Subramaniam, Chair

RE1-Silencing Transcription Factor (REST), a member of the Kruppel-type zinc finger transcription factor family, is believed to act as a master negative regulator of neurogenesis. During neurogenesis, the decreasing expression of REST leads to increased expression of neuronal

genes and the emergence of neuronal processes over time. The temporal patterns of REST-controlled processes and transcriptional regulatory events underlying neural induction and early neural development have not been studied utilizing a systems-level approach on high-throughput time course data.

Human-derived induced Pluripotent Stem Cells (iPSCs) combined with well-established neural differentiation protocols allow for the in-vitro elucidation of gene expression patterns characteristic of human neurodevelopment during this crucial early in-vivo developmental phase.

Using time series data capturing genome-wide transcriptome information from human iPSCs differentiating into i) Cortical and ii) Hypothalamic neurons, REST-controlled gene regulatory networks (GRNs) were generated. These GRNs captured transcription factor-target regulatory interactions across the time series and were investigated in order to elucidate early neurodevelopment towards the two neuronal phenotypes. Functional enrichment analysis of gene sets obtained from these GRNs was used to determine where along the time series different REST-controlled neuronal processes emerged.

The systems-level approach allowed for temporal resolution of genome-wide trans-regulatory interactions over the time course, the dynamics of which underlie neural differentiation and development. The outcome of the research is a novel qualitative kinetic model consisting of the time-varying GRNs under the control of REST that gives insight into the i) temporal sequence of emergent neuronal processes accompanying neurodifferentiation and ii) the temporal

sequence of key transcriptional regulatory events underlying the early neural differentiation of hiPSCs.

CHAPTER 1: INTRODUCTION

hiPSCs:

Human induced pluripotent stem cells (hiPSCs) have shown potential in the elucidation of the underlying molecular mechanisms driving early human neurodevelopment. Studies have shown that human iPSCs and human ESCs utilize similar transcriptional networks over the same developmental time course to generate neuroepithelia and neuronal types in response to the same differentiation protocols [1].

Given the difficulties in studying neonatal brain development in-vivo in humans, iPSCs and ESCs have become important tools for studying brain development in-vitro in recent times. Studies have concluded that hESCs and hiPSCs are molecularly and functionally equivalent [2]. The importance of various pathways and transcription factors in driving neural differentiation and establishing neural identity have been established, although there is very little work on the interplay between these signaling cues and the expression of transcriptional regulators driving neurodevelopment over time.

Neurons from iPSCs have mainly been studied because of the promise they show in the study of the molecular mechanisms driving neurological diseases. With respect to single-step induction methods driving fibroblasts to become functional neurons, one of the major drawbacks is the establishment of a homogenous neuronal population with little or no subtype specificity. Although a homogenous neuronal culture could be seen to reflect in-vivo neuronal tissue

homogeneity, it also precludes us from understanding the molecular mechanisms driving differentiation into specific neuronal subtypes [3].

Neural Differentiation:

Our current understanding of the mechanisms driving neural induction in vertebrates is based on a traditional model consisting of the temporal order of signals driving tissue specification of dividing cells. In this traditional model, the neural plate emerges from dorsal ectoderm via blockade of BMP signaling. In the absence of BMP signaling, the emerging neural plate gives rise to the anterior forebrain with additional signals such as FGF, WNT, Notch and RAR required to give rise to more posterior regions [4, 5]. This cascade of signals driving neural induction has been shown to be conserved across mammals and has also been demonstrated in human embryonic stem cells (hESCs). However, there is very little knowledge on how these different signaling cues influence gene expression and emergence of phenotype across time during development.

REST and its co-repressor CoREST (RCOR1: REST Co-Repressor 1) are known to target members of the SHH gradient morphogen neurodevelopmental signaling pathway [6].

One of the signaling pathways required for neural differentiation is the Wnt signaling pathway (both Canonical and Non-canonical). This pathway, also responsible for neural crest induction, is known to act via LEF1-mediated transcriptional modification [7].

Literature has reported that BAM factors- ASCL1, POU3F and MYT1L play an important regulatory role in converting fibroblasts to functional neurons. It has also been established that transcription factors are drivers of neural development. Yamanaka factors are known to play a central role in the reprogramming of iPSCs into neural stem cells/progenitor cell. These observations serve to highlight the importance of transcriptional regulatory networks in driving neural differentiation and the elucidation of these dynamic networks across the differentiation time course [8].

bHLH factors play key roles in development and regeneration of the nervous system. Hes factors directly repress the expression of proneural genes such as ASCL1. The downregulation/absence of Hes factors results in an increase of proneural gene expression resulting in accelerated neurodevelopment [9]. ASCL1 alone is known to drive conversion of human fibroblasts to action potential-firing neuronal cells [10].

REST:

RE1 Silencing Transcription factor (REST [11]) is a major regulator in neuronal development by the repression of RE1 motif which is associated with neuron specific genes.

There are about 1300 RE1 sites in the human and murine genomes. REST was initially identified as a neuronal gene inhibitor in non-neuronal cells and has a very high association with non-coding RNAs such as miRNA 9 and

miRNA 124 implicated to have a direct role in neuronal development by targeting other developmental genes including ASCL1 [12].

Major functions of REST include axon guidance, vesicle trafficking, maintenance of the cytoskeleton, regulation of the ECM, neurite outgrowth and signaling. REST regulates the transitions from the pluripotent state to neural stem/neural progenitor cell states and to mature neuron [13].

REST expression decreases with neural lineage commitment. REST-mediated chromatin remodeling is required in neural progenitors for proper S-phase dynamics, as part of its well-established role in repressing neuronal genes until terminal differentiation [14].

In neural stem cells REST is seen to recruit corepressors (RCOR1/Co-REST, HDAC1) and forms a repressive protein complex that is then able to silence the neuronal genes by reversible chromatin remodeling. When REST is repressed, activation of neuronal genes was observed. It has been found that REST is required for the correct execution of neuronal differentiation programmers, but is not required for neural commitment per se or for maintaining the multipotent status of stem cells [15].

Gene Regulatory Networks:

Candidate-gene approaches to understanding development have their limitations. Developmental changes in cell phenotype or state are fundamentally underwritten, not by a single molecule or pathway, but by multiple genes and interactions. Systems level approaches on high-throughput transcriptomic data

allow us to glean insight on the global expression patterns of expressed genes at a given point in time.

Intuitively, the prior transcriptomic state of a differentiating cell sets a context in which a set of regulatory TFs establishes the current transcriptomic state (representative of phenotype). This occurs through regulatory interactions of these TFs with their target genes. These interactions result in a change in target gene expression driving the cell from prior to current transcriptomic state.

The set of all these TF-target gene regulatory interactions along the developmental axis represent the underlying molecular interactions that govern the differentiation process driving the cell closer to the final expression state and phenotype (fate commitment) [16, 17]. Time series data allows us to elucidate differentiation dynamics through the construction of GRNs comprised of transcriptomic states and regulatory interactions across time. GRNs are a graphical representation of changing transcriptomic states, where nodes represent genes, edges represent regulatory interactions and kinetics are captured through the changing expression levels of genes that constitute the transcriptome.

The portions of the transcriptome that are involved in TF-target regulatory interactions and/or showing statistically significant change in expression (compared to the prior transcriptomic state) can be qualitatively annotated through use of time-series data. These annotations would give perspective on the temporal changes in known biological function that accompany the development of the cell towards the final phenotype.

CHAPTER 2: METHODS

Neural Differentiation Protocols:

Neural differentiation protocols between datasets differ only in the treatments between Day 0 and Day 7. The difference in protocols has been outlined below:

i) **Hypothalamic Protocol:** 1microM Dorsomorphin (BMP inhibitor), 2microM XAV939 (Wnt inhibitor), 10microM SB431542 (TGFbeta inhibitor)

ii) **Cortical Protocol:** 100nM LDN193189 (This is a more potent BMP inhibitor compared to Dorsomorphin used in Hypothalamic protocol), 10microM SB431542 (TGFbeta inhibitor)

Addition of DAPT, a Notch inhibitor, was a part of both protocols and was done at Day 22. The difference in differentiation protocols between Hypothalamic and Cortical datasets is highlighted in **Figure 1.**

					Neuralization Hypothal	Neuralization Cortical		Neuron Differentiation				Maintenance
Sample #		Day	Passage	Stage	SB31542+ Dorsomorphin + XAV	SB31542+ LDN193189	Y27632	DAPT	DMEM/F12 + N2 50:50 w Neurobasal + B27 + glutamax	Neurobasal + B27 + Glutamax	AA2P	BDNF
		-1	Day -1	iPSC								
		0			SB+D,X	SB+L			D/F/N2-NB,B27,Glu			
		1			SB+D,X	SB+L			D/F/N2-NB,B27,Glu			
		2			SB+D,X	SB+L			D/F/N2-NB,B27,Glu			
		3			SB+D,X	SB+L			D/F/N2-NB,B27,Glu			
		4			SB+D,X	SB+L			D/F/N2-NB,B27,Glu			
		5			SB+D,X	SB+L			D/F/N2-NB,B27,Glu			
		6			SB+D,X	SB+L			D/F/N2-NB,B27,Glu			
1		7	Day 7		SB+D,X	SB+L	Y27		D/F/N2-NB,B27,Glu			
2		10							D/F/N2-NB,B27,Glu			
3		12	Day 12				Y27		D/F/N2-NB,B27,Glu		AA2P	
4		15							D/F/N2-NB,B27,Glu		AA2P	
		17	Day 17				Y27		D/F/N2-NB,B27,Glu		AA2P	
5		20							D/F/N2-NB,B27,Glu		AA2P	
6		22	Day 22	Neuroepithelial			Y27	DAPT		NB,B27,Glu	AA2P	
7		23						DAPT		NB,B27,Glu	AA2P	
8		24						DAPT		NB,B27,Glu	AA2P	
9		25						DAPT		NB,B27,Glu	AA2P	
10		26						DAPT		NB,B27,Glu	AA2P	
11		28		Neuron				DAPT		NB,B27,Glu	AA2P	
12		30								NB,B27,Glu	AA2P	
13		35								NB,B27,Glu	AA2P	
14		40								NB,B27,Glu	AA2P	BDNF
15		45								NB,B27,Glu	AA2P	BDNF
16		50								NB,B27,Glu	AA2P	BDNF

Figure 1: Overview of differentiation protocols used for neural specification of iPSCs. Table outlining protocols used to generate data for Hypothalamic (in blue box) and Cortical (in orange box) datasets. Note the addition of XAV939, a Wnt-inhibitor, between Day 0 - Day 7 in the Hypothalamic protocol.

Dataset:

Baseline sample consisted of untreated iPSCs obtained from human foreskin fibroblasts from a normal subject. hiPSCs were differentiated to cortical neuron and hypothalamic neuron (2 datasets). Data provided was collected at over 3 individual sequencing runs (Day 0 (3 timepoints, common to both datasets), Days 7-50 Hypothalamic and Days 7-50 Cortical (16 timepoints for each dataset). RNA-Seq data obtained were single replicate, single end reads consisting of 50 bases/read. The number of input reads ranged from 11.8M – 30.2M.

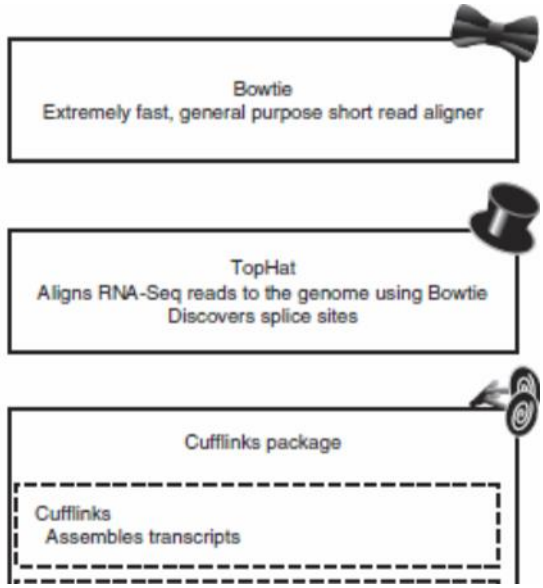

Figure 2: RNA-Seq data processing pipeline. [18]

Data Processing:

Alignment of raw RNA-Seq reads to the genome was done using Bowtie. Mapping to the genome was automatically done using TopHat aligner [18] guided by an annotated gene models (GTF file) obtained from Ensembl (http://www.ensembl.org). The overall read mapping rate ranged from 96.6% - 97.7%. Uniquely aligned reads/sample ranged from 10.7M – 27.5M. The resulting alignment data from TopHat were then fed to an assembler Cufflinks to assemble aligned RNA-Seq reads into transcripts. Annotated transcripts were obtained from the Ensembl database. Gene symbols (HGNC symbols) annotated to samples were used for further analysis. All samples were merged into the 19 timepoint timecourse after Cufflinks analysis. Transcript abundances were measured in Fragments Per Kilobase of exon per Million fragments

mapped (FPKMs). Cortical and Hypothalamic datasets were merged by gene identifier and expression data (across all timepoints, both datasets) were quantile normalized together before separating data into individual datasets.

Initial Data Visualization:

Hierarchical clustering of gene expression profiles across the time course was performed using the (pheatmaps) package in R. A dendrogram comparison of time course gene expression profiles for all expressed genes was done across datasets. The Euclidean distance matrix was used in clustering data for dendrogram generation.

Grouping of Expression Data:

The datasets were individually grouped into 6 different groups, with each group consisting of a set of 3 consecutive timepoints (Control + Groups 1-5). Grouping of timepoints and relation of timepoints/groups to the profile of REST and the number of days in culture are shown in **Fig 3**.

Generating Differentially Regulated Transcripts Across the Time Series:

Variance modeling-based t-test (Cyber-T) [19] was used to identify significant and differentially expressed transcripts between pairwise consecutive groups across the time series. Cyber-T uses a Bayesian estimate of the variance among the transcript expression values. An unpaired t-test was used to find the differentially regulated transcripts in each window (where window refers to the timepoints spanned by group n-1 and group n). In this way, we were able to identify significant and DE genes across different windows along the time series. The time windows spanned to identify genes differentially regulated consisted of data collected between the following days along the time course: Days 0 – 12, Days 7 – 22, Days 15 – 25, Days 23 – 30, Days 26 – 45.

Analysis of Regulatory Events Involved in Increase of Target Gene Expression:

For mapping regulatory interactions and constructing GRNs across consecutive pairs of time windows, only the subset of genes reported as significant, DE with fold change > 0 (upregulated) was selected from each set of significant and DE genes. This was done in order to capture the temporal regulatory networks and emergent functionalization (in being reflected by enrichment of the specific subset of upregulated genes) across the time series as differentiation proceeds towards a neuronal phenotype.

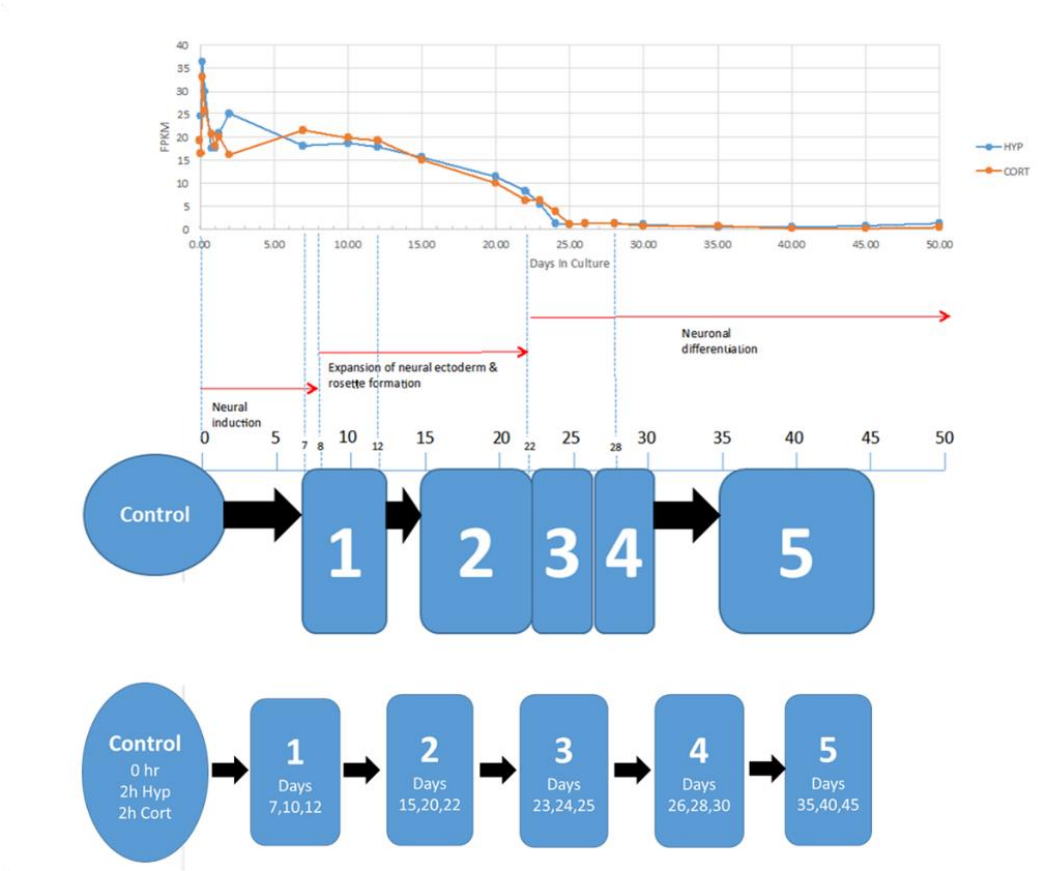

Figure 3: Grouping of timepoints and relation of timepoints/groups to the profile of REST and the number of days in culture. Cortical data (Orange), Hypothalamic data (Blue)

TF-Target Mapping across the Time Series:

REST target lists were assembled from TRANSFAC (Qiagen), the ChEA interaction database [20] and a recent study that expanded the REST cistrome in human ESCs [21]. A master TF-Target interaction list was assembled from the same sources, in addition to MSigDB [22].

Constructing Gene Regulatory Networks:

Gene regulatory networks were constructed by selecting all upregulated targets of REST from among the genes reported as significant and differentially expressed across consecutive time windows along the time series. Further expansion of these networks was done for a) upregulated targets of REST which were themselves transcription factors with upregulated targets in the immediate future time window and b) upregulated targets of the targets in (a) which are in turn transcription factors with upregulated targets in their immediate future time window.

This expansion of regulatory networks across the time course was done to generate GRNs that could be used to i) identify emergent neuronal processes that appear as a direct or downstream effect of REST ii) inspect upregulated transcriptional regulators, their upstream regulators, and the processes they upregulate across the time series (in both Hypothalamic and Cortical datasets).

Functional Enrichment Analyses:

GOrilla [23, 24] was used to generate rooted trees of enriched GO terms from gene sets of interest. Enriched GO terms with p-values $\leq 1 \times 10^{-3}$ were selected. The option to search for enriched GO terms in a target list of genes compared to a background list of genes was selected. The background list used for functional enrichment was the list of all genes measured in both datasets

(Hypothalamic and Cortical). GOrilla used a hypergeometric distribution to find the exact probabilities to compute enrichment likelihoods [25]:

$$P(\geq s) = \sum_{i=s}^{b} \frac{\binom{b}{i}\binom{N-b}{k-i}}{\binom{N}{k}}$$

where *b* is the number of 'background' genes annotated with the GO term/pathway, *s* is the number of 'selected' genes annotated with the term/pathway, *N* is the total number of 'background' genes and *k* is the total number of 'selected' genes.

Results were then exported to REVIGO [26]. REViGO takes lists of Gene Ontology terms as input and summarizes them by removing redundant terms. REVIGO was used to select high-level enriched GO Terms from the list of all GO Terms exported from Gorilla (similarity score cutoff = 0.7), and utilized the p-values input from GOrilla for assigning significance. Key terms enriched across the time series common to both datasets were selected and displayed in figures accompanying the analysis.

For functional annotation of smaller gene lists using relaxed p-value cutoffs (p-values $\leq 1 \times 10^{-2}$), ConsesusPathDB-human [27] was employed using our background gene list and selecting the option for over-representation analysis.

Comparative Analysis of GRNs across Cortical and Hypothalamic datasets:

For GRNs generated for Cortical and Hypothalamic datasets, summaries of enriched processes/significant genes common to both datasets was provided in the Figures and Tables accompanying the results. Qualitative comparisons across datasets, where done, were across corresponding time windows along the time course. Commonalities/differences in the temporal aspects of i) REST-controlled regulatory networks driving neural differentiation and ii) the functional contribution of the genes regulated by the above mentioned regulators to the emergence of neuronal phenotype were identified across datasets.

Identifying Protein-Level Interactions:

TFs under REST-control reported as common across datasets were uploaded to STRING-db [26] along with a REST-significance list consisting of members of the REST-repressive complex (REST, RCOR1, HDAC1) and beta-catenin (CTNNB1) in order to identify protein level interactions between regulatory genes and the REST-regulatory complex within the same transcriptomic state (static in time). Beta-catenin was included due to its importance as a downstream effector of Wnt-signaling (difference in protocols between datasets). A high confidence cutoff of 0.7 was used to select known protein interactions reported within our list of genes. This protein-level interaction identification was done in order to identify (based on existing

knowledge) any common protein regulators/complexes that could facilitate colocalization/coregulation events explaining the temporal patterns of regulation captured by our GRNs in the context of neural differentiation from iPSCs to i) Cortical and ii) Hypothalamic neurons.

CHAPTER 3: ANALYSIS AND RESULTS

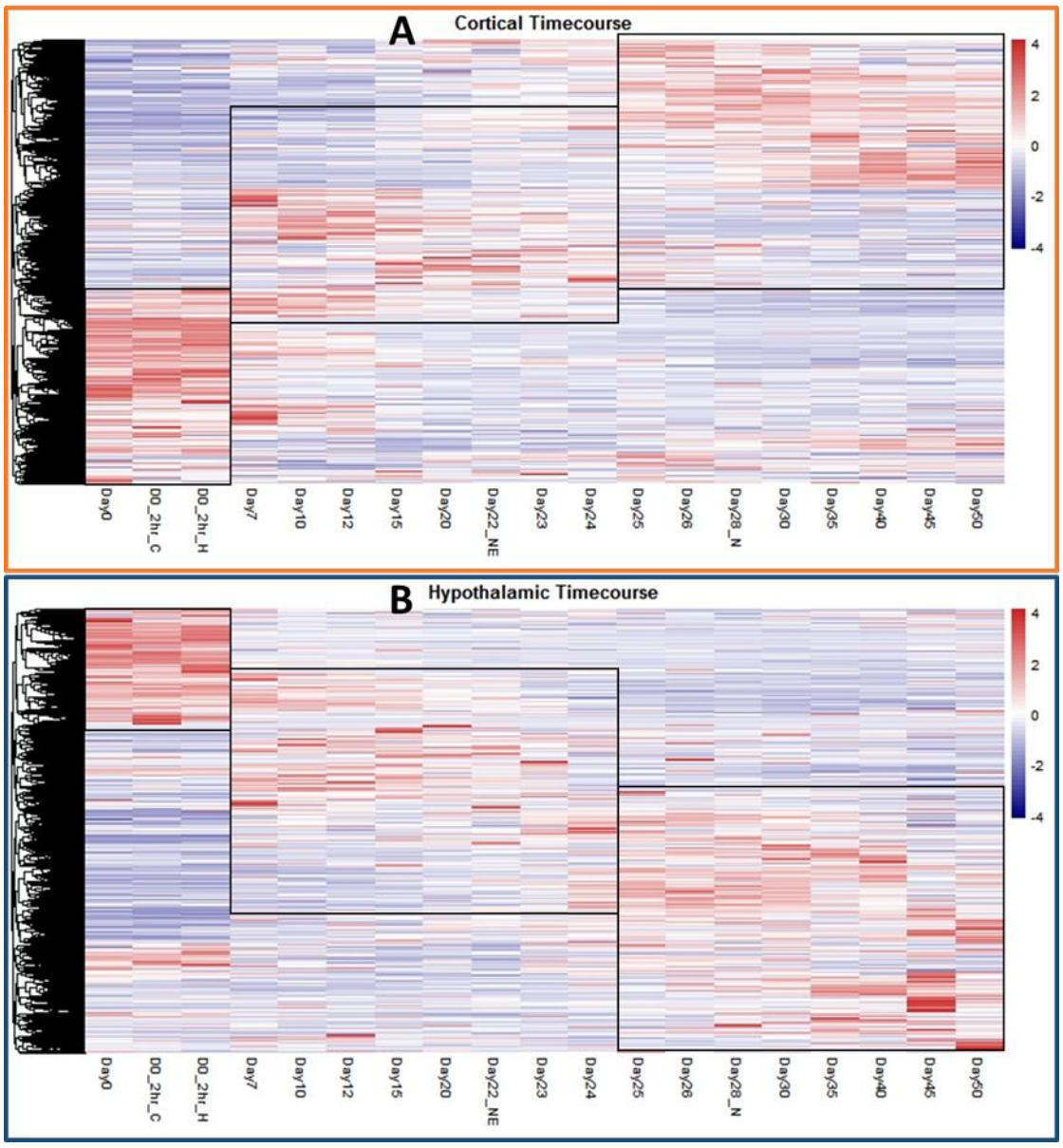

Figure 4: Dendrograms of gene expression profiles for both datasets for all expressed genes across the entire timecoure: A) Cortical dataset (outlined in orange) and B) Hypothalamic dataset (outlined in blue). Boxes outline gene clusters across the timecourse. Red indicates high expression, Blue indicates low expression. Both datasets showed a similar expression trend across the time series, with expression patterns across the developmental axis reflective of the iPSC, neuroepithelial and neuronal phenotypic states.

Visualization of Overall Gene Expression Across the Time Series:

Hierarchical clustering of normalized expression values for all expressed genes was carried out in order to visually inspect transcriptomic changes across the course of neurodevelopment in both datasets. Results are shown in Fig 4. Both datasets showed roughly 3 distinct clusters of genes across the time course (see Fig 4), and the distinct transcriptomic states represented by the clusters correspond to the 3 distinct phenotypic states of iPSC (early), neuroepithelial (mid) and neuronal (late) in their appearance along the time series.

Summary of DE Genes and REST Regulation Across Time:

Summary of genes reported as differentially expressed across the time series shown in Table 1 (both datasets). The Cortical time series shows a consistently higher number of upregulated REST targets in each stage compared to the Hypothalamic dataset, except between Days 15 – 25 (summarized in Table 1 A,B). Upregulated DE genes were used for construction of GRNs and mapping regulatory interactions to identify processes and genes emerging with time.

In order to generate GRNs under the control of REST by mapping regulatory interactions across the time series, we used the subset of upregulated targets of REST which were TFs themselves for the first 4 time

windows (Days 0 to 30) (Table 2). This selection of windows was based on the profile of REST in both datasets (Figure 3). There was a decrease in the number of TFs reported in lists of upregulated targets of REST across the time course in both datasets (Table 2 A,B).

Table 1: Number of DE Genes, Upregulated DE Genes and Upregulated REST targets across the entire time course. A) Cortical dataset (Orange) B) Hypothalamic dataset (Blue)

A	Cortical	Days 0 to 12	Days 7 to 22	Days 15 to 25	Days 23 to 30	Days 26 to 45
	Total DE genes reported	7896	6193	2675	4637	4124
	Upregulated DE genes	3927	3146	1131	2492	2259
	Upregulated REST targets	767	810	294	693	575
B	**Hypothalamic**	**Days 0 to 12**	**Days 7 to 22**	**Days 15 to 25**	**Days 23 to 30**	**Days 26 to 45**
	Total DE genes reported	9871	3309	4208	3389	2975
	Upregulated DE genes	4825	1729	1884	2014	1287
	Upregulated REST targets	677	450	466	475	316

Table 2: Number of Upregulated REST targets and Upregulated TFs across Days 0 – 30. A) Cortical dataset (Orange) B) Hypothalamic dataset (Blue)

A	Cortical	Days 0 to 12	Days 7 to 22	Days 15 to 25	Days 23 to 30
	Upregulated REST targets	767	810	294	693
	Upregulated TFs among REST targets	25	10	3	2
B	**Hypothalamic**	**Days 0 to 12**	**Days 7 to 22**	**Days 15 to 25**	**Days 23 to 30**
	Upregulated REST targets	677	450	466	475
	Upregulated TFs among REST targets	21	9	5	4

Functional Annotation of Upregulated DE Genes Across the Time Series:

Functional enrichment analysis was carried out on upregulated genes reported as significant in each window across the time course. This FA allowed us to qualitatively annotate the progression of cell fate commitment from iPSC towards Cortical and Hypothalamic neuron across time.

Functional annotation found to be common between datasets across are shown in Figure 5. Morphological changes and synaptic localization are seen to reflect key annotation across Days 0 - 22, followed by the emergence of annotation for Neurotransmitter Transport from days 15 – 45.

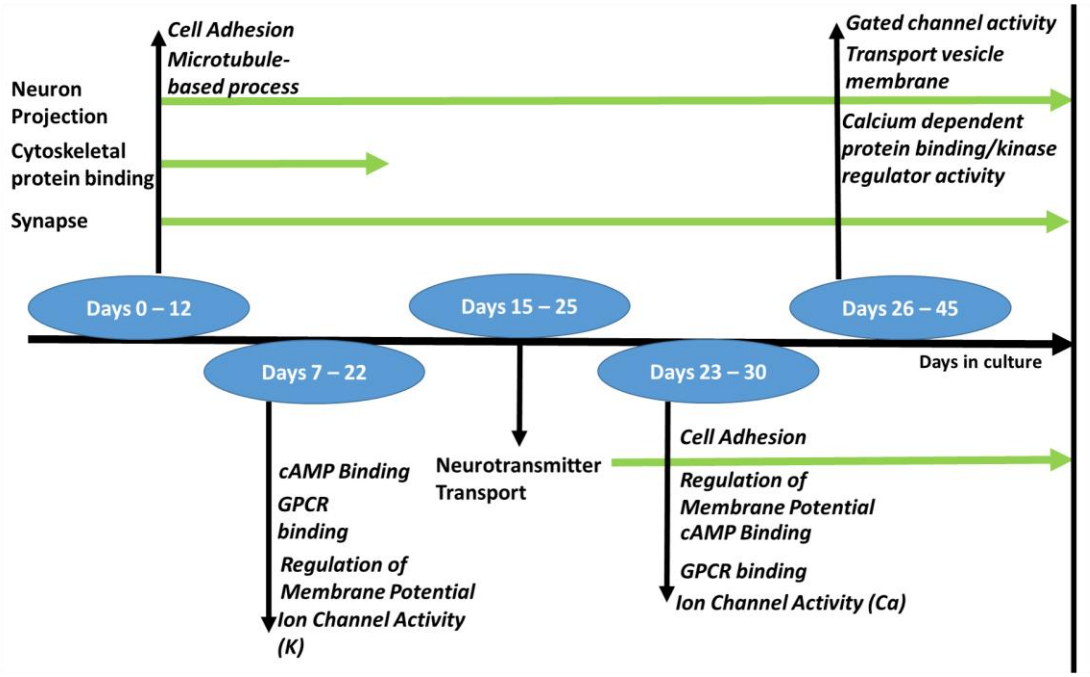

Figure 5: Functional annotations common to upregulated DE genes across Cortical and Hypothalamic datasets along the time series. Displayed terms appear at similar time windows across the time series in both Cortical and Hypothalamic data. Terms in bold persist across windows spanned by arrows. Italicized terms are specific to time window they are reported in.

In addition to the FAs listed in Figure 5 along the time series, there were annotations unique to Cortical and Hypothalamic datasets across the time windows (Figure 6, 7).

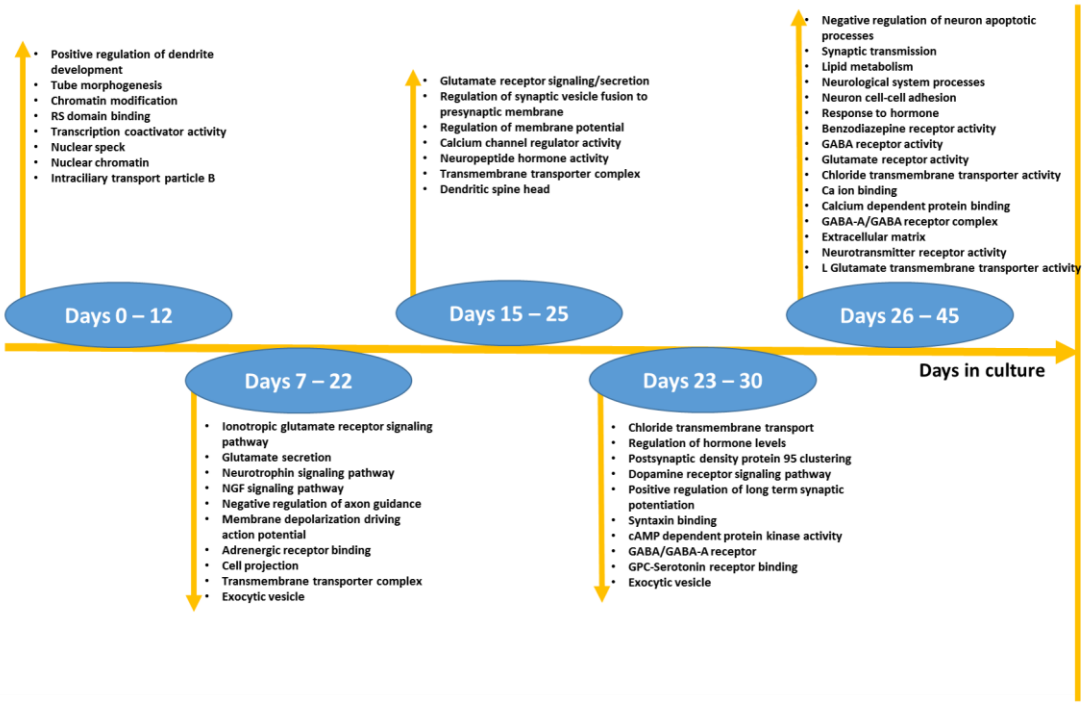

Figure 6: Dataset-specific functional annotations of upregulated DE genes from the Cortical dataset across the time course.

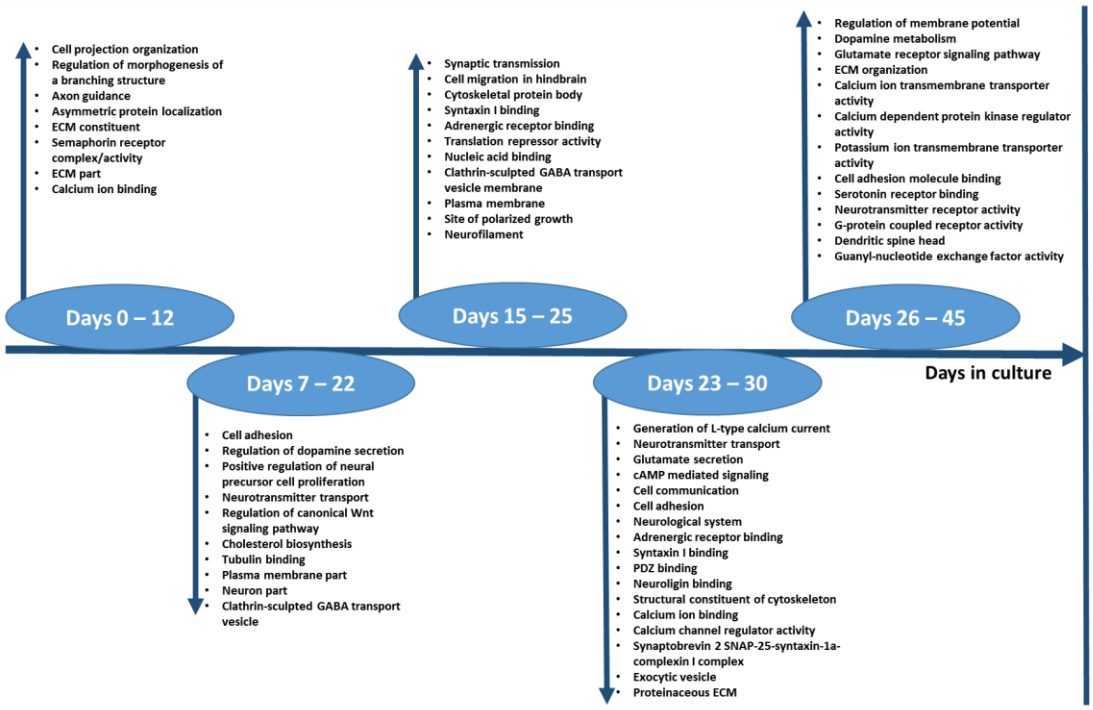

Figure 7: Dataset-specific functional annotations of upregulated DE genes from the Hypothalamic dataset across the time course.

Functional Annotation of Upregulated REST Targets Across the Time Series:

Functional enrichment analysis was carried out on upregulated REST target genes reported as significant in each window across the time course (Table 2). This FA allowed us to qualitatively annotate the direct control of REST in progression of cell fate commitment from iPSC towards Cortical and Hypothalamic neuron across time.

Annotated processes found to be common between datasets across are shown in Figure 8. In addition to the FAs listed in Figure 8 along the time series,

there were annotations unique to Cortical and Hypothalamic datasets across the time windows (Figures 9, 10).

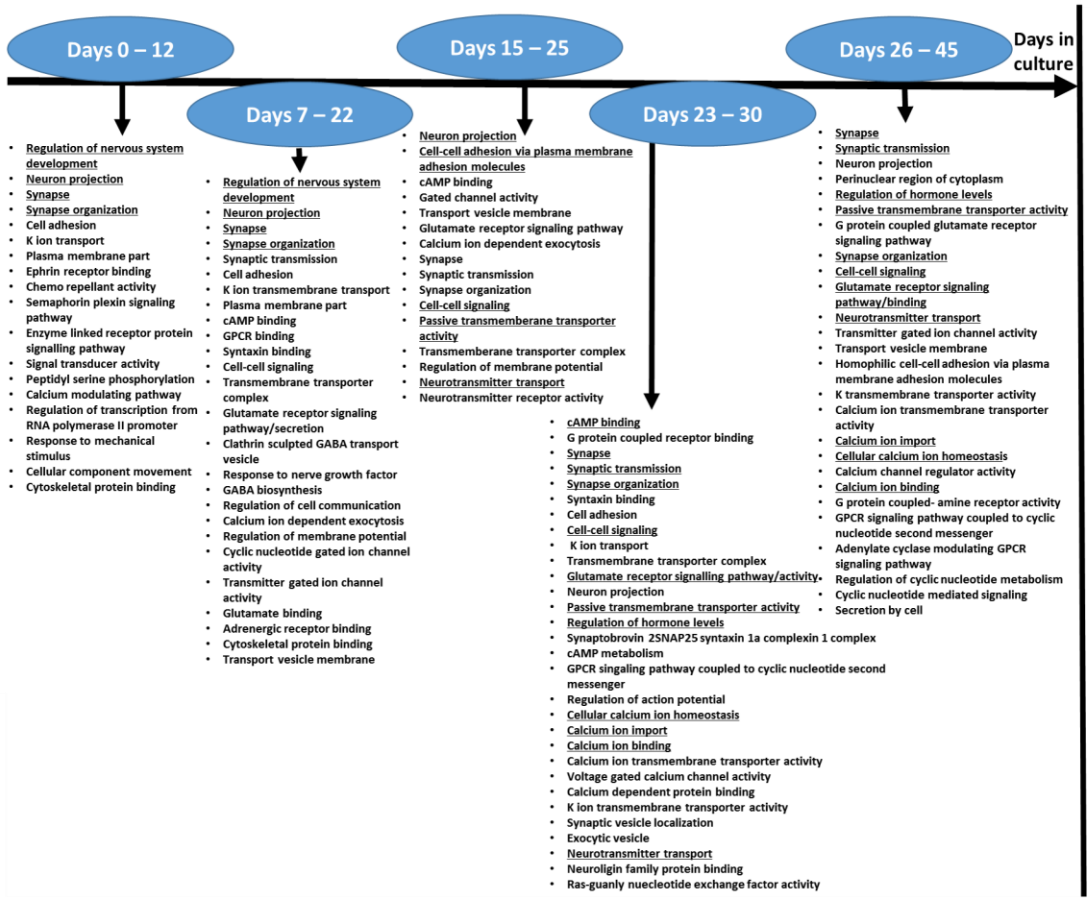

Figure 8: Functional annotations common to upregulated REST target genes across Cortical and Hypothalamic datasets along the time series. Terms common across consecutive windows are underlined.

FAs of REST target lists over windows spanning the time series showed emergent neuronal functionalization. Across the time series, the sequential emergence of FAs related to axon guidance, axonogenesis, synapse organization, cell-cell adhesion, ion homeostasis and synaptic transmission via neurotransmitters captured the functional role of direct targets of REST across time. Hormone regulation, GPCR binding, K and Ca ion transport and glutamate

receptor signaling are FA terms common to both Hypothalamic and Cortical datasets reported between Days 23 – 45.

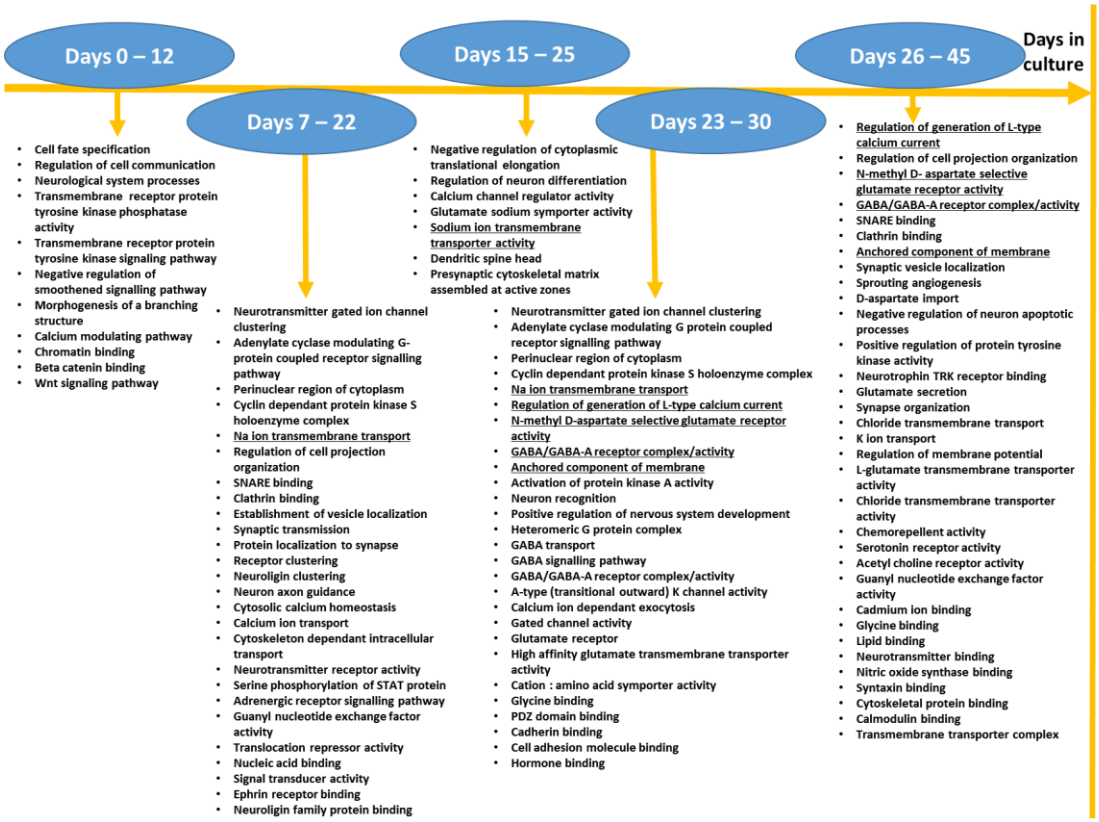

Figure 9: Dataset-specific functional annotations of upregulated REST targets from the Cortical dataset across the time course. Terms common across consecutive windows are underlined.

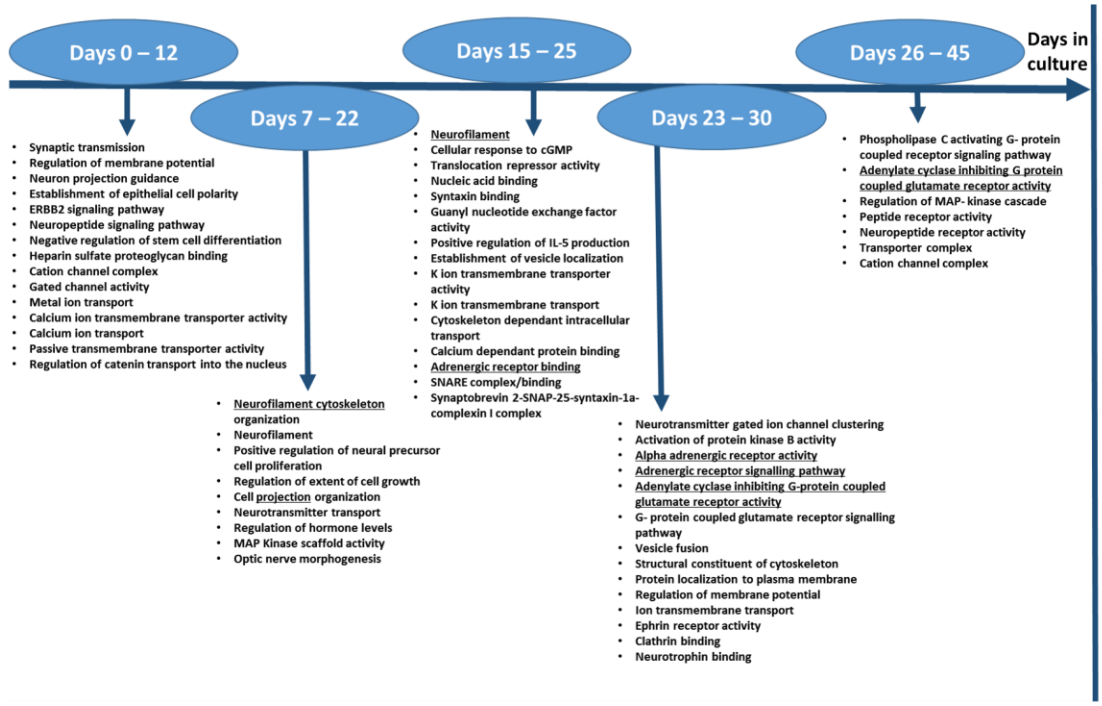

Figure 10: Dataset-specific functional annotations of upregulated REST targets from the Hypothalamic dataset across the time course. Terms common across consecutive windows are underlined.

Analysis of TF-Target Interactions for Upregulated TFs Among REST Targets:

For each window along the time course, the upregulated TFs of interest (which are REST targets) reported as significant and differentially expressed were selected and mapped to their subset of upregulated target genes that were reported as significant and DE in the immediate future window. Since REST is a known repressor with decreasing expression across the time series, REST repression causes upregulation (derepression) of target genes. This allowed us

to generate TF-Target maps spanning the time windows of significance across the time series.

Summaries of TFs reported among upregulated targets of REST in both datasets are presented in Tables 3 – 6.

Table 3: Summary of common upregulated TFs Days 0 - 12 (which are targets of REST) and their upregulated targets Days 7 – 22. A) Cortical dataset (Orange) B) Hypothalamic dataset (Blue). Upregulated targets common to both datasets are displayed in bold.

Upregulated TFs (REST target genes) Days 0 – 12	Days 7 – 22	
	Upregulated targets	Upregulated targets
CREM	626	271
LEF1	533	303
FOXP1	430	204
SMAD3	331	168
POU3F2	218	120
ASXL1	68	21
SALL1	10	4
PRDM16	15	9
RXRA	10	9
NFKB1	• **VCAM1** • **CRMP1** • **SLC1A2** • OLFM4 • SERPINB9	• **VCAM1** • **CRMP1** • **SLC1A2** • SOD2
EBF1	• **TH** • **STMN2**	• **TH** • **STMN2**
STAT5B	• MAF	• PIM1
SFPQ	• **TH**	• **TH**
NFIC	• JUN	• CBS • TFAP2A
E2F7	• **HCN3**	• **HCN3** • PIM1

Protein Interaction Networks Across the Time Series:

STRING-db was used to identify protein-level interactions between REST significance list (REST, RCOR1, HDAC1), beta-catenin (CTNNB1) and groups of TFs with biological/regulatory significance across the time series based on their order of emergence (window in which they are reported as significant, DE and upregulated). Results for TFs common to both datasets across time (reported in Tables 3 – 6) shown in Figures 11 – 14.

Table 4: Summary of common upregulated TFs Days 7 - 22 (which are targets of REST) and their upregulated targets Days 15 – 25. Cortical dataset (Orange), Hypothalamic dataset (Blue). Upregulated targets common to both datasets are displayed in bold.

Upregulated TFs (REST target genes) Days 7 – 22	Days 15 - 25	
	Upregulated targets	Upregulated targets
POU3F2	72	111
PRDM16	• **CAMK1D** • PSMD14 • NPAS2 • LRRC7	• **CAMK1D** • NALCN • ROBO2 • MAPRE3 • SPAG6
NR2F1	• CDKN1A	• CYP2D6
INSM1	• **INSM1**	• **INSM1**

Table 5: Summary of common upregulated TFs Days 15 - 25 (which are targets of REST) and their upregulated targets Days 23 – 30. A) Cortical dataset (Orange) B) Hypothalamic dataset (Blue)

Upregulated TFs (REST target genes) Days 15 - 25	Days 23 - 30	
	Upregulated targets	Upregulated targets
BCL11B	• **ADAM11** • **RUNDC3B** • RAPGEF1 • SH3RF3	• **ADAM11** • **RUNDC3B** • BCL11B • PDE4D

Table 6: Summary of common upregulated TFs Days 23 - 30 (which are targets of REST) and their upregulated targets Days 26 – 45. A) Cortical dataset (Orange) B) Hypothalamic dataset (Blue). Upregulated targets common to both datasets are displayed in bold.

Upregulated TFs (REST target genes) Days 23 - 30	Days 26 - 45	
	Upregulated targets	Upregulated targets
NR3C1	236	119

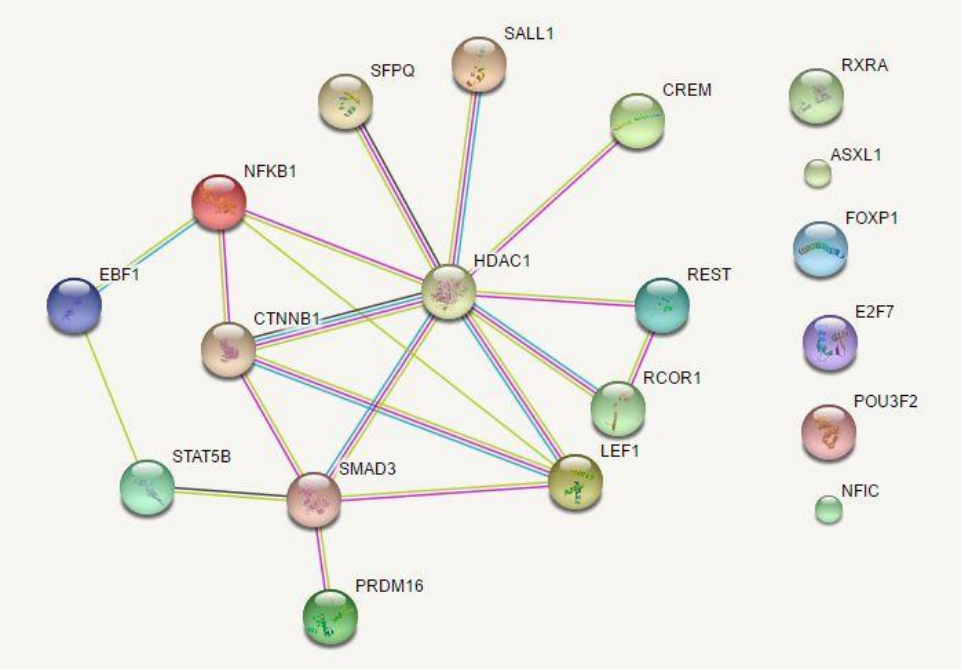

Figure 11: Protein interactions between upregulated TFs (Days 0 – 12) under REST control common to Cortical and Hypothalamic datasets. List of common TFs shown in Table 3.

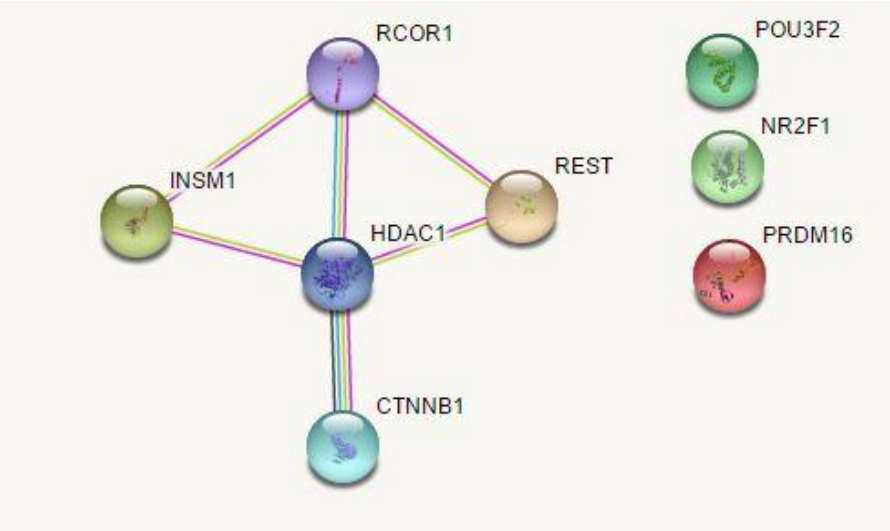

Figure 12: Protein interactions between upregulated TFs (Days 7 – 22) under REST control common to Cortical and Hypothalamic datasets. List of common TFs shown in Table 4.

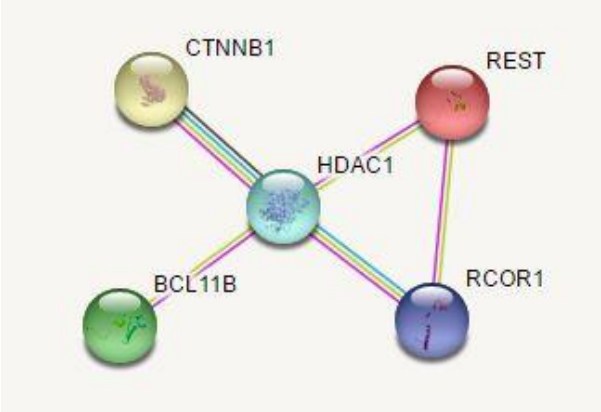

Figure 13: Protein interactions between upregulated TFs (Days 15 – 25) under REST control common to Cortical and Hypothalamic datasets. List of common TFs shown in Table 5.

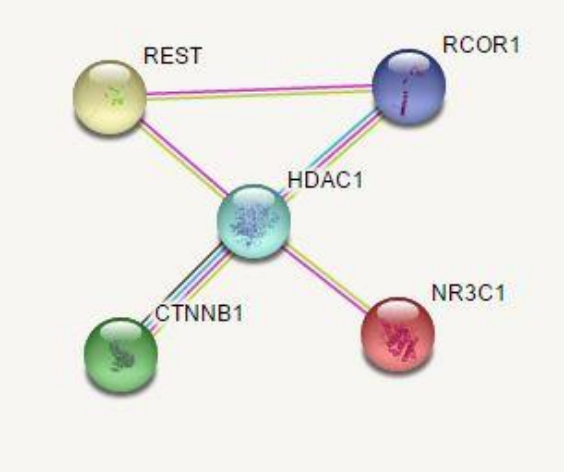

Figure 14: Protein interactions between upregulated TFs (Days 15 – 25) under REST control common to Cortical and Hypothalamic datasets. List of common TFs shown in Table 6.

PPIs consisting of TFs common to both datasets showed similar trends across the time series: TFs interacting with the REST repressive complex interact primarily through HDAC1 .

TFs unique to datasets across the time series were summarized based on window of significance as well as dataset they belong to. Summary shown in Table 7. Figures 15 and 16 contain information on upregulated targets of upregulated TFs across the time series in Cortical and Hypothalamic datasets respectively

.

Table 7: Summary of dataset-specific upregulated TFs across the time series (which are targets of REST). Upregulated targets of these TFs are in the immediate future window: A) Cortical dataset (Orange) B) Hypothalamic dataset (Blue)

A

Cortical	Days 0 to 12	Days 7 to 22	Days 15 to 25	Days 23 to 30
	• DCP1A • NR2F1 • BCL6 • TFDP1 • NR2C1 • LHX2 • INSM1	• GATA2 • RCOR3 • NR3C1 • BCLL1B • TRPS1	• NR1D1	• NR0B1

B

Hypothalamic	Days 0 to 12	Days 7 to 22	Days 15 to 25	Days 23 to 30
	• TCF4 • USF2 • CHD1L • GLI2 • HES1	• TCF7 • BCL3 • LHX2	• POU3F2 • RARA	• KLF5 • EBF1 • BCL11B

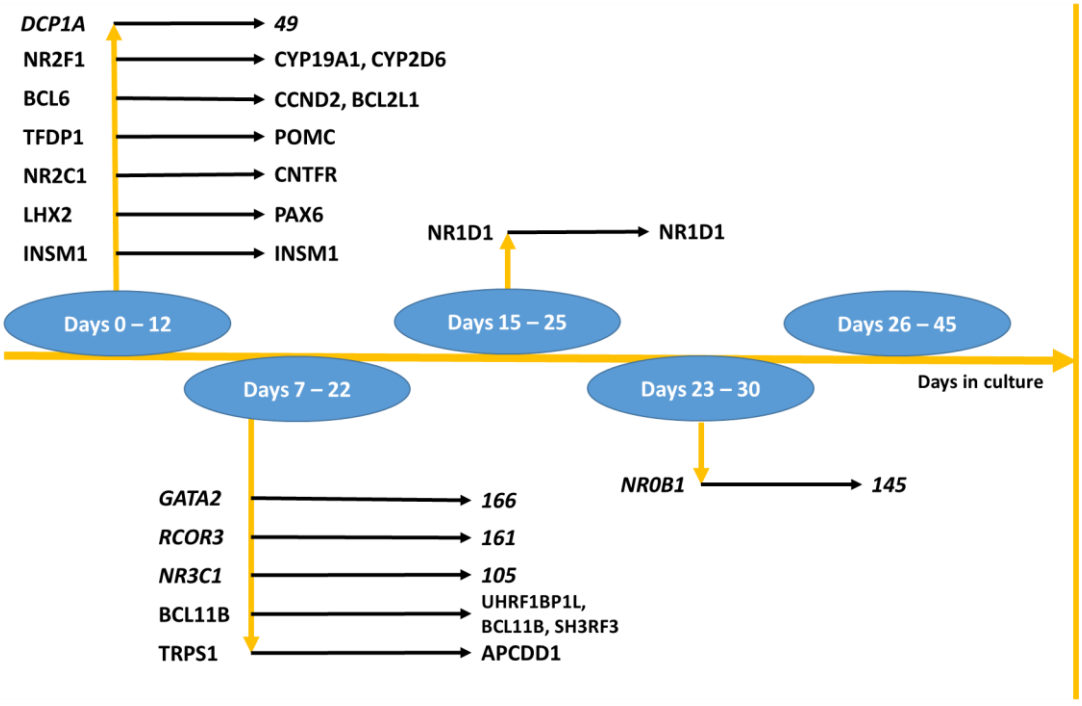

Figure 15: Summary of upregulated TFs unique to Cortical dataset across the time series.

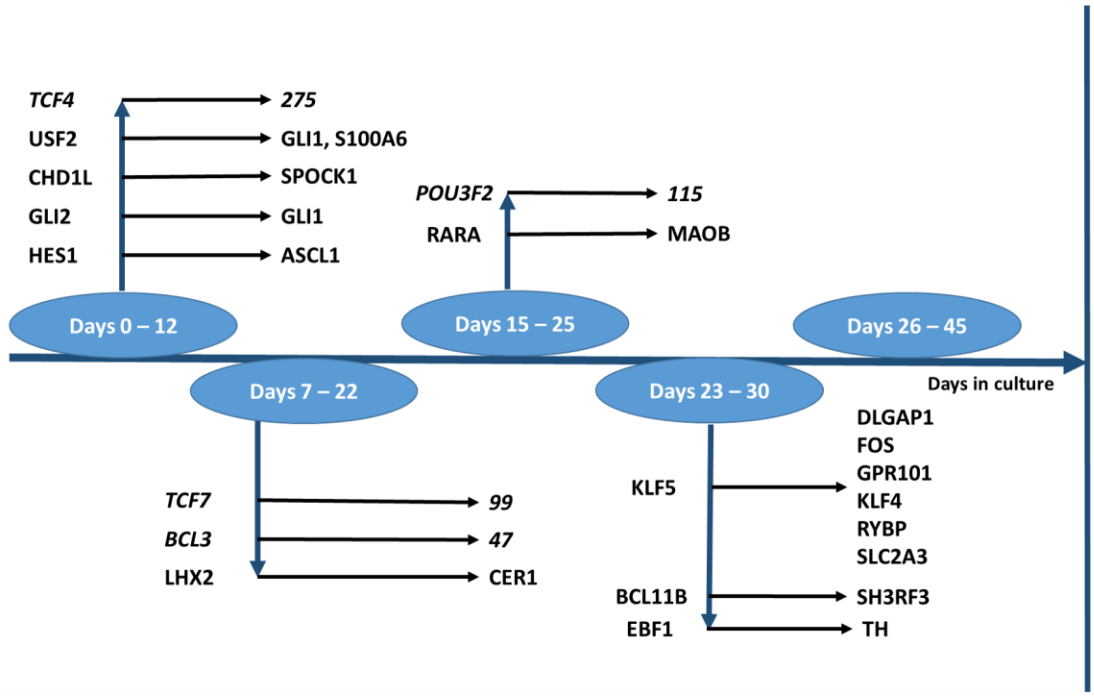

Figure 16: Summary of upregulated TFs unique to Hypothalamic dataset across the time series.

Dataset-Specific Protein Interaction Networks Across the Time Series:

PPI networks (Figures 15-18) for Cortical dataset specific TFs across the time series (from Table 7A) were generated using methods discussed earlier, and are displayed below:

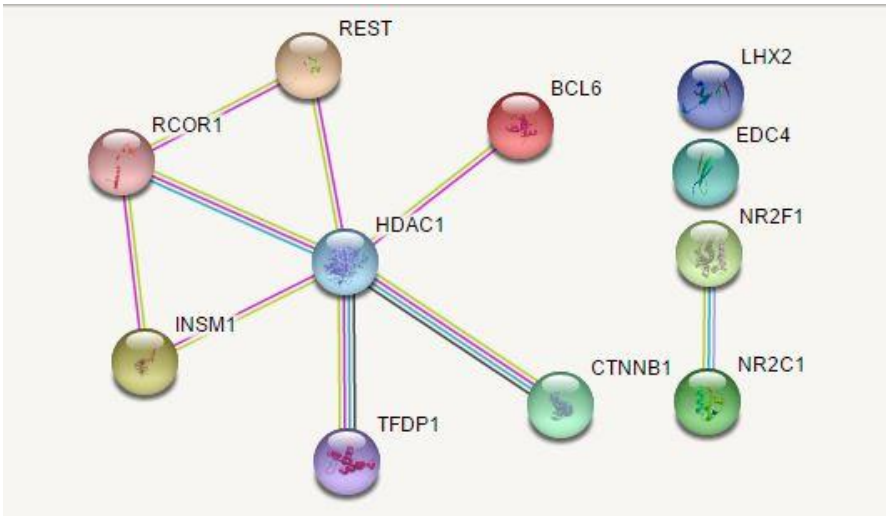

Figure 17: Protein interactions between upregulated TFs (Days 0 – 12) under REST control: Cortical dataset.

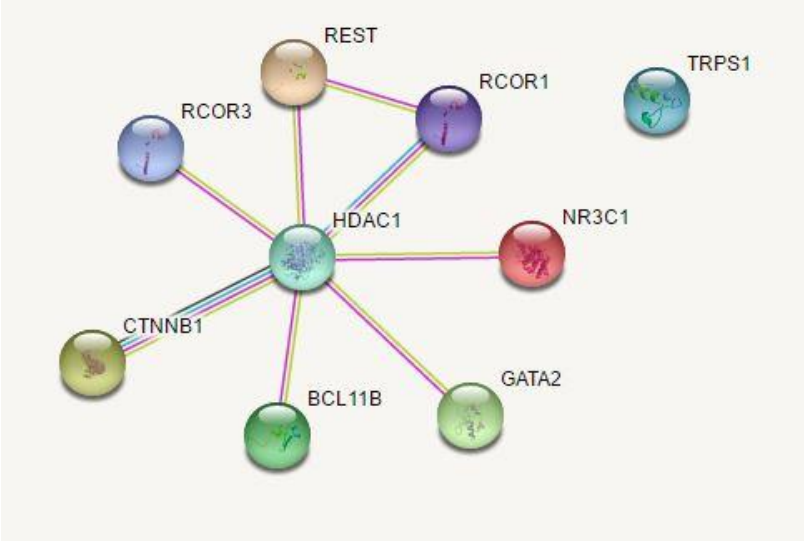

Figure 18: Protein interactions between upregulated TFs (Days 7 – 22) under REST control: Cortical dataset.

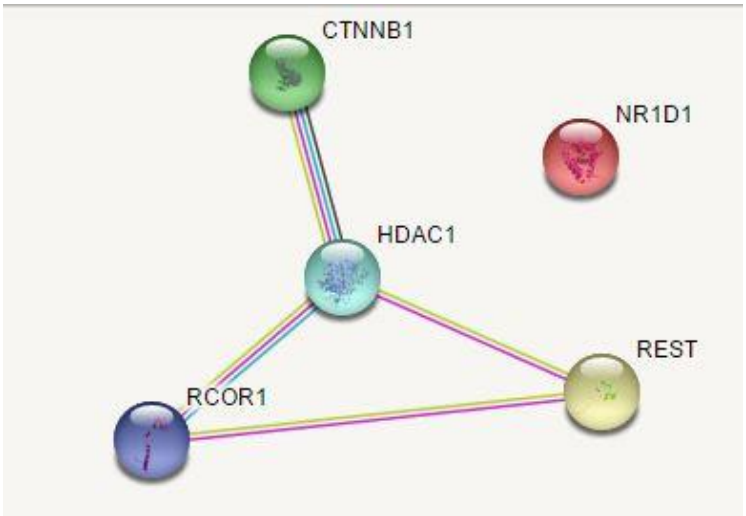

Figure 19: Protein interactions between upregulated TFs (Days 15 – 25) under REST control: Cortical dataset.

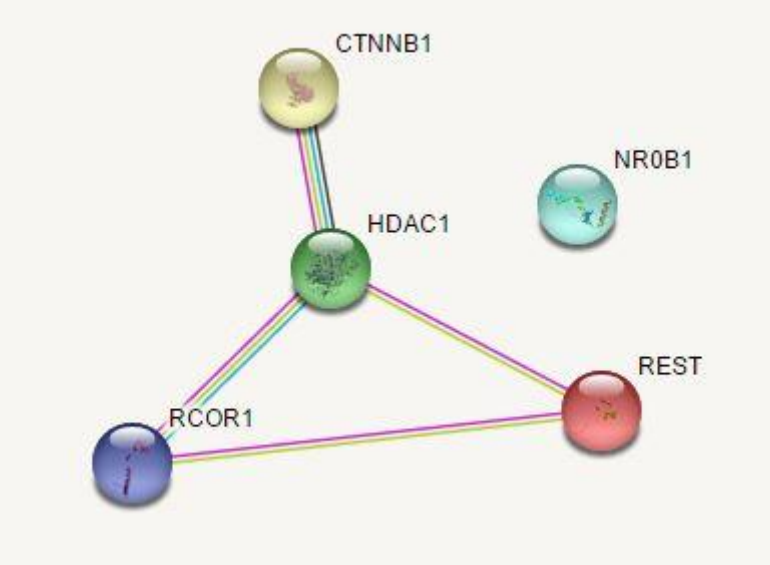

Figure 20: Protein interactions between upregulated TFs (Days 23 – 30) under REST control: Cortical dataset.

The Cortical PPIs generated using stage-specific TFs showed interaction with the REST-repressive complex through HDAC1, a trend seen in the case of TFs common to both datasets.

PPI networks (Figures 19-22) for Hypothalamic dataset-specific TFs across the time series (from Table 7B) were generated using methods discussed earlier, and are displayed below:

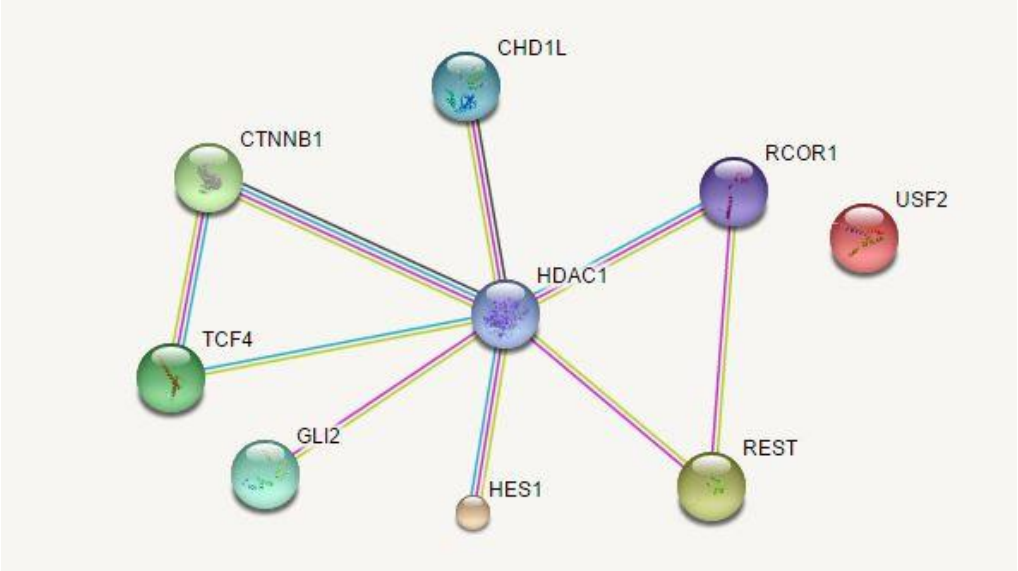

Figure 21: Protein interactions between upregulated TFs (Days 0 – 12) under REST control: Hypothalamic dataset.

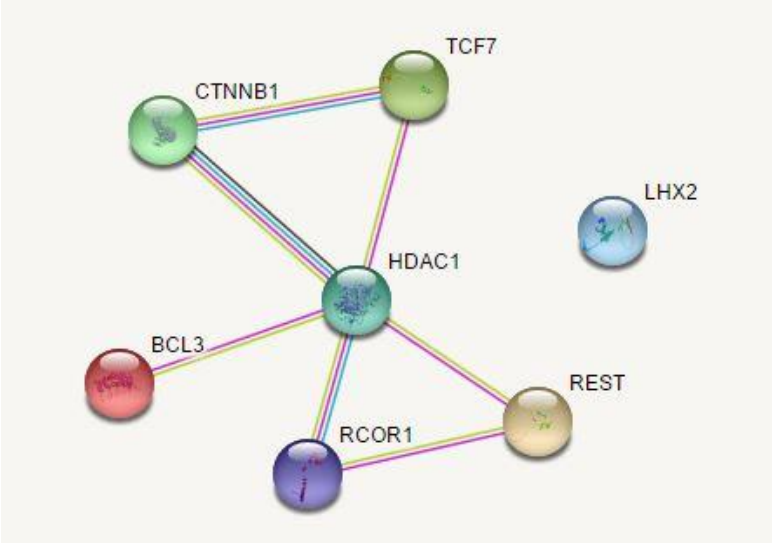

Figure 22: Protein interactions between upregulated TFs (Days 7 – 22) under REST control: Hypothalamic dataset.

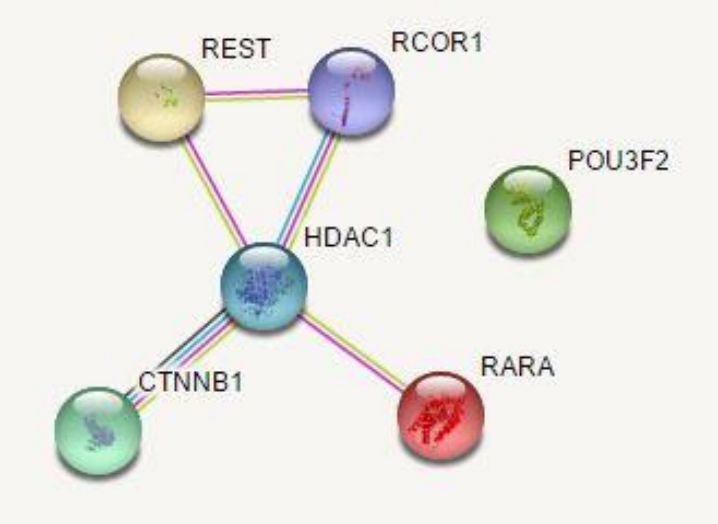

Figure 23: Protein interactions between upregulated TFs (Days 15 – 25) under REST control: Hypothalamic dataset.

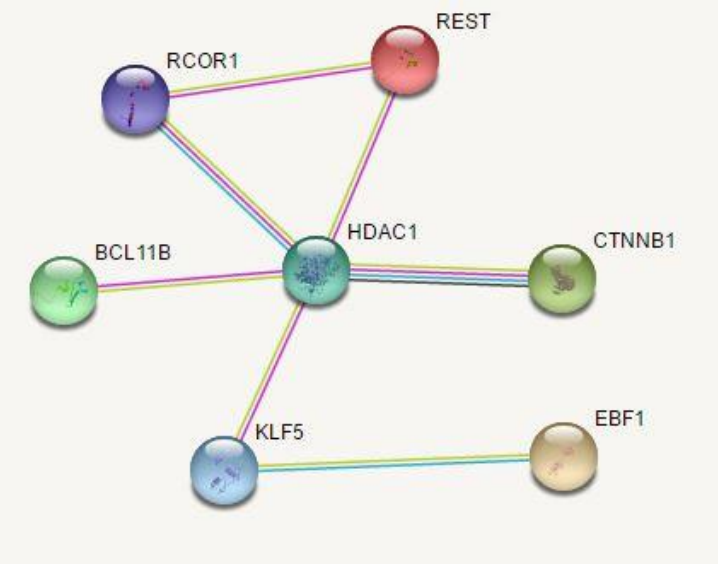

Figure 24: Protein interactions between upregulated TFs (Days 23 – 30) under REST control: Hypothalamic dataset.

The Hypothalamic PPIs generated using stage-specific TFs showed interaction with the REST-repressive complex through HDAC1, a trend similar to what was observed in the case of TFs specific to the Cortical dataset, as well as in the case of TFs common to both datasets.

Analysis of TF-Target interactions for targets of REST-controlled TFs which are TFs themselves with upregulated targets reported in their immediate future window:

Summaries of interactions with time windows and corresponding lists of upstream REST-controlled regulators presented in Figures 23-26. Interactions captured the number of upregulated targets controlled by TFs which are targets of REST-controlled TFs from previous windows along the time series (shown in column listing “Direct REST targets”, Tables 8-11).

Table 8: Regulatory interactions between upregulated TFs (Days 7 – 22) and their upregulated targets (Days 15 – 25): Cortical dataset. Genes common to both datasets highlighted in yellow. Genes in bold are also direct targets of REST.

Upregulated TFs (Direct REST targets) Days 0 – 12	Upregulated targets (TFs) Days 7 - 22	Upregulated targets (of Day 7 – 22 TFs) Days 15 – 25
LEF1, SMAD3, CREM	**GATA2**	166
FOXP1, CREM	**RCOR3**	161
SMAD3, POU3F2, PRDM16	**NR3C1**	105
SMAD3, FOXP1, NFIC	JUN	70
LEF1, POU3F2, CREM	PHC1	68
SMAD3, FOXP1	TFAP2C	64
LEF1, SMAD3	EOMES	64
LEF1, SMAD3, FOXP1	MEIS1	44
LEF1	DACH1	40
FOXP1	AUTS2	38
LEF1, SMAD3, LHX2	PAX6	23
LEF1, POU3F2	NFIB	20
LEF1, SMAD3	NOTCH1	17
SMAD3	THRA	10
LEF1, FOXP1	**BCL11B**	• UHRF1BP1L • BCL11B • SH3RF3
LEF1	USF1	• PLA2G4C • SPP1
LEF1, CREM	ZBTB17	• CDKN1A
POU3F2	**TRPS1**	• APCDD1
SMAD3, CREM	SREBF1	• CDKN1A
LEF1	**NR2F1**	• CDKN1A
FOXP1, CREM	MZF1	• PRKCA
LEF1, SMAD3, POU3F2, STAT5B	MAF	• GCLM
LEF1, INSM1	**INSM1**	• INSM1
LEF1, SMAD3, FOXP1	ARX	• LMO1

Table 9: Regulatory interactions between upregulated TFs (Days 15 – 25) and their upregulated targets (Days 23 – 30): Cortical dataset. Genes common to both datasets highlighted in yellow. Genes in bold are also direct targets of REST.

Upregulated targets (TFs) Days 7 – 22	Upregulated targets (TFs) Days 15 - 25	Upregulated targets (of Day 15 – 25 TFs) Days 23 – 30
GATA2, **RCOR3**, PHC1, AUTS2, THRA	LMO2	145
GATA2 , TFAP2C, AUTS2, NR3C1, **POU3F2**	NFIB	47
PHC1	GBX2	45
PAX6, NFIB, JUN, POU3F2	NR4A2	27
RCOR3	THRA	19
PHC1, TFAP2C, **BCL11B**	**BCL11B**	• RAPGEF1 • RUNDC3B • SH3RF3 • ADAM11
RCOR3, MEIS1	**NR1D1**	• NR1D1
Direct REST targets (TFs) Days 7 - 22		
POU3F2	DLX5	• MYC

Table 10: Regulatory interactions between upregulated TFs (Days 7 – 22) and their upregulated targets (Days 15 – 25): Hypothalamic dataset. Genes common to both datasets highlighted in yellow. Genes in bold are direct targets of REST.

Upregulated TFs (Direct REST targets) Days 0 – 12	Upregulated targets (TFs) Days 7 - 22	Upregulated targets (of Day 7 – 22 TFs) Days 15 – 25
TCF4	OLIG2	190
LEF1, SMAD3	EOMES	156
LEF1	**TCF7**	99
LEF1, FOXP1, SMAD3, NFIC	TFAP2A	89
LEF1, FOXP1, SMAD3, TCF4	MEIS1	74
FOXP1, TCF4	AUTS2	45
LEF1, POU3F2, TCF4	NFIB	36
LEF1, SMAD3	NOTCH1	20
LEF1	ZIC3	18
LEF1, FOXP1, GLI2, USF2	GLI1	• CDH7 • KCND3 • NCAM1 • ROBO2
LEF1	**NR2F1**	• CYP2D6
LEF1, CREM	**LHX2**	• CER1
LEF1	**INSM1**	• INSM1
LEF1, TCF4	ESRRG	• ESRRA

Table 11: Regulatory interactions between upregulated TFs (Days 15 – 25) and their upregulated targets (Days 23 – 30): Hypothalamic dataset. Genes common to both datasets highlighted in yellow. Genes in bold are direct targets of REST.

Upregulated targets (TFs) Days 7 – 22	Upregulated targets (TFs) Days 15 - 25	Upregulated targets (of Day 15 – 25 TFs) Days 23 – 30
TCF7, MEIS1	KDM5B	154
OLIG2	PBX1	108
AUTS2	LMO2	94
OLIG2, EOMES	MEIS1	66
TCF7, TFAP2A	THRA	15
TCF7	**BCL11B**	• BCL11B • PDE4D • RUNDC3B • ADAM11
TCF7	USF1	• ID4 • PLA2G4C
TCF7, TFAP2A	RARA	• MAOB

DISCUSSION:

Functional Annotation Patterns Across the Time Series for Upregulated DE Genes:

The developmental progression of iPSCs towards neuronal phenotypes in the case of both Hypothalamic and Cortical protocols was captured across time (Figure 5). In terms of functional annotations common to both datasets across the time series for lists of upregulated DE genes, neuronal outgrowth and establishment of neuronal morphology emerged first (between Days 0 – 12) and persisted through the time course. Developmentally, this was followed by the emergence of FAs related to GPCR binding, cAMP binding, Potassium ion channel activity and regulation of membrane potential (between Days 7 – 22). Neurotransmitter transport appeared in the next time window (between Days 15 – 25) and this was followed by the emergence of FAs related to the regulation membrane potential and Calcium ion transport/activity (between Days 23 – 45), indicating the functionalization of synapses formed by the differentiating neurons.

In the FAs unique to the Cortical dataset across the time series (Fig 6), annotation indicative of chromatin modification and transcriptional coactivator activity were reported first (between Days 0 – 12). This was followed by emergence of glutamate/adrenergic receptor binding and firing of action potentials (between Days 7 – 22). Neurotrophin/NGF signaling, indicative of neuronal survival, also appeared in the same time window and persisted across

the subsequent window as well (Days 15 – 25). Synaptic potentiation, dopamine receptor signaling pathway, serotonin receptor binding and hormone sensitivity appear later in the time course (between Days 23 – 30). Glutamate/GABA receptor activity, synaptic transmission, Calcium ion regulation and neuron specific cell-cell adhesion emerged last across the time series (between Days 26 – 45), indicating of a class of neurons expressing Glutamate, GABA and Serotonin receptors.

In the FAs unique to the Hypothalamic dataset across the time series (Figure 7), Semaphorin receptor activity involved in axon guidance appeared first (between Days 0 – 12), followed by regulation of Wnt signaling, positive regulation of neural precursor cell proliferation, GABA transport and dopamine secretion in the time window that follows (between Days 7 – 22). Synaptic transmission, adrenergic receptor binding and neurofilament (a component of the neuronal cytoskeleton) appeared next (between Days 15 – 25). Glutamate secretion and neuroligin appeared in the subsequent time window (between Days 23 – 30), indicating synaptic maintenance and synaptic activity. Serotonin/glutamate receptor activity, dopamine metabolism and cell adhesion molecule binding emerged last across the time series (between Days 26 – 45), indicating of a class of neurons expressing Glutamate, Dopamine and Serotonin receptors.

Functional Annotation Patterns Across the Time Series for Upregulated REST Target Genes:

The role of REST in regulating developmental progression of iPSCs towards neuronal phenotypes in the case of both Hypothalamic and Cortical protocols was captured across time (Figure 8). In terms of functional annotations common to both datasets across the time series for lists of upregulated REST target genes, axonal guidance (Semaphorin-plexin signaling pathway, Ephrin receptor binding) and synapse organization emerged first (between Days 0 – 12). Adrenergic receptor binding, glutamate binding and GABA transport appear next (between Days 7 – 22), in addition to cAMP/GPCR binding, Potassium ion transport and synaptic transmission. This could explain the FAs seen in the corresponding stage of the common FA for upregulated DE genes discussed in the previous section. Neurotransmitter transport appeared in the next time window (Day 15 – 25), which is the same window in which the same term is enriched in FA for all upregulated DE genes. Calcium ion homeostasis, response to hormone, neuroligin binding, cAMP/GPCR binding, glutamate receptor signaling and synaptic transmission emerged next (between Days 23 – 30), once again explaining the overall functionalization of the differentiating neurons in both datasets across this time window. Most terms reported between Days 23 – 30 were seen between Days 26 – 45 as well: this can be explained by the overall profile of REST in both datasets across the time series (Figure 3). The expression of REST decreases over the course of time

until it reaches a minimum around Day 25. Derepressed target genes would therefore remain derepressed after Day 25, resulting in persistence of FA seen between Days 23 – 30 over the future window (Days 26 – 45).

In the FAs unique to upregulated targets of REST in the Cortical dataset across the time series (Fig 9), annotation indicative of chromatin binding and Wnt signaling/Beta catenin binding were reported first (between Days 0 – 12). This was followed by emergence of adrenergic receptor binding and synaptic transmission (between Days 7 – 22). Sodium ion transmembrane transporter activity also appeared in the same time window and persisted across the following time windows as well (Days 15 – 30). GABA receptor activity and Glycine binding appear later in the time course (between Days 23 – 30). Most of these terms persisted to the next time window (between Days 26 – 45), with the appearance of Neurotrophin TRK receptor binding and synapse organization/chemorepellent activity.

In the FAs unique to upregulated targets of REST in the Hypothalamic dataset across the time series (Figure 10), ERBB2 and neurotrophin signaling pathways appeared first (between Days 0 – 12), followed by neurofilament organization, positive regulation of neural precursor cell proliferation, neurotransmitter transport and regulation of hormone levels in the time window that follows (between Days 7 – 22). Synaptic vesicle transport, adrenergic receptor binding and neurofilament (a component of the neuronal cytoskeleton) appeared next (between Days 15 – 25). Adrenergic receptor activity persisted

to the subsequent time window (between Days 23 – 30), with the emergence of Ephrin receptor activity involved in axonal guidance and neurotrophin binding, which promotes neuronal cell survival. Phospholipase C activating GPCR signaling, regulation of MAPK cascade and neuropeptide receptor activity emerged last across the time series (between Days 26 – 45).

Emergent functionalization captured across the time series through sets of upregulated REST targets could explain the trends reflected by the overall DE genelists across different stages along neurodifferentiation. This highlights the importance of REST in influencing neuronal development, and presents a strong case for further functional elucidation of GRNs consisting of regulatory interactions of REST-related transcriptional regulators and their targets.

PPI Networks:

The PPI networks generated (Figures 11 – 14, 17 – 24) showing interactions between the members of the REST-repressive complex, beta-catenin and the TFs of significance (Tables 3 – 7) showed interactions of REST-target genes which are TFs themselves with the REST-repressive complex primarily through HDAC1, a repressive histone deacetylase. Preliminary inspection of FAs for upregulated targets of these TFs of significance (results

not shown) shows their significant contribution in upregulating genes indicative of a neuronal phenotype. Interactions driving recruitment of TFs to promoter regions of neuronal target genes through HDAC1 could be an interesting hypothesis regarding the concerted molecular mechanisms driving REST regulation of neural differentiation.

Future Work:

Further annotation of GRNs with FAs with temporal information for TF-Target lists as being under direct or indirect control of REST would help elucidate regulatory networks (TF-target interactions across time) driving neuronal functionalization across the time series. The potential of HDAC1 in recruiting TFs activating repressed neuronal gene expression in a time and context dependent manner to REST-repressive complex bound promoters seems to have potential for further elucidation. Such a mechanism involving protein interactions between HDAC1, beta-catenin and LEF1 (LEF1 is reported in Table 3 as a REST-target with TF function common to both datasets) has been reported in literature [28], and involves switching transcription of target genes on/off depending on the cellular context.

REFERENCES

1. Hu, B.Y., J.P. Weick, J. Yu, L.X. Ma, X.Q. Zhang, J.A. Thomson, and S.C. Zhang, *Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency.* Proc Natl Acad Sci U S A, 2010. **107**(9): p. 4335-40.
2. Choi, J., S. Lee, W. Mallard, K. Clement, G.M. Tagliazucchi, H. Lim, I.Y. Choi, F. Ferrari, A.M. Tsankov, R. Pop, G. Lee, J.L. Rinn, A. Meissner, P.J. Park, and K. Hochedlinger, *A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs.* Nat Biotechnol, 2015. **33**(11): p. 1173-81.
3. Zhang, Y., C. Pak, Y. Han, H. Ahlenius, Z. Zhang, S. Chanda, S. Marro, C. Patzke, C. Acuna, J. Covy, W. Xu, N. Yang, T. Danko, L. Chen, M. Wernig, and T.C. Sudhof, *Rapid single-step induction of functional neurons from human pluripotent stem cells.* Neuron, 2013. **78**(5): p. 785-98.
4. Cox, W.G. and A. Hemmati-Brivanlou, *Caudalization of neural fate by tissue recombination and bFGF.* Development, 1995. **121**(12): p. 4349-58.
5. Wills, A.E., V.M. Choi, M.J. Bennett, M.K. Khokha, and R.M. Harland, *BMP antagonists and FGF signaling contribute to different domains of the neural plate in Xenopus.* Dev Biol, 2010. **337**(2): p. 335-50.
6. Qureshi, I.A., S. Gokhan, and M.F. Mehler, *REST and CoREST are transcriptional and epigenetic regulators of seminal neural fate decisions.* Cell Cycle, 2010. **9**(22): p. 4477-86.
7. Leung, A.W., B. Murdoch, A.F. Salem, M.S. Prasad, G.A. Gomez, and M.I. Garcia-Castro, *WNT/beta-catenin signaling mediates human neural crest induction via a pre-neural border intermediate.* Development, 2016. **143**(3): p. 398-410.
8. Mertens, J., M.C. Marchetto, C. Bardy, and F.H. Gage, *Evaluating cell reprogramming, differentiation and conversion technologies in neuroscience.* Nat Rev Neurosci, 2016.
9. Imayoshi, I. and R. Kageyama, *bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells.* Neuron, 2014. **82**(1): p. 9-23.
10. Chanda, S., C.E. Ang, J. Davila, C. Pak, M. Mall, Q.Y. Lee, H. Ahlenius, S.W. Jung, T.C. Sudhof, and M. Wernig, *Generation of induced neuronal cells by the single reprogramming factor ASCL1.* Stem Cell Reports, 2014. **3**(2): p. 282-96.
11. Chong, J.A., J. Tapia-Ramirez, S. Kim, J.J. Toledo-Aral, Y. Zheng, M.C. Boutros, Y.M. Altshuller, M.A. Frohman, S.D. Kraner, and G. Mandel, *REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons.* Cell, 1995. **80**(6): p. 949-57.

12. Yoo, A.S., A.X. Sun, L. Li, A. Shcheglovitov, T. Portmann, Y. Li, C. Lee-Messer, R.E. Dolmetsch, R.W. Tsien, and G.R. Crabtree, *MicroRNA-mediated conversion of human fibroblasts to neurons.* Nature, 2011. **476**(7359): p. 228-31.
13. Baldelli, P. and J. Meldolesi, *The Transcription Repressor REST in Adult Neurons: Physiology, Pathology, and Diseases(1,2,3).* eNeuro, 2015. **2**(4).
14. Nechiporuk, T., J. McGann, K. Mullendorff, J. Hsieh, W. Wurst, T. Floss, and G. Mandel, *The REST remodeling complex protects genomic integrity during embryonic neurogenesis.* Elife, 2016. **5**: p. e09584.
15. Jorgensen, H.F., Z.F. Chen, M. Merkenschlager, and A.G. Fisher, *Is REST required for ESC pluripotency?* Nature, 2009. **457**(7233): p. E4-5; discussion E7.
16. Huang, S., G. Eichler, Y. Bar-Yam, and D.E. Ingber, *Cell fates as high-dimensional attractor states of a complex gene regulatory network.* Phys Rev Lett, 2005. **94**(12): p. 128701.
17. Huang, S., I. Ernberg, and S. Kauffman, *Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective.* Semin Cell Dev Biol, 2009. **20**(7): p. 869-76.
18. Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn, and L. Pachter, *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.
19. Baldi, P. and A.D. Long, *A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes.* Bioinformatics, 2001. **17**(6): p. 509-19.
20. Lachmann, A., H. Xu, J. Krishnan, S.I. Berger, A.R. Mazloom, and A. Ma'ayan, *ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments.* Bioinformatics, 2010. **26**(19): p. 2438-44.
21. Rockowitz, S. and D. Zheng, *Significant expansion of the REST/NRSF cistrome in human versus mouse embryonic stem cells: potential implications for neural development.* Nucleic Acids Res, 2015. **43**(12): p. 5730-43.
22. Subramanian, A., P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, and J.P. Mesirov, *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.
23. Eden, E., R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini, *GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists.* BMC Bioinformatics, 2009. **10**: p. 48.

24. Eden, E., D. Lipson, S. Yogev, and Z. Yakhini, *Discovering motifs in ranked lists of DNA sequences.* PLoS Comput Biol, 2007. **3**(3): p. e39.
25. Hsiao, A., T. Ideker, J.M. Olefsky, and S. Subramaniam, *VAMPIRE microarray suite: a web-based platform for the interpretation of gene expression data.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W627-32.
26. Supek, F., M. Bosnjak, N. Skunca, and T. Smuc, *REVIGO summarizes and visualizes long lists of gene ontology terms.* PLoS One, 2011. **6**(7): p. e21800.
27. Kamburov, A., U. Stelzl, H. Lehrach, and R. Herwig, *The ConsensusPathDB interaction database: 2013 update.* Nucleic Acids Res, 2013. **41**(Database issue): p. D793-800.
28. Billin, A.N., H. Thirlwell, and D.E. Ayer, *Beta-catenin-histone deacetylase interactions regulate the transition of LEF1 from a transcriptional repressor to an activator.* Mol Cell Biol, 2000. **20**(18): p. 6882-90.