Exploring Genes Associated with Memory and Learning in the Thalamus

Isabel Joyce, Newton North High School, Newton, MA 02460 & BioScience Project, Wakefield, MA, 01880

Introduction

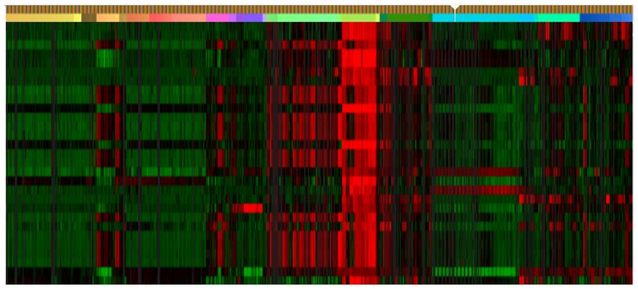
- The Thalamus, a small structure located in the brain, relays sensory messages from other parts of the body to the cerebral cortex, making the structure vital for a person's perception.
- Symptoms of loss of perception come in learning or memory, conduct disorder, ADHD, and response behaviors.
- The aim of the project is to identify candidate genes associated with Thalamus mediated learning and memory. (Thalamic control of human attention driven by memory and learning. de Bourbon-Teles J, et al. 2014)

Methods

Gene expression data for the Thalamus was collected from the The Allen Brain Atlas (http://www.brain-map.org) using the differential search option. Data was collected from four available donors: H0351.2001, H0351.2002 H0351.1009, H0351.1012 using a cutoff value for fold expression of 3.0 and greater.
Venny 2.1.0 (http://bioinfogp.cnb.csic.es/tools/venny/) was used to compare the gene lists from four chosen brain donors to identify genes that are common and different across each donor.
Statistical analysis was done in Python Anywhere (https://www.pythonanywhere.com) an online programming tool.
Cluster analysis and Gene Ontology classifications were obtained with DAVID (https://david.ncifcrf.gov)
The STRING database (http://string-db.org) was used to identify potential interacting partners, pathways, and other genes relating to learning.
GeneWeaver (http://www.geneweaver.org) was used to find further information for relevant genes and identify novel gene candidates.associated with learning.

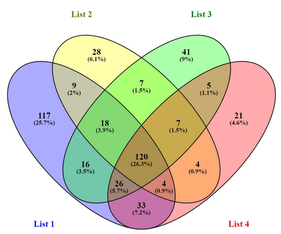
Results

Gene expression profiling



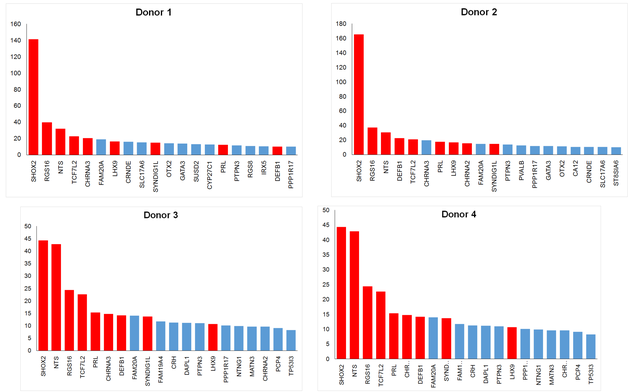
Gene expression data heatmap for the Thalamus of six donor brains. A region of high gene expression is centralized to the dorsal Thalamus as indicated in red.
Key:
-Red: the overrepresented genes in the Thalamus region relative to grey matter
-Green: the underrepresented genes.
-Black: the genes that share no difference in the Thalamus region versus grey matter.

Conserved Genes



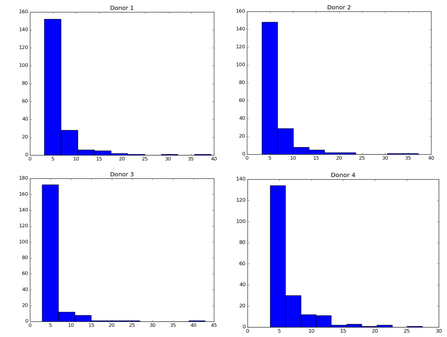
The Venn diagram shows the genes that were common among the four donors. About 26% of the genes within the four donors were shared between them all.

Top 20 genes with highest expression values



The graphs show the top 20 genes with the highest fold change values. The red bars represent the genes shared between all of the donors, while the blue bars represent the genes specific to each donor.

Distributions



The distribution of the gene expression data for each donor is right skewed, with most of the values falling in the lower range.

Statistical analysis

	Donor 1	Donor 2	Donor 3	Donor 4
mean	5.793276	6.114357	5.210719	6.217587
std	4.477543	4.350196	4.050425	3.594923
min	3.127000	3.269000	2.989000	3.570000
25%	3.520750	3.750500	3.415000	4.027000
50%	4.115500	4.772000	3.959000	4.863500
75%	6.476250	6.623250	5.354500	6.891500
max	39.378000	37.223000	42.819000	27.458000

variance
Donor 1 20.048389
Donor 2 18.924202
Donor 3 16.405940
Donor 4 12.923473
dtype: float64
>>>

The mean fold expression for each donor is in the same range; the standard deviation for each donor is in the same range as well, though the standard deviation is fairly large compared to the mean fold expression indicating variation in the data. The minimum and maximum for all four donors are relatively consistent as well.

Common genes classification

Functional enrichment

transcription factor activity	RT	16	13.8	6.9E-4	3.3E-2
Homeobox	RT	8	6.9	7.2E-4	9.6E-2
regulation of amine transport	RT	4	3.4	9.6E-4	2.1E-1
regulation of heart contraction	RT	5	4.3	1.1E-3	2.0E-1
HOX	RT	8	6.9	1.1E-3	8.0E-2
regulation of secretion	RT	7	6.0	1.2E-3	1.9E-1
signal	RT	33	28.4	1.6E-3	5.3E-2
adult behavior	RT	5	4.3	1.7E-3	2.1E-1
postsynaptic membrane	RT	6	5.2	1.7E-3	9.5E-2
behavior	RT	10	8.6	1.8E-3	2.0E-1
signal peptide	RT	33	28.4	1.8E-3	5.0E-1
synapse	RT	9	7.8	1.9E-3	7.2E-2
regulation of neurotransmitter transport	RT	3	2.6	1.0E-2	3.9E-1
cell surface receptor linked signal transduction	RT	20	17.2	1.1E-2	3.8E-1
chemical homeostasis	RT	9	7.8	1.1E-2	3.7E-1

KEGG pathways

KEGG_PATHWAY	Wnt signaling pathway	RT	7	6.0	1.7E-4	8.8E-3
KEGG_PATHWAY	Melanogenesis	RT	6	5.2	2.1E-4	5.5E-3
KEGG_PATHWAY	Calcium signaling pathway	RT	7	6.0	3.9E-4	6.8E-3
KEGG_PATHWAY	Basal cell carcinoma	RT	4	3.4	3.6E-3	4.6E-2
KEGG_PATHWAY	Vascular smooth muscle contraction	RT	4	3.4	2.5E-2	2.4E-1
KEGG_PATHWAY	Alzheimer's disease	RT	4	3.4	6.4E-2	4.4E-1

33 genes: signaling (27.5%)
16 genes: transcription factor (13.3%)
20 genes: receptor activity (16.7%)

31 genes involved in neural related processes /concentration in development, signaling, morphology
ASCL2, ADRA1D, ADM,CACNA1G, CALML3, CHRM2, CHRNA2, CHRNA3, CHRNA4,COL4A3, CPLX3, CRH, GBX2, GDNF, HCN4, IRX3, IRX5, LRRTM1, NTNG1, NTS, OTX2, PVALB, PROK2, PRL, PROX1, RGS16, RPE65, RDH12, SLC6A11, TCF7L2, VANGL1

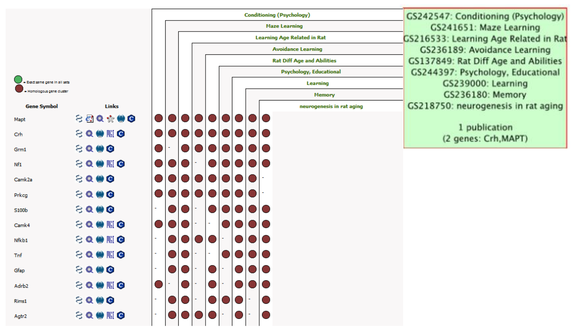
Behavior related categories: feeding behavior, circadian rhythm behavior, locomotory behavior, chemosensory behavior adult behavior, response to nicotine

Candidate genes

CRH	corticotropin releasing hormone	Related Genes	Homo sapiens
BBID	110.CRHandLC-NA.		
GOTERM_BP_FAT	steroid biosynthetic process, glucocorticoid biosynthetic process, induction of apoptosis, defense response, inflammatory response, cell surface receptor linked signal transduction, G-protein coupled receptor protein signaling pathway, cell-cell signaling, synaptic transmission, female pregnancy, parturition, behavior, response to corticosterone, memory, long-term memory, rhythmic behavior, circadian rhythm, feeding behavior, steroid metabolic process, glucocorticoid metabolic process, lipid biosynthetic process, induction of apoptosis by extracellular signals, induction of apoptosis by hormones, response to wounding, regulation of hormone levels, regulation of circadian sleep/wake cycle, wakefulness, positive regulation of circadian sleep/wake cycle, wakefulness, regulation of cell death, positive regulation of cell death, induction of programmed cell death, transmission of nerve impulse, circadian sleep/wake cycle process, respiratory tube development, lung development, adrenal gland development, cellular hormone metabolic process, induction of programmed cell death by hormones, endocrine system development, tube development, regulation of circadian sleep/wake cycle, REM sleep, negative regulation of circadian sleep/wake cycle, sleep, negative regulation of circadian sleep/wake cycle, REM sleep, hormone metabolic process, hormone biosynthetic process, circadian sleep/wake cycle, regulation of circadian sleep/wake cycle, regulation of circadian rhythm, positive regulation of circadian rhythm, negative regulation of circadian rhythm, regulation of apoptosis, positive regulation of apoptosis, regulation of programmed cell death, positive regulation of programmed cell death, regulation of neuron apoptosis, neuroprotection, regulation of system process, regulation of circadian sleep/wake cycle, sleep, regulation of hormone secretion, positive regulation of hormone secretion, rhythmic process, circadian behavior, positive regulation of behavior, negative regulation of behavior, positive regulation of response to stimulus, negative regulation of response to stimulus, gland development, regulation of behavior, neurological system process, cognition, regulation of secretion, positive regulation of secretion, positive regulation of transport, positive regulation of multicellular organismal process, negative regulation of multicellular organismal process, regulation of cortisol secretion, positive regulation of cortisol secretion, regulation of cellular localization, respiratory system development,	

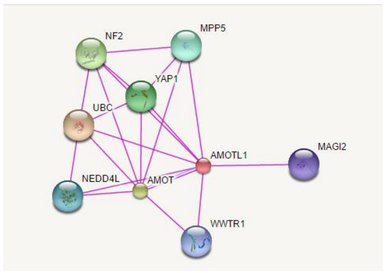
AMOTL1	angiomotin like 1	Related Genes	Homo sapiens
GOTERM_CC_FAT	plasma membrane, cell-cell junction, tight junction, apicolateral plasma membrane, cell junction, apical junction complex, plasma membrane part, pollulting junction,		
GOTERM_MF_FAT	identical protein binding,		
INTERPRO	Angiomotin,		
KEGG_PATHWAY	Tight junction,		
OMIM_DISEASE	Conduct disorder and ADHD,		

Gene set analysis



- Due to the limited protein interactions for CRH in the STRING database, the GeneWeaver database was queried with "CRH and learning" to find further evidence for this association and to identify candidates genes.
- 22 gene sets resulted from the query and 9 of the most relevant were analyzed for gene set interactions (above right.) 1 gene along with CRH is associated with all 9 categories (MAPT2): 4 genes; 8 categories (GRM1, NF1, CAM2KA, PRKCG), 3 genes; 7 categories (S100B, CAMK4, NFKB1) and 5 genes; 6 categories (TNF, GFAP, ADRB2, RIMS1, AGTR2).

Network analysis



The AMOTL1 protein interaction network based on physical interactions.
Ubc: Neuronal developmental and differentiation
Nf2: Cytoskeletal dynamics
Nedd4L: Metal ion homeostasis
Mpp5, AMOT, Cell junction,
Amotl1: Cell junction, ADHD, Conduct disorder
Magi2: Signaling
Wwtr1: Transcription factor

Conclusions

In conclusion, among the common genes between donors, the ones most strongly linked to memory and learning are AMOTL1 and CRH. AMOTL1 is implicated in conduct disorder and ADHD and CRH is associated with learning and memory.

Using AMOTL1 and CRH in a Systems Biology approach, we identified genes of interest that may be involved with learning/memory processes mediated by the Thalamus

The Thalamus, being known for secreting perception signals, is an interesting part of the brain to study for memory/learning, which is typically associated with the hippocampus.

Information regarding the thalamus and learning/memory is a relatively new finding that needs further investigation.