Chapter 5

Global analysis of the interplay between CREB *O*-GlcNAc glycosylation and phosphorylation

Portions of this chapter will be published as:

Jensen EH, Neve R, Hsieh-Wilson LC. "Global analysis of the interplay between site-specific CREB *O*-GlcNAc glycosylation and phosphorylation." *Manuscript in preparation.*

5.1 Abstract

The post-translational modification (PTM) "code" refers to the control of transcription factor (TF) activity through post-translational modifications (PTMs).[1] CREB (cAMP response element binding protein) is a TF that is controlled by a variety of PTMs to regulate neuronal metabolism, activity, differentiation, development, and survival.[2] CREB phosphorylation at serine 133 has been shown to enhance CREB-mediated transcription while CREB glycosylation at serine 40 has been shown to decrease CREB-mediated transcription.[3] The exact gene networks modulated by and potential interplay between CREB glycosylation and phosphorylation have not been explored. Through differential expression analysis with glycosylation-deficient (S40A) and phosphorylation-deficient (S133A) CREB mutants, we show that CREB *O*-GlcNAcylation is important for neuronal activity and excitability while phosphorylation at serine 133 regulates the expression of genes involved in neuronal differentiation. Furthermore, many of the S40A and S133A differentially-expressed genes were directly bound by (1) CREB and its co-activators, CREB-binding protein and p300, (2) activating histone modifications, (3) OGT and *O*-GlcNAc and (4) Tet1, an critical regulator of neuronal activity and differentiation. Finally, we found that CREB *O*-GlcNAcylation regulates activity- and excitotoxicity-related gene networks while CREB phosphorylation regulates neuronal differentiation and amino and fatty acid metabolism-related gene networks. This study demonstrates that CREB *O*-GlcNAcylation at serine 40 and phosphorylation mediate different gene networks. Together, *O*-GlcNAc and phosphorylation impart a TF code,

which CREB must integrate and decode to modulate neuronal activity, differentiation, and metabolism.

4.2 General approach and validation

Despite an understanding of the importance of CREB phosphorylation and glycosylation in mediating particular CREB-regulated genes, the relative importance of each of CREB's PTMs and their potential interplay in negotiating neuronal metabolism, survival, development, and excitability are not well understood. What are the global gene networks regulated by phosphorylation and glycosylation? Are the gene networks distinct or do the PTMs globally affect CREB-mediated transcription across all gene networks? To identify the global transcriptional changes regulated by specific CREB PTMs, we sought to find a system that would allow us to express different CREB modification-defective mutants with minimal interference from endogenous WT CREB.

While the total Creb1 knockout is postnatal lethal, we obtained the Creb1$^{\alpha\delta}$ knockout mice, which lack the Creb1 α and δ isoforms, but express 5% of the β isoform of CREB.[4] We cultured E15-16 Creb1$^{\alpha\delta-/-}$ cortical neurons and used replication-defective herpes simplex virus (HSV) expressing various CREB mutants and GFP (Figure 5.1). Expression of GFP, WT CREB (CREB), a glycosylation-deficient mutant (S40A-CREB), a phosphorylation-deficient mutant (S133A-CREB), and a glycosylation- and phosphorylation-deficient mutant (S40A-S133A-CREB) began at 2 hours with increasing levels up to 12 hours after transduction as observed by immunohistochemical (IHC) staining (Figure 5.2). We chose two time points: 4 hours, when CREB is beginning to be expressed, and 8 hours, when CREB is fully expressed (Figure 5.3). Our previous study showed that neuronal depolarization induced CREB glycosylation, so we KCl

depolarized neurons for 2 hours followed by RNA extraction and RNA-Seq analysis as previously described.[5]

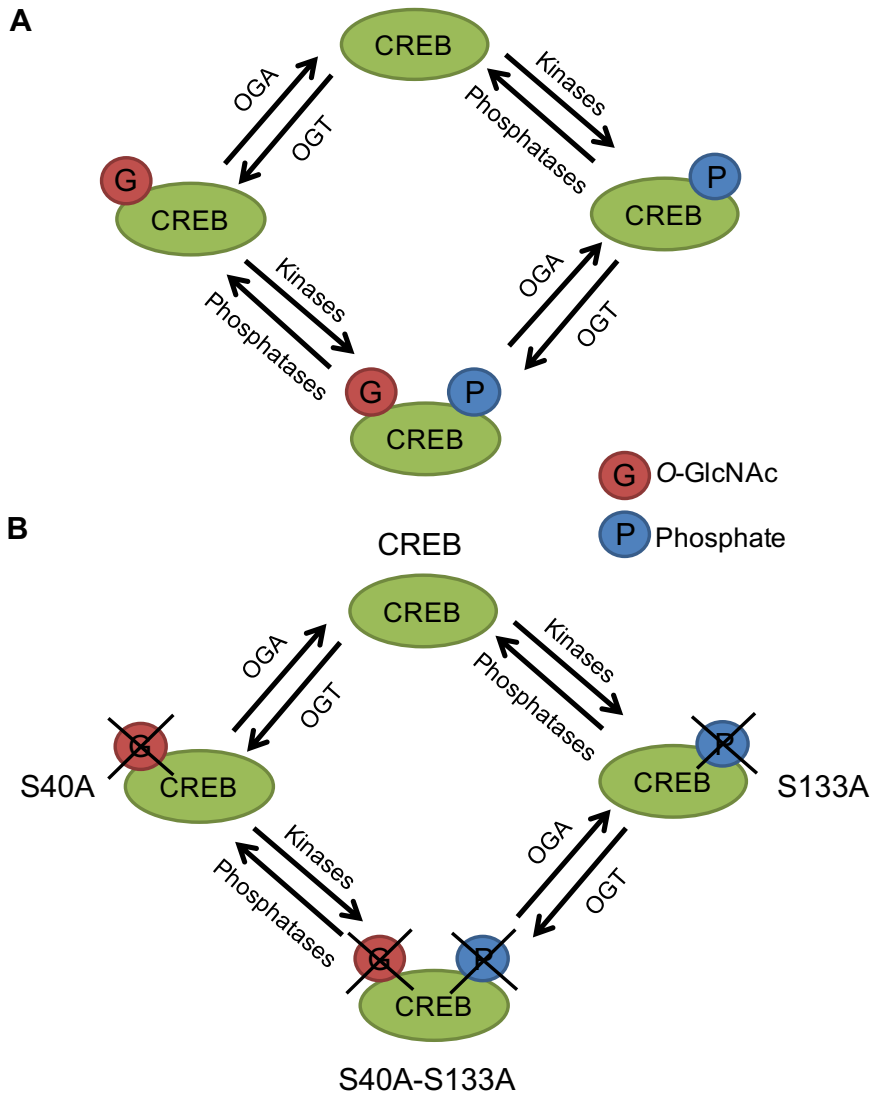


Figure 5.1 Overview of CREB mutants. (A) CREB can be glycosylated at S40 and/or phosphorylated S133. (B) Generation of glycosylation- or phosphorylation- deficient mutants allow us to tease apart contribution of these post-translational modifications to CREB transcription. In particular, we create a non-glycosylatable mutant, S40A-CREB, a non-phosphorylatable mutant, S133-CREB, and a S40A-S133A-CREB mutant that can neither be glycosylated nor phosphorylated.

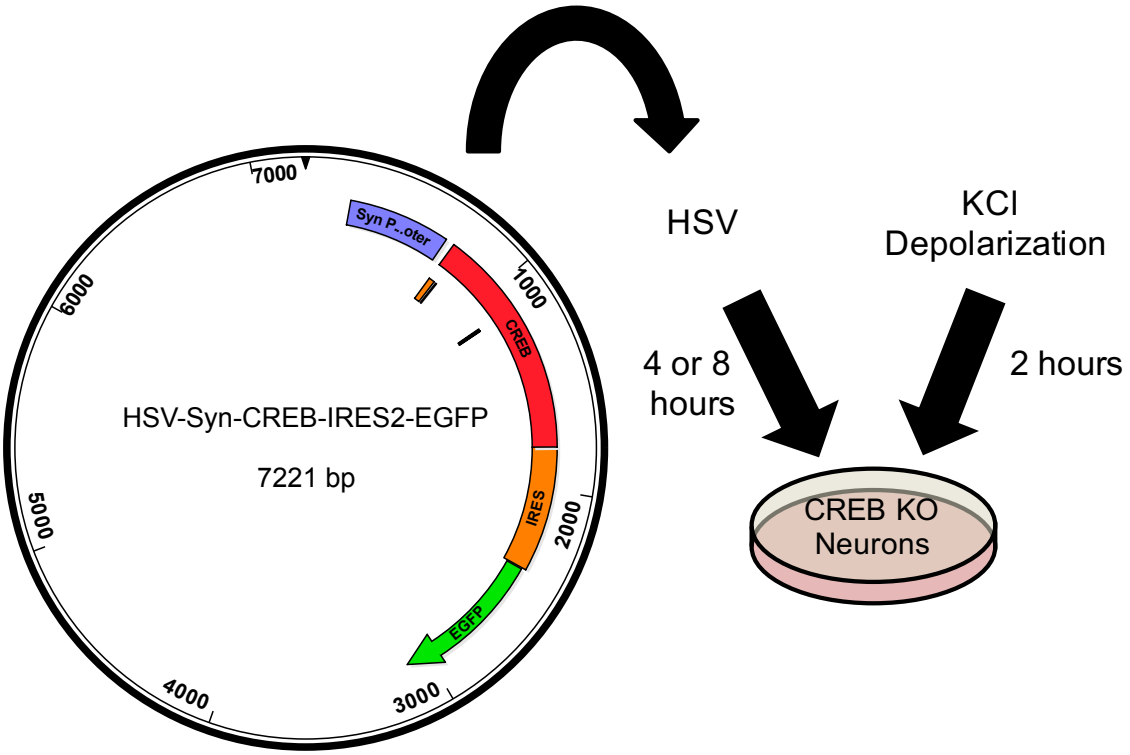


Fig. 5.2. Schematic of experimental overview. Shown here is the overall schematic of the expression of the various CREB mutants using HSV on the CREB KO cortical neuron background. Based on ICC, we chose 2 times points to explore the early and late stage changes. At 4 hours, CREB is beginning to be expressed (incipient changes). At 8 hours, CREB has had several hours to express and sufficient time has elapsed for significant translational changes to occur. We also depolarized neurons with KCl for a total of 2 hours and then we extracted RNA and submitted to RNA-Seq.

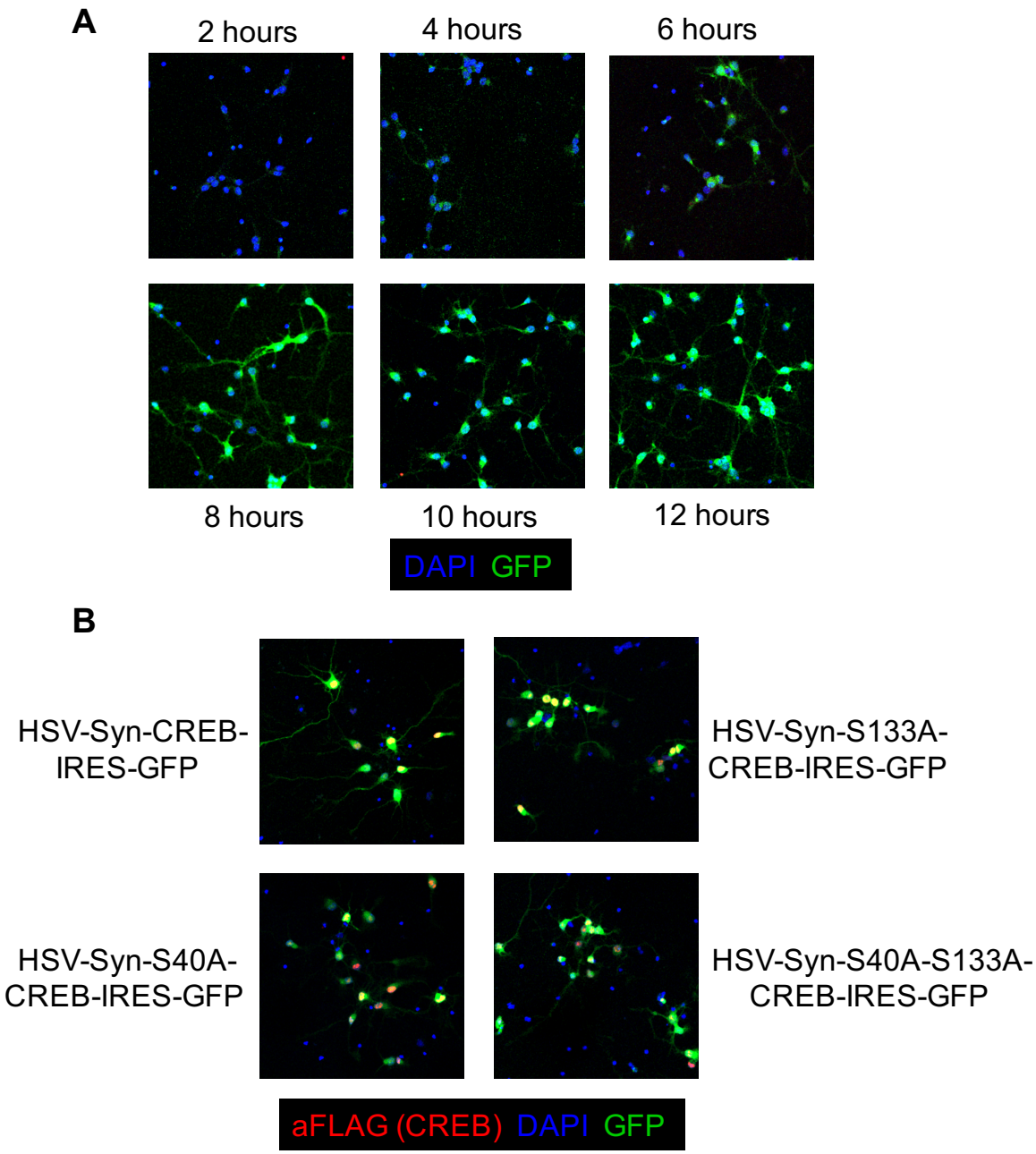


Figure 5.3 HSV transduces neurons rapidly. (A) E16.5 mouse cortical neurons were transfected at 5DIV with HSV-Syn-IRES-GFP for 2, 4, 6, 8, 10, and 12 hours. Expression of GFP is observed starting at 2 hours and increases at 4 hours post-transduction. By 8 hours, the neuronal processes have GFP throughout. (B) E16.5 cortical neurons (5DIV) were transfected with the following HSV for 8 hours: HSV-Syn-CREB-IRES-GFP, HSV-Syn-S40A-CREB-IRES-GFP, HSV-Syn-S133A-CREB-IRES-GFP, HSV-Syn-S40A-S133A-CREB-IRES-GFP. We observed the robust nuclear staining of FLAG (CREB) at 8 hours with all CREB-expressing HSV.

We also verified that the CREB expression levels were consistent using qPCR and RNA-Seq. Indeed, the S40A/WT levels of *Creb1* expression at 4 hours were found to be the same (1.00-fold, q-value = 1.00). There was a 19-fold increase in CREB expression in both the 4-hour WT and S40A conditions when compared to the GFP condition on CREB KO background (q-value = 0.004). Similarly, we verified using qPCR the CREB levels across the 8 hour WT, S40A, S133A, and S40A-S133A within the replicates (Figure 5.4).

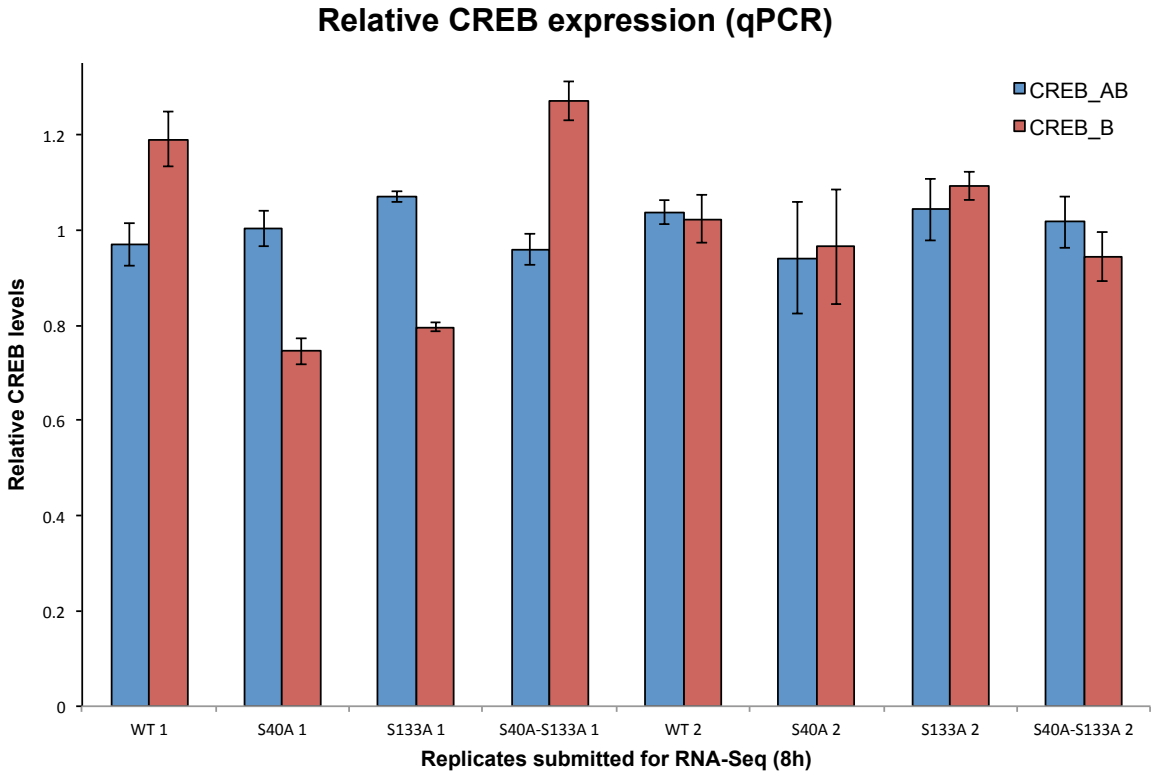


Figure 5.4 Relative CREB expression (qPCR). We monitored the CREB expression levels for the first and second replicates in 8 hours. Shown here are the Creb1 α and β isoform levels (CREB_AB) and the Creb1 β isoform levels (CREB_B). Each replicate group is normalized by the within group average CREB_AB or CREB_B expression levels. The error bars represent the relative standard error of the mean.

5.3 Neuronal polarization and axonogenesis genes are upregulated in the S40A CREB condition at 4 hours

RNA-Seq analysis was first performed with GFP, CREB, and S40A-CREB expression for 4 hours in order to identify the incipient gene expression changes (2

replicates each) (Table 5.1). Differentially-expressed (DE) genes were identified using the Cuffdiff pipeline on the Galaxy interface, and significance is reported as q-values (FDR).[6,7] At 4 hours, we observed upregulation of genes involved in neuronal growth and polarization in the glycosylation-deficient S40A-CREB mutant when compared to the WT CREB condition (FDR <0.05). Specifically, kinesin family member 1C (*Kif1c*, 17-fold), rotatin (*Rttn*, 36-fold), and brain-specific serine/threonine-protein kinase 2 (*Brsk2*, 4.3-fold) are involved in neuronal polarization and axonogenesis (Figure 5.5A).[8] The other two genes were myelin regulatory factor (*Myrf*, 3.5-fold) and CCR4-NOT transcription complex, subunit 3 (*Cnot3*, 4.8-fold), which are critical (1) for differentiation and pluripotency and (2) for synaptic rewiring and mRNA regulation respectively.[9,10]

Table 5.1 S40A/WT differentially-expressed genes at 4 hours

Gene names	Gene Description	S40A/WT log2(FC)	q-value
Actr5	ARP5 Actin Related Protein 5 Homolog	6.6	0.004
Rttn	Rotatin	5.2	0.004
Kif1c	Kinesin Family Member 1C	4.1	0.004
Cnot3	CCR4-NOT Transcription Complex Subunit 3	2.2	0.004
Brsk2	BR Serine/Threonine Kinase 2	2.1	0.004
Myrf	Myelin Regulatory Factor	1.8	0.05
Pdia4	Protein Disulfide Isomerase Family A Member 4	-1.2	0.012
Irak1	Interleukin 1 Receptor Associated Kinase 1	-3.5	0.004
2010111I01Rik	Aminopeptidase O	-6.6	0.004

Table 5.1 shows the differentially-expressed genes at 4 hours including the gene names, descriptions, $\log_2$(FC) where "FC" refers to fold-change, and q-values for the S40A/WT comparison at 4 hours. The upregulated genes are highlighted in pink while the downregulated genes are highlighted in green. q-values (FDR) < 0.05.

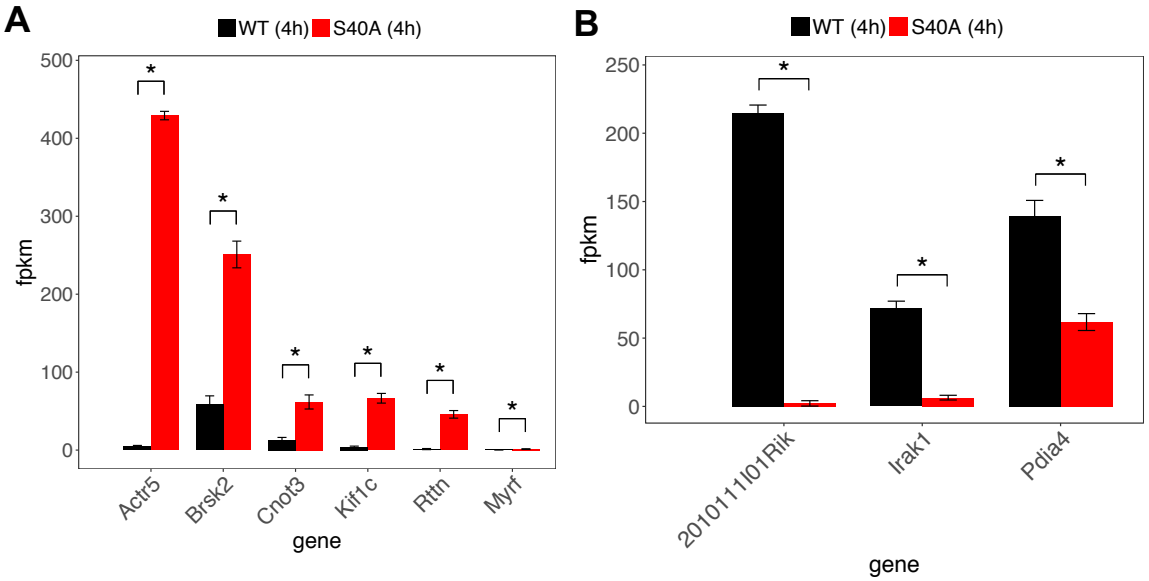


Figure 5.5 Differentially-expressed genes in S40A/WT at 4 hours. (A) The upregulated genes in the S40A condition over the WT CREB condition are involved in neuronal polarization, cytoskeletal rearrangement, and axonogenesis. (B) The downregulated genes in the S40A condition over the WT CREB condition are stress- and immune response-related genes. *FDR < 0.1.

Given the overrepresentation of genes important for neuronal growth and polarization amongst the upregulated genes, we expressed S40A and WT CREB on DIV1 E16.5 CREB KO cortical neurons. After 2 days of HSV expression, we observed enhanced neurite outgrowth in the S40A-expressing neurons consistent with the neurite outgrowth observed in our previous paper (Fig 5.6).[3] Our current and previous results show that ablation of the glycosylation site at serine 40 results in enhanced neuronal growth.

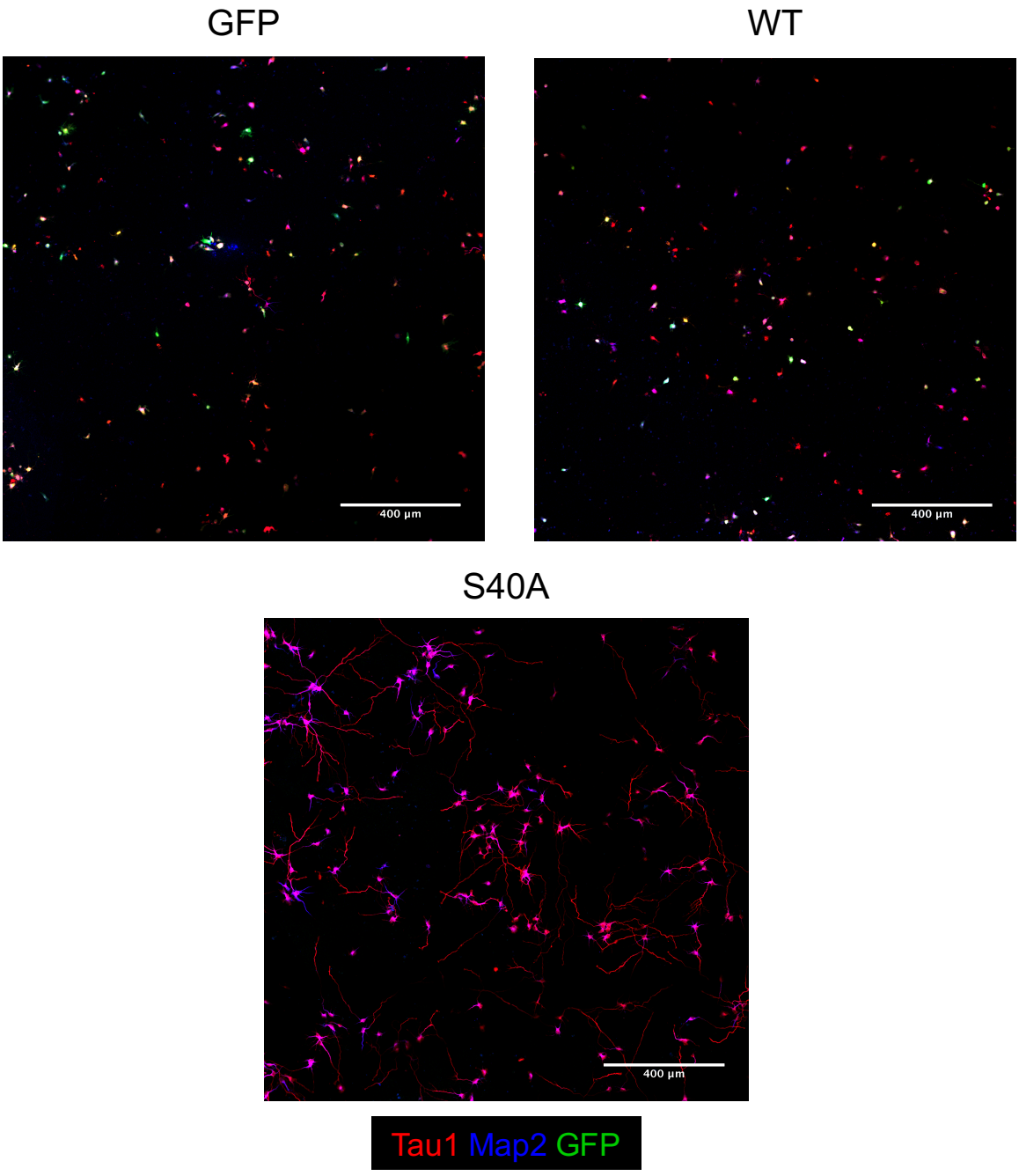


Figure 5.6 S40A-CREB produces enhanced neurite outgrowth. We transduced DIV1 E16.5 CREB KO cortical neurons with HSV-Synapsin-IRES-GFP, HSV-Synapsin-CREB-IRES-GFP, or HSV-Synapsin-S40A-CREB-IRES-GFP (green is GFP). After 2 days of expression, the neurons were fixed and then stained for Tau1 (red) and Map2 (blue). Tau1 is a marker for axons while Map2 stains dendrites and neuronal cell bodies. We observed enhanced neurite outgrowth in the S40A-CREB condition only. Scale bars represent 400

The S40A/WT downregulated genes at 4 hours were broadly involved in stress and immune response, including protein disulfide isomerase family A member 4 (*Pdia4*,

2.2-fold), interleukin 1 receptor associated kinase 1 (*Irak1*, 11-fold), and aminopeptidase O (*2010111I01Rik*, 98-fold) (Figure 5.5B).[11,12] Irak1 is a serine/threonine kinase, which increases the stability of interleukin-1 and is important for immune response.[12] Pdia4 is upregulated in response to endoplasmic reticulum stress and can protect cells from oxidative stress.[11] Aminopeptidase O (ApO) has not been fully characterized, but ApO is believed to be important for angiogenesis.[13] In summary, four hours of expression yielded changes in few genes, but overall, the S40A mutant increased expression of neuronal growth genes and decreased expression of stress and immune response genes.

5.4 Neuronal excitability genes are upregulated in the S40A CREB condition at 8 hours

We next explored the longer-term CREB-induced changes in gene expression. After 8 hours of expression, we identified 87 upregulated genes when we compared S40A-CREB to WT CREB (FDR < 0.1) (Tables 5.2, 5.3). The upregulated genes were enriched for genes encoding synaptic proteins ($P = 2.4\text{x}10^{-4}$), calcium signaling proteins ($P = 3.6\text{x}10^{-3}$), and voltage-gated channels ($P = 8.4\text{x}10^{-4}$) (Table 5.2). Upregulated calcium signaling genes included Ca^{2+}/calmodulin kinase II (*Camk2a,* 1.6-fold) and calcium voltage-gated channel subunit alpha1 I (*Cacna1i*, 1.6-fold) (Figure 5.7A). Calcium signaling is very important for modulating neuronal excitability and LTP.[14] The expression of cAMP signaling genes was also increased in the S40A/WT comparison, including protein kinase A (PKA, *Prkacb*, 1.4-fold), *Camk2a*, and brain-derived neurotrophic factor (*Bdnf,* 1.6-fold), a canonical CREB target important for neuronal growth, LTP, and long-term depression (LTD) (Figure 5.7B).[15]

Table 5.2 DAVID GO annotation of S40A/WT upregulated genes at 8 hours

GO term	#	Gene names	FE
Dendrite	14	*Kcnh1, Ache, Kcnc4, Cplx1, Ddn, Shh, Crhr1, Kcnj4, Bdnf, Lynx1, Chrm1, Negr1, Camk2a, Synpo*	7.0
Neuronal cell body	13	*Kcnh1, Ache, Kcnc4, Cplx1, Ptprn, Shh, Crhr1, Kcnj4, Bdnf, Crh, Nrsn1, Camk2a, Negr1*	5.9
Synapse	12	*Kcnh1, Kcnj4, Ache, Cplx1, Syndig1, Chrm1, Grin2d, Psd3, Lgi3, Ptprn, Camk2a, Synpo*	5.8
Voltage-gated channel	7	*Kcnh1, Kcnj4, Kcnc4, Kcns2, Kcnj9, Cacna1i, Kcnh3*	14
Glycoprotein	31	*Kcnh1, Kcnc4, Ache, Slc6a1, Spock3, Epha10, Shh, Col26a1, Bdnf, Rspo1, Creg2, Grin2d, Lgi3, Cntnap1, Etl4, Loxl2, Negr1, Synpo, Phyhip, Cckbr, Cdhr1, Ai593442, Ptprn, Crhr1, Lynx1, Slc6a7, Chrm1, Cemip, Wif1, Car4, Kcnh3*	2.1
Circadian entrainment	6	*Kcnj9, Grin2d, Cacna1i, Prkacb, Gng4, Camk2a*	14
Potassium transport	6	*Kcnh1, Kcnj4, Kcnc4, Kcns2, Kcnj9, Kcnh3*	15
Calcium signaling pathway	7	*Cckbr, Chrm1, Grin2d, Cacna1i, Prkacb, Itpka, Camk2a*	9.1
Cholinergic synapse	6	*Kcnj4, Ache, Chrm1, Prkacb, Gng4, Camk2a*	12
Axon	9	*Kcnh1, Bdnf, Ache, Kcnc4, Slc6a1, Cntnap1, Camk2a, Shh, Synpo*	5.9
Cell junction	11	*Kcnj4, Ache, Syndig1, Chrm1, Grin2d, Psd3, Lgi3, 9430020k01rik, Ptprn, Camk2a, Synpo*	4.5
Postsynaptic membrane	7	*Kcnj4, Ache, Syndig1, Chrm1, Grin2d, Psd3, Synpo*	7.7
Membrane	44	*Kcnh1, Ache, Kcnc4, Syndig1, Slc6a1, Ildr2, Epha10, Tmem151b, Tmem151a, Shh, Tmcc2, Kcns2, Pacsin3, Grin2d, Rasl10b, Cntnap1, Prkacb, Loxl2, Gng4, Negr1, Camk2a, Synpo, Ngef, Cckbr, Cdhr1, Ai593442, Psd3, Ptprn, Ddn, Crhr1, Kcnj4, Lynx1, Slc6a7, Kcnj9, Mtfp1, Chrm1, Vsnl1, Cemip, Nrsn1, Car4, Cend1, Parp1, Fam163b, Kcnh3*	1.5
Lipoprotein	11	*Ache, Lynx1, Cckbr, Ncald, Vsnl1, Rasl10b, Prkacb, Car4, Gng4, Negr1, Shh*	3.7
Secreted	16	*Ache, Spock3, Fam24a, Shh, Bdnf, Col26a1, Lynx1, Rspo1, Creg2, Cemip, Crh, Cartpt, Wif1, Lgi3, Loxl2, Scg2*	2.5
Phosphoprotein	43	*Kcnh1, Gda, Kcnc4, Syndig1, Tcap, Slc6a1, Ankrd34c, Cnot3, Ildr2, 9430020k01rik, Itpka, Tmcc2, Rasal1, Fbxw7, Pacsin3, Inpp5j, Grin2d, Osbpl1a, Apba3, Etl4, Cntnap1, Prkacb, Camk2a, Negr1, Scg2, Synpo, Ngef, Map1a, Esrrg, Psd3, Ptprn, Ddn, Rcan2, Crhr1, Dact2, Slc6a7, Nab2, Chrm1, Zbtb4, Cartpt, Parp1, Cend1, Fam163b*	1.5
Ion transport	9	*Kcnh1, Kcnj4, Kcnc4, Kcns2, Kcnj9, Grin2d, Cacna1i, Atp6v1g2, Kcnh3*	3.8
Postsynaptic density	6	*Syndig1, Chrm1, Map1a, Psd3, Camk2a, Synpo*	6.1

Table 5.2 contains the DAVID functional gene ontology annotations of the S40A-CREB/ CREB upregulated differentially-expressed genes. Benjamini-corrected p-values are all less than 0.05. FE = fold enrichment.

Table 5.3 List of differentially-expressed S40A/WT genes at 8 hours

Gene names	Gene description	S40A/WT log2(FC)	q-value
Loxl2	lysyl oxidase-like 2(Loxl2)	7.2	0.01
Susd5	sushi domain containing 5(Susd5)	2.4	0.03
Rcan3	regulator of calcineurin 3(Rcan3)	2.2	0.01
Cartpt	CART prepropeptide(Cartpt)	2.1	0.01

Car4	carbonic anhydrase 4(Car4)	2.0	0.01
Wif1	Wnt inhibitory factor 1(Wif1)	1.8	0.01
Rspo1	R-spondin 1(Rspo1)	1.6	0.03
Ankrd34c	ankyrin repeat domain 34C(Ankrd34c)	1.6	0.03
Cckbr	cholecystokinin B receptor(Cckbr)	1.4	0.01
Itpka	inositol 1,4,5-trisphosphate 3-kinase A(Itpka)	1.2	0.01
Apba3	amyloid beta (A4) precursor protein-binding, family A, member 3(Apba3)	1.2	0.01
Parp1	poly (ADP-ribose) polymerase family, member 1(Parp1)	1.2	0.01
Tcap	titin-cap(Tcap)	1.1	0.02
Crh	corticotropin releasing hormone(Crh)	1.1	0.05
Creg2	cellular repressor of E1A-stimulated genes 2(Creg2)	1.0	0.01
Pacsin3	protein kinase C and casein kinase substrate in neurons 3(Pacsin3)	1.0	0.01
Mybpc1	myosin binding protein C, slow-type(Mybpc1)	1.0	0.02
Lynx1	Ly6/neurotoxin 1(Lynx1)	0.85	0.01
Cdhr1	cadherin-related family member 1(Cdhr1)	0.85	0.03
Kcns2	K+ voltage-gated channel, subfamily S, 2(Kcns2)	0.85	0.02
Lgi3	leucine-rich repeat LGI family, member 3(Lgi3)	0.84	0.02
Nab2	Ngfi-A binding protein 2(Nab2)	0.81	0.04
Nrsn1	neurensin 1(Nrsn1)	0.77	0.01
Cntnap1	contactin associated protein-like 1(Cntnap1)	0.74	0.03
Rasal1	RAS protein activator like 1 (GAP1 like)(Rasal1)	0.74	0.05
Mtfp1	mitochondrial fission process 1(Mtfp1)	0.74	0.04
Gng4	guanine nucleotide binding protein (G protein), gamma 4(Gng4)	0.73	0.01
Cemip	cell migration inducing protein, hyaluronan binding(Cemip)	0.73	0.01
Slc6a7	solute carrier family 6 (neurotransmitter transporter, L-proline), member 7(Slc6a7)	0.73	0.01
Inpp5j	inositol polyphosphate 5-phosphatase J(Inpp5j)	0.73	0.02
Cnot3	CCR4-NOT transcription complex, subunit 3(Cnot3)	0.72	0.02
Dact2	dishevelled-binding antagonist of beta-catenin 2(Dact2)	0.71	0.04
Bdnf	brain derived neurotrophic factor(Bdnf)	0.68	0.01
Kcnh3	potassium voltage-gated channel, subfamily H (eag-related), member 3(Kcnh3)	0.68	0.06
Kcnj9	potassium inwardly-rectifying channel, subfamily J, member 9(Kcnj9)	0.67	0.01
Camk2a	calcium/calmodulin-dependent protein kinase II alpha(Camk2a)	0.66	0.01
Esrrg	estrogen-related receptor gamma(Esrrg)	0.66	0.09
Rcan2	regulator of calcineurin 2(Rcan2)	0.65	0.01
Ddn	dendrin(Ddn)	0.65	0.01
Kcnc4	potassium voltage gated channel, Shaw-related subfamily, member 4(Kcnc4)	0.65	0.03
Cacna1i	calcium channel, voltage-dependent, alpha 1I subunit(Cacna1i)	0.65	0.01

Spock3	sparc/osteonectin, cwcv and kazal-like domains proteoglycan 3(Spock3)	0.65	0.01
Etl4	enhancer trap locus 4(Etl4)	0.65	0.02
9430020K01 Rik	RIKEN cDNA 9430020K01 gene(9430020K01Rik)	0.64	0.01
Fam163b	family with sequence similarity 163, member B(Fam163b)	0.64	0.03
Vsnl1	visinin-like 1(Vsnl1)	0.63	0.01
Shh	sonic hedgehog(Shh)	0.63	0.03
Col26a1	collagen, type XXVI, alpha 1(Col26a1)	0.62	0.06
Synpo	synaptopodin(Synpo)	0.62	0.03
Rassf3	Ras association (RalGDS/AF-6) domain family member 3(Rassf3)	0.61	0.05
Ncald	neurocalcin delta(Ncald)	0.60	0.01
Phyhip	phytanoyl-CoA hydroxylase interacting protein(Phyhip)	0.59	0.01
Ngef	neuronal guanine nucleotide exchange factor(Ngef)	0.59	0.01
Cend1	cell cycle exit and neuronal differentiation 1(Cend1)	0.59	0.01
Syndig1	synapse differentiation inducing 1(Syndig1)	0.58	0.03
Tmem151a	transmembrane protein 151A(Tmem151a)	0.58	0.01
Ildr2	immunoglobulin-like domain containing receptor 2(Ildr2)	0.58	0.06
Epha10	Eph receptor A10(Epha10)	0.57	0.06
Nrip3	nuclear receptor interacting protein 3(Nrip3)	0.57	0.02
Tmcc2	transmembrane and coiled-coil domains 2(Tmcc2)	0.57	0.02
Kcnh1	potassium voltage-gated channel, subfamily H (eag-related), member 1(Kcnh1)	0.56	0.01
Grin2d	glutamate receptor, ionotropic, NMDA2D (epsilon 4)(Grin2d)	0.56	0.02
A830018L16 Rik	RIKEN cDNA A830018L16 gene(A830018L16Rik)	0.55	0.10
Me3	malic enzyme 3, NADP(+)-dependent, mitochondrial(Me3)	0.55	0.10
Fam81a	family with sequence similarity 81, member A(Fam81a)	0.55	0.03
Slc6a1	solute carrier family 6 (neurotransmitter transporter, GABA), member 1(Slc6a1)	0.54	0.04
Crhr1	corticotropin releasing hormone receptor 1(Crhr1)	0.54	0.09
Atp6v1g2	ATPase, H+ transporting, lysosomal V1 subunit G2(Atp6v1g2)	0.54	0.02
Kcnj4	potassium inwardly-rectifying channel, subfamily J, member 4(Kcnj4)	0.54	0.05
Osbpl1a	oxysterol binding protein-like 1A(Osbpl1a)	0.54	0.01
Map1a	microtubule-associated protein 1 A(Map1a)	0.54	0.03
AI593442	expressed sequence AI593442(AI593442)	0.53	0.03
Chrm1	cholinergic receptor, muscarinic 1, CNS(Chrm1)	0.52	0.08
Scg2	secretogranin II(Scg2)	0.52	0.06
1700020I14Rik	RIKEN cDNA 1700020I14 gene(1700020I14Rik)	0.51	0.07
Zbtb4	zinc finger and BTB domain containing 4(Zbtb4)	0.51	0.09
D3Bwg0562e	phospholipid phosphatase related 4(Plppr4)	0.51	0.05

Ache	acetylcholinesterase(Ache)	0.50	0.09
Cplx1	complexin 1(Cplx1)	0.49	0.09
Ptprn	protein tyrosine phosphatase, receptor type, N(Ptprn)	0.49	0.06
Rasl10b	RAS-like, family 10, member B(Rasl10b)	0.49	0.06
Prkacb	protein kinase, cAMP dependent, catalytic, beta(Prkacb)	0.48	0.09
Fbxw7	F-box and WD-40 domain protein 7(Fbxw7)	0.48	0.08
Gda	guanine deaminase(Gda)	0.48	0.10
Psd3	pleckstrin and Sec7 domain containing 3(Psd3)	0.48	0.08
Tmem151b	transmembrane protein 151B(Tmem151b)	0.47	0.10
Negr1	neuronal growth regulator 1(Negr1)	0.46	0.08
Grik3	glutamate receptor, ionotropic, kainate 3(Grik3)	-0.48	0.09
Islr2	immunoglobulin superfamily containing leucine-rich repeat 2(Islr2)	-0.51	0.08
Draxin	dorsal inhibitory axon guidance protein(Draxin)	-0.51	0.07
Shb	src homology 2 domain-containing transforming protein B(Shb)	-0.51	0.10
Dkk3	dickkopf WNT signaling pathway inhibitor 3(Dkk3)	-0.52	0.06
D8Ertd82e	DNA segment, Chr 8, ERATO Doi 82, expressed(D8Ertd82e)	-0.53	0.09
Bcl2l11	BCL2-like 11 (apoptosis facilitator)(Bcl2l11)	-0.53	0.06
Ezh2	enhancer of zeste 2 polycomb repressive complex 2 subunit(Ezh2)	-0.54	0.09
Nefm	neurofilament, medium polypeptide(Nefm)	-0.54	0.04
Sox11	SRY (sex determining region Y)-box 11(Sox11)	-0.55	0.04
Tle4	transducin-like enhancer of split 4(Tle4)	-0.57	0.02
Fst	follistatin(Fst)	-0.57	0.10
Epha2	Eph receptor A2(Epha2)	-0.58	0.04
Gpc2	glypican 2 (cerebroglycan)(Gpc2)	-0.62	0.01
Ebf3	early B cell factor 3(Ebf3)	-0.64	0.08
Igf2	insulin-like growth factor 2(Igf2)	-0.66	0.08
Nckap5	NCK-associated protein 5(Nckap5)	-0.70	0.01
Trp53i11	transformation related protein 53 inducible protein 11(Trp53i11)	-0.70	0.02
Carhsp1	calcium regulated heat stable protein 1(Carhsp1)	-0.72	0.01
Nxph3	neurexophilin 3(Nxph3)	-0.74	0.03
Arhgap28	Rho GTPase activating protein 28(Arhgap28)	-0.75	0.04
Id1	inhibitor of DNA binding 1(Id1)	-0.75	0.03
Tshz2	teashirt zinc finger family member 2(Tshz2)	-0.76	0.01
Fn1	fibronectin 1(Fn1)	-0.77	0.01
Foxp2	forkhead box P2(Foxp2)	-0.78	0.08
Plin2	perilipin 2(Plin2)	-0.79	0.02
Kif26a	kinesin family member 26A(Kif26a)	-0.81	0.01
Adam8	a disintegrin and metallopeptidase domain 8(Adam8)	-0.81	0.05
Zfhx3	zinc finger homeobox 3(Zfhx3)	-0.83	0.03

Id3	inhibitor of DNA binding 3(Id3)	-0.84	0.01
Hmox1	heme oxygenase 1(Hmox1)	-0.87	0.01
Bcas1	breast carcinoma amplified sequence 1(Bcas1)	-0.89	0.09
Tcf7l2	transcription factor 7 like 2, T cell specific, HMG box(Tcf7l2)	-0.92	0.02
Txnip	thioredoxin interacting protein(Txnip)	-0.95	0.01
Myrf	myelin regulatory factor(Myrf)	-1.0	0.08
Nts	neurotensin(Nts)	-1.1	0.01
Plxnb3	plexin B3(Plxnb3)	-1.1	0.09
Plp1	proteolipid protein (myelin) 1(Plp1)	-1.1	0.05
Igfbpl1	insulin-like growth factor binding protein-like 1(Igfbpl1)	-1.2	0.01
Cldn11	claudin 11(Cldn11)	-1.3	0.02
Nckap1l	NCK associated protein 1 like(Nckap1l)	-1.5	0.06
Laptm5	lysosomal-associated protein transmembrane 5(Laptm5)	-1.6	0.01
Mag	myelin-associated glycoprotein(Mag)	-1.6	0.04
H19	H19, imprinted maternally expressed transcript(H19)	-1.6	0.01
Lgals3	lectin, galactose binding, soluble 3(Lgals3)	-1.7	0.01
Cyba	cytochrome b-245, alpha polypeptide(Cyba)	-1.7	0.02
Rsad2	radical S-adenosyl methionine domain containing 2(Rsad2)	-1.7	0.08
Pdia4	protein disulfide isomerase associated 4(Pdia4)	-1.8	0.01
Pou2f2	POU domain, class 2, transcription factor 2(Pou2f2)	-1.8	0.01
Enpp6	ectonucleotide pyrophosphatase/phosphodiesterase 6(Enpp6)	-1.9	0.01
Slc11a1	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 1(Slc11a1)	-2.0	0.05
C3ar1	complement component 3a receptor 1(C3ar1)	-2.1	0.01
Ctss	cathepsin S(Ctss)	-2.1	0.01
E130102H24 Rik,Mir101a	microRNA 101a(Mir101a)	-2.1	0.04
Cybb	cytochrome b-245, beta polypeptide(Cybb)	-2.5	0.02
Itgb2	integrin beta 2(Itgb2)	-2.5	0.01
Cd36	CD36 antigen(Cd36)	-2.5	0.01
Itgam	integrin alpha M(Itgam)	-2.5	0.02
Gpnmb	glycoprotein (transmembrane) nmb(Gpnmb)	-2.6	0.01
Ttr	transthyretin(Ttr)	-2.9	0.01
Spp1	secreted phosphoprotein 1(Spp1)	-2.9	0.01
C1qb	complement component 1, q subcomponent, beta polypeptide(C1qb)	-3.0	0.03
Lyz2	lysozyme 2(Lyz2)	-3.0	0.01
2010111I01Rik	RIKEN cDNA 2010111I01 gene(2010111I01Rik)	-3.2	0.01
C1qa	complement component 1, q subcomponent, alpha polypeptide(C1qa)	-3.3	0.03
Tyrobp	TYRO protein tyrosine kinase binding protein(Tyrobp)	-3.4	0.01
C1qc	complement component 1, q subcomponent, C chain(C1qc)	-3.7	0.02

Kif1c	kinesin family member 1C(Kif1c)	-3.8	0.01
Mmp12	matrix metallopeptidase 12(Mmp12)	-4.6	0.01
Tlcd1	TLC domain containing 1(Tlcd1)	-5.7	0.01

Table 5.3 shows the gene names, descriptions, $\log_2$(FC) where "FC" refers to fold-change, and q-values for the differentially-expressed genes in the S40A/WT comparison at 8 hours. The upregulated genes are highlighted in pink while the downregulated genes are highlighted in green. q-values (FDR) < 0.1.

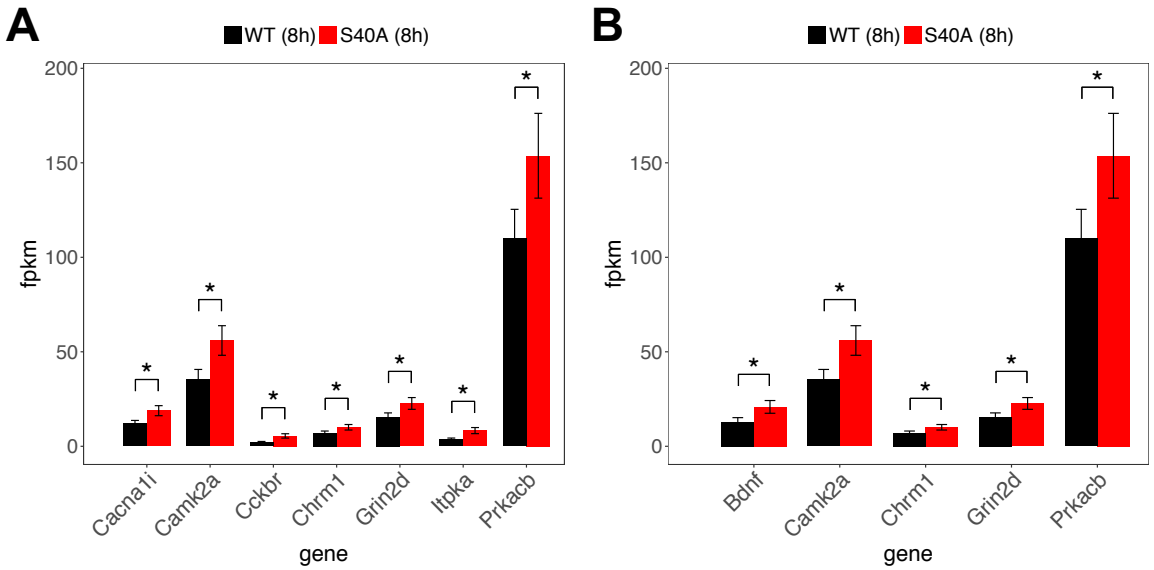


Figure 5.7 Expression levels of upregulated genes in S40A/WT involved in calcium and cAMP signaling pathways at 8 hours. (A) Calcium signaling pathway genes and (B) cAMP signaling pathway genes are upregulated in the S40A condition over the WT CREB condition. *FDR < 0.1.

LTP-related genes were also upregulated in the S40A condition, including glutamate ionotropic receptor NMDA type subunit 2D (*Grin2D*, 1.5-fold) in addition to *Bdnf* and *Prkacb*. Voltage-gated channels such as several potassium (*Kcnc4, Kcnh1, Kcnh3*, *Kcnj4, Kcnj9, Kcns2,* 1.5-1.8-fold) and calcium (*Cacna1i*, 1.6-fold) channels were enriched in the S40A condition (Figure 5.7A). Increased levels of voltage-gated channels lower the barrier for neuronal depolarization, thereby facilitating LTP and synaptic plasticity.[14] Finally, the S40A mutant increased the expression of genes that encode for proteins found in the dendrite (P = 1.0x10^{-5}), axon (P = 3.9x10^{-3}), neuronal cell body (P = 9.7x10^{-5}), and synapse (P = 2.4x10^{-4}), suggesting that S40A leads to enhanced neuronal growth and activity (Figures 5.8B-E, 5.9, Table 5.2). Overall, these

results suggest that CREB glycosylation at serine 40 plays a critical role in calibrating homeostatic neuronal excitability and LTP.

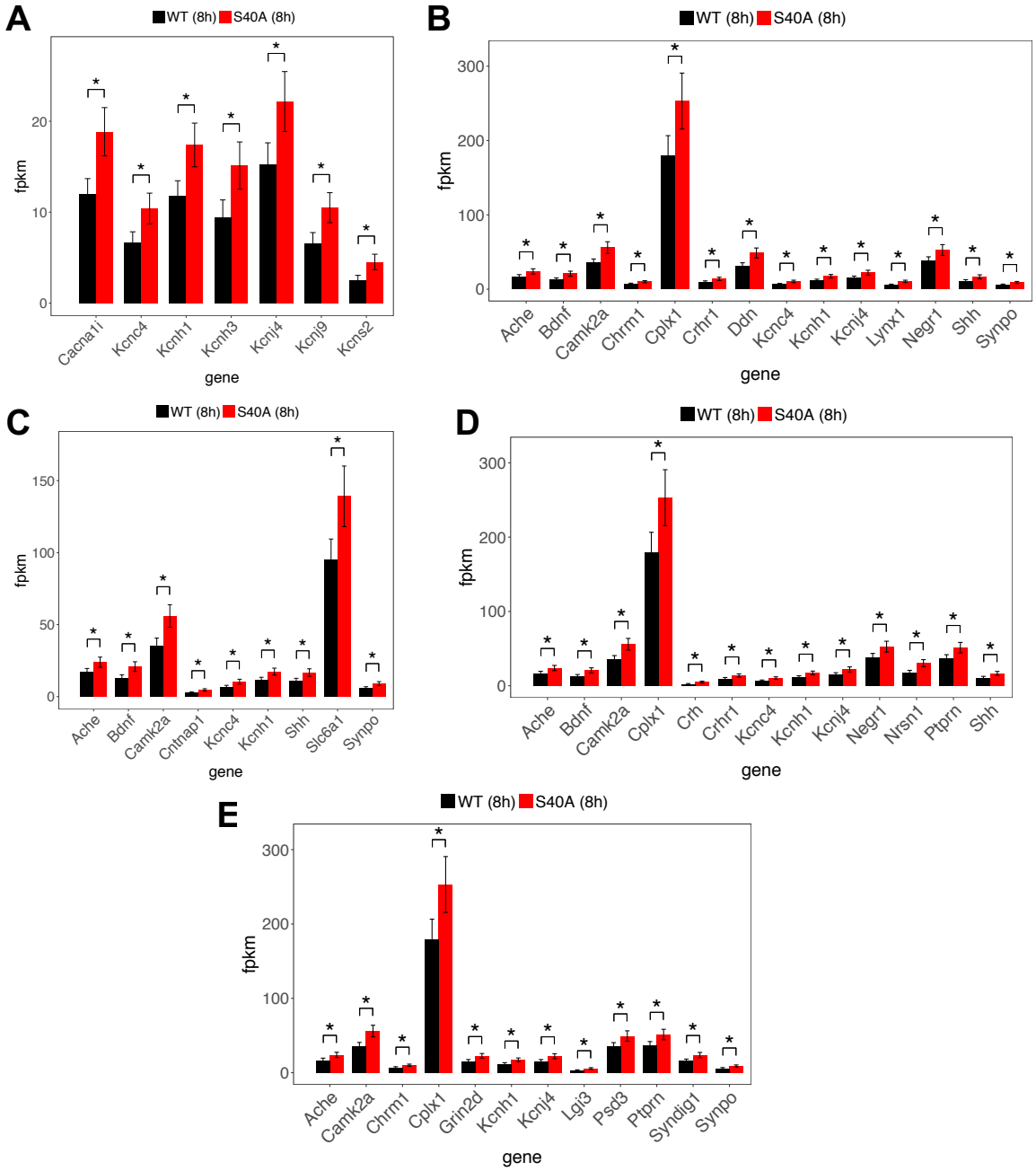


Figure 5.8 Expression levels for neuronal activity upregulated genes in S40A/WT at 8 hours. (A) Voltage-gated channels, (B) dendritic genes, (C) axonal genes, (D) neuronal cell body, and (E) synaptic genes are upregulated in S40A/WT after 8 hours of HSV treatment. *FDR < 0.1.

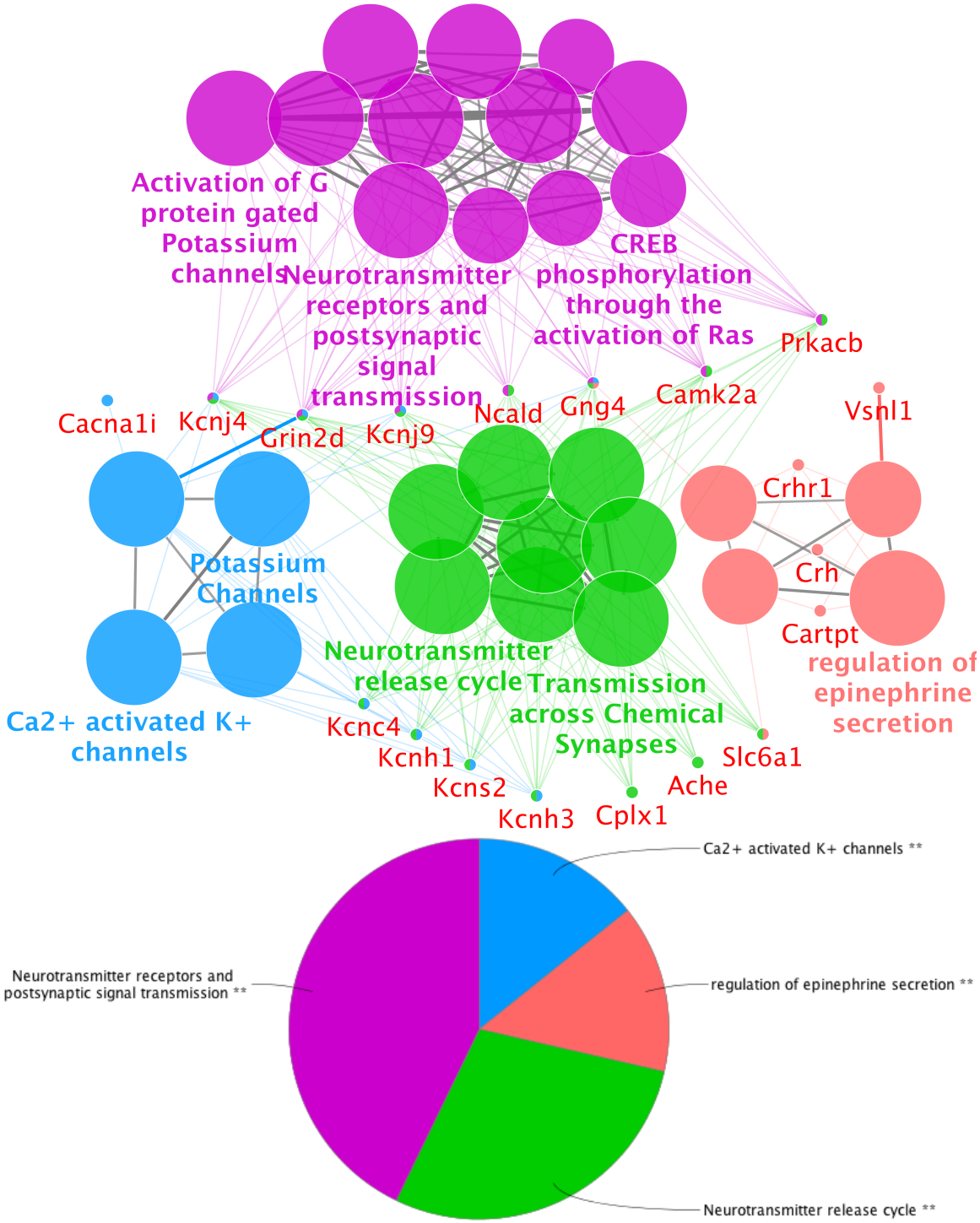


Figure 5.9 Cytoscape gene ontology annotations for the S40A/WT upregulated genes at 8 hours. Shown here are the top gene ontology categories for (A) the upregulated genes and (B) the downregulated genes in the S40A/WT comparison. The genes that belong to and are shared by these gene ontology categories are also shown. The gene annotations were created performed using the Reactome Pathways database for the downregulated genes and GO Molecular Functions, GO Biological Process, and Reactome Pathways databases for the upregulated genes with the ClueGO extension in Cytoscape. The significance for all listed categories are Bonferroni step down corrected $P < 0.0029$.

Our paper corroborates our previous study showing that S40A-CREB primes neurons for memory formation.[3] In particular, our previous study showed that important memory-related genes such as *Bdnf* showed increased expression with S40A-CREB expression when compared to CREB expression. This study also found that *Bdnf* expression was enhanced in the S40A-CREB condition in addition to neuronal genes from the dendrites to the axons to the neuronal cell body. Indeed, our current results support the findings in our previous paper where we demonstrated that injecting S40A-CREB-expressing HSV into the lateral amygdala lead to enhanced memory consequences at 2h, which leveled off at 24 hours.[3] Increased *Bdnf* expression was upregulated in other studies expressing other more active CREB mutants, VP16-CREB, Y134F-CREB, and DIEDML-CREB.[16,17]

5.5 Innate immune response and phagosome genes are downregulated in the glycosylation-deficient mutant at 8 hours

Next, we found that the 70 downregulated genes (Table 5.3) in the S40A/WT comparison were involved in innate-immune response ($P = 4.8\text{x}10^{-2}$) and the phagosome ($P = 2.3\text{x}10^{-2}$) at 8 hours (Table 5.4). S40A displayed a reduction in innate immune response-related genes, including complement component 1 subunits (*C1qa*, *C1qb*, *C1qc*, 7.8-13-fold); lectin, galactose-binding, soluble 3 (galectin-3, *Lgals3*, 3.2-fold); and radical *S*-adenosyl methionine domain containing 2 (*Rsad2*, 3.3-fold) (Figure 5.10A).[18] Immune activation is critical for synaptic pruning especially during early stages of neuronal development.[18] Therefore, impaired immune response at this stage could lead to inappropriate synaptic connections and pruning. Finally, phagosomal genes were downregulated in the S40A-CREB condition, including cytochrome b-245 light chain (*Cyba*, 3.2-fold) and cytochrome b-245 heavy chain (*Cybb*, 5.2-fold), which are the major

components of the phagocytic oxidase responsible for generating superoxide (Figure 5.10B).[19] Altogether, the upregulated genes in the S40A over WT CREB condition were involved in neuronal activation and enhancement of neuronal excitability while the downregulated genes were related to immune response and phagocytosis (Figure 5.11).

Table 5.4 DAVID functional annotation of S40A/WT downregulated genes

GO term	#	Gene names	FE
Glycoprotein	31	*Nxph3, Enpp6, C3ar1, Igfbpl1, Grik3, Fst, Ezh2, Itgb2, Pdia4, Itgam, Slc11a1, Ttr, Gpc2, Adam8, Gpnmb, Nefm, Spp1, Fn1, Islr2, Mag, Draxin, Plxnb3, Ctss, Mmp12, Epha2, C1qa, Dkk3, C1qb, Cybb, Cd36, Myrf*	2.7
Signal	32	*Nxph3, Enpp6, Igfbpl1, Grik3, Fst, Itgb2, Pdia4, C1qc, Itgam, Ttr, Gpc2, Adam8, Gpnmb, Spp1, Tyrobp, Fn1, Islr2, Mag, Lyz2, Lgals3, Draxin, Plxnb3, Igf2, Ctss, Mmp12, Epha2, Bcl2l11, C1qa, Dkk3, C1qb, Tlcd1, Nts*	2.4
Staphylococcus aureus infection	6	*C1qa, C3ar1, C1qb, Itgb2, C1qc, Itgam*	26
Secreted	17	*Nxph3, Lyz2, Lgals3, Igfbpl1, Draxin, Fst, Igf2, C1qc, Mmp12, C1qa, Dkk3, C1qb, Ttr, Gpc2, Nts, Spp1, Fn1*	3.4
Cell surface	12	*Slc11a1, Cd36, Lgals3, Plxnb3, Itgb2, Ctss, Pdia4, Adam8, Itgam, Epha2, Tyrobp, Islr2*	5.4
Extracellular region	18	*Nxph3, Enpp6, Lyz2, Lgals3, Igfbpl1, Draxin, Fst, Igf2, C1qc, Mmp12, C1qa, Dkk3, C1qb, Ttr, Gpc2, Nts, Spp1, Fn1*	2.9
Pertussis	5	*C1qa, C1qb, Itgb2, C1qc, Itgam*	15
Phagosome	6	*Cyba, Cybb, Cd36, Itgb2, Ctss, Itgam*	7.4
Repressor	8	*Tshz2, Id1, Ezh2, Tle4, Id3, Zfhx3, Tcf7l2, Foxp2*	5.0
Leukocyte transendothelial migration	5	*Cyba, Cybb, Itgb2, Cldn11, Itgam*	8.9
Respiratory burst	3	*Slc11a1, Cyba, Cybb*	105
Integrin-mediated signaling pathway	5	*Plp1, Itgb2, Adam8, Itgam, Tyrobp*	15
Cell adhesion	9	*Mag, Cd36, Itgb2, Cldn11, Gpnmb, Itgam, Epha2, Fn1, Spp1*	5.2
Neutrophil chemotaxis	5	*Lgals3, Nckap1l, Itgb2, Itgam, Spp1*	19
Negative regulation of transcription from RNA polymerase II promoter	11	*Txnip, Cd36, Id1, Sox11, Fst, Ezh2, Tle4, Id3, Zfhx3, Tcf7l2, Foxp2*	4.2
Innate immune response	8	*C1qa, C1qb, Cyba, Cybb, Lgals3, Rsad2, C1qc, Tyrobp*	5.6
Complement pathway	3	*C1qa, C1qb, C1qc*	36

Table 5.4 contains the DAVID functional gene ontology annotations of the S40A/WT downregulated differentially-expressed genes at 8 hours. Benjamini-corrected p-values are all less than 0.05. FE=fold enrichment.

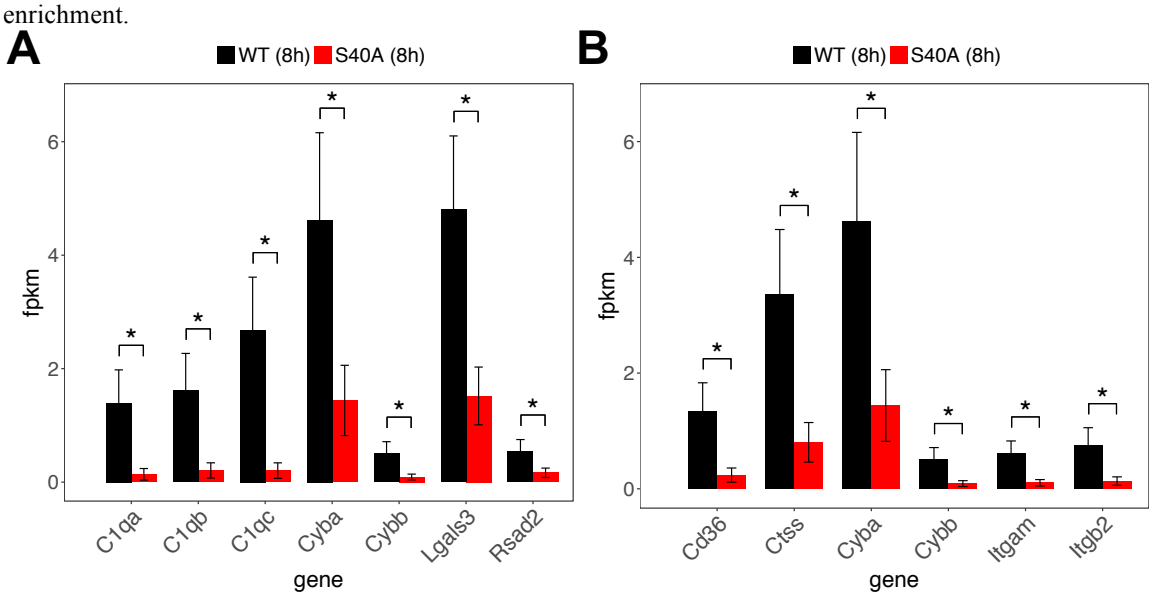


Figure 5.10 Downregulated genes in S40A/WT are involved in innate immune response and phagosome at 8 hours. (A) Innate immune response and (B) phagosome-related genes are downregulated in the S40A condition over the WT CREB condition. *FDR < 0.1.

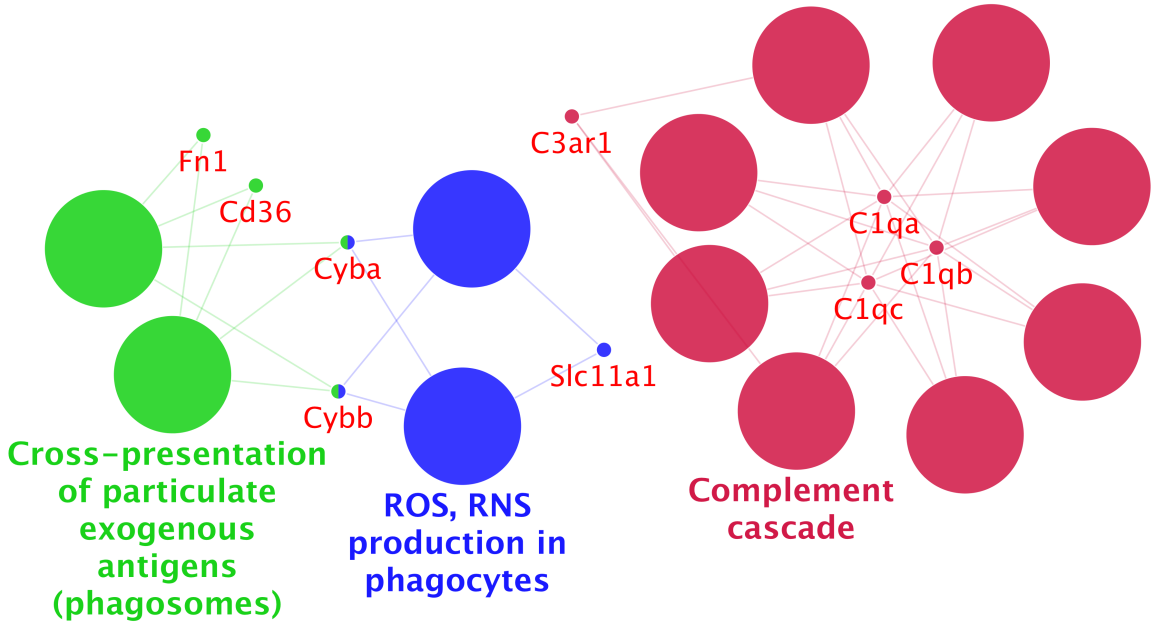


Figure 5.11 Gene ontology annotations for the downregulated genes in the S40A/WT comparison at 8 hours. Shown here are the top gene ontology categories for the downregulated genes in the S40A/WT comparison. The top categories include complement cascade innate immune response as well as phagosomal related categories. The genes that belong to and are shared by these gene ontology categories are also shown. The gene annotations were created performed using the Reactome Pathways database with the ClueGO extension in Cytoscape. The significance for all listed categories are Bonferroni step down corrected $P < 0.00046$.

5.6 Loss of CREB phosphorylation at serine 133 affects nervous system development at 8 hours

After exploring the glycosylation-related changes in transcription, we delved into the phosphorylation-dependent CREB gene expression changes. At 8 hours, abrogation of serine 133 phosphorylation resulted in increased expression of 17 genes (Table 5.5), which were enriched for genes involved in nervous system development ($P = 9.0\text{x}10^{-8}$), including wingless-type MMTV integration site family, member 7B (*Wnt7b*, 1.8-fold), netrin G2 (*Nntg2*, 2.2-fold), plexin D1 (*Plxnd1*, 1.7-fold), huntingtin-associated protein 1 (*Hap1*, 2.8-fold), dachshund homolog 2 (*Dach2*, 1.7-fold), and oligodendrocyte TFs 1 and 2 (*Olig1*, *Olig2*, 1.7- and 2.0-fold respectively) (Figures 5.12A, 5.13A, Table 5.6). In addition, the S133A/WT comparison displayed decreased expression of 40 genes (Table 5.5), which were enriched for genes found in the extracellular matrix ($P = 2.0\text{x}10^{-4}$) and important for cellular differentiation ($P = 4.0\text{x}10^{-2}$) (Figure 5.12B-C, Table 5.7). The DE extracellular matrix genes included aggrecan (*Acan*, 3.2-fold), collagen, type XIX, 1 (*Col19a1*, 1.9-fold), nephronectin (*Npnt* 1.8-fold), and *Lgals3* (3.9-fold) (Figure 5.12B, Figure 5.13B). Among the 18 downregulated genes were genes associated with differentiation into glia and other cell types, such as semaphorin 3C (*Sema3c*, 1.7-fold), eyes absent homolog 4 (*Eya4*, 2.1-fold), alanyl (membrane) aminopeptidase N (*Anpep*, 4.1-fold), and nephronectin (*Npnt*, 1.8-fold) (Figure 5.12C, Figure 5.13B). Because our primary cells were obtained from E16.5 cortices, our cell population was comprised of neural progenitor cells (NPCs) and neurons. We observe that loss of phosphorylation at serine 133 resulted in loss of pluripotency in this mixed cellular population and progression toward neuronal development.

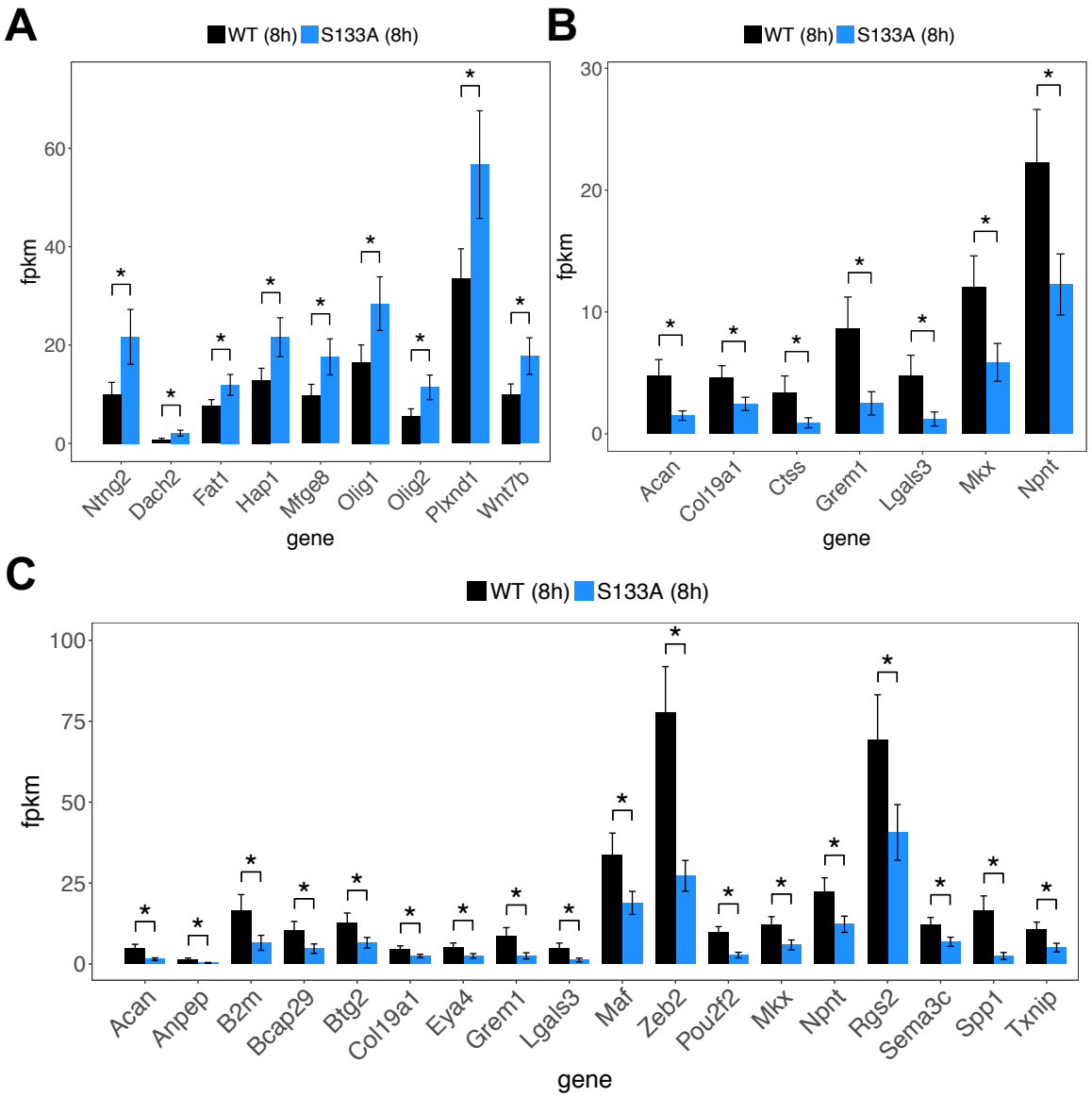


Figure 5.12 Differentially-expressed genes in S133A/WT at 8 hours are involved in neuronal differentiation and development. (A) Genes involved in the nervous system development are upregulated in the S133A-CREB condition over the CREB condition. (B) Extracellular matrix organization genes and (C) cell differentiation genes are downregulated in S133A-CREB condition compared to the CREB condition. *FDR < 0.1.

Table 5.5 List of differentially-expressed S133A/WT genes at 8 hours

Gene names	Gene description	S133A/WT log2(FC)	q-value
Col7a1	collagen, type VII, alpha 1(Col7a1)	6.4	0.004
Pdgfra	platelet derived growth factor receptor, alpha polypeptide(Pdgfra)	5.4	0.004
Rttn	rotatin(Rttn)	4.3	0.004
Rcan3	regulator of calcineurin 3(Rcan3)	3.2	0.004
Dach2	dachshund 2 (Drosophila)(Dach2)	1.5	0.07

Fn1	fibronectin 1(Fn1)	1.1	0.004
Ntng2	netrin G2(Ntng2)	1.1	0.01
Olig2	oligodendrocyte transcription factor 2(Olig2)	1.0	0.05
Aldh1l1	aldehyde dehydrogenase 1 family, member L1(Aldh1l1)	1.0	0.07
Mfge8	milk fat globule-EGF factor 8 protein(Mfge8)	0.86	0.10
Cebpb	CCAAT/enhancer binding protein (C/EBP), beta(Cebpb)	0.85	0.05
Wnt7b	wingless-type MMTV integration site family, member 7B(Wnt7b)	0.83	0.06
Olig1	oligodendrocyte transcription factor 1(Olig1)	0.78	0.08
Plxnd1	plexin D1(Plxnd1)	0.76	0.05
Hap1	huntingtin-associated protein 1(Hap1)	0.75	0.05
Fxyd6	FXYD domain-containing ion transport regulator 6(Fxyd6)	0.70	0.07
Fat1	FAT atypical cadherin 1(Fat1)	0.66	0.09
Nufip2	nuclear fragile X mental retardation protein interacting protein 2(Nufip2)	-0.69	0.09
Rgs2	regulator of G-protein signaling 2(Rgs2)	-0.77	0.09
Sema3c	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3C(Sema3c)	-0.81	0.07
Me3	malic enzyme 3, NADP(+)-dependent, mitochondrial(Me3)	-0.82	0.10
Maf	avian musculoaponeurotic fibrosarcoma oncogene homolog(Maf)	-0.83	0.04
Npnt	nephronectin(Npnt)	-0.86	0.02
Igfbpl1	insulin-like growth factor binding protein-like 1(Igfbpl1)	-0.88	0.05
Col19a1	collagen, type XIX, alpha 1(Col19a1)	-0.91	0.04
Btg2	B cell translocation gene 2, anti-proliferative(Btg2)	-1.0	0.08
Gabrd	gamma-aminobutyric acid (GABA) A receptor, subunit delta(Gabrd)	-1.0	0.03
Mkx	mohawk homeobox(Mkx)	-1.0	0.03
Eya4	EYA transcriptional coactivator and phosphatase 4(Eya4)	-1.1	0.07
Txnip	thioredoxin interacting protein(Txnip)	-1.1	0.03
Nts	neurotensin(Nts)	-1.1	0.08
Dlx6os1	distal-less homeobox 6, opposite strand 1(Dlx6os1)	-1.1	0.04
Hmgcll1	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase-like 1(Hmgcll1)	-1.1	0.05
Bcap29	B cell receptor associated protein 29(Bcap29)	-1.1	0.08
Lpl	lipoprotein lipase(Lpl)	-1.2	0.004
Zfp612	zinc finger protein 612(Zfp612)	-1.2	0.08
Prss35	protease, serine 35(Prss35)	-1.3	0.05
B2m	beta-2 microglobulin(B2m)	-1.3	0.05
Plin2	perilipin 2(Plin2)	-1.5	0.010
Zeb2	zinc finger E-box binding homeobox 2(Zeb2)	-1.5	0.004
Acan	aggrecan(Acan)	-1.7	0.004
H19	H19, imprinted maternally expressed transcript(H19)	-1.8	0.02

Grem1	gremlin 1, DAN family BMP antagonist(Grem1)	-1.8	0.02
Pou2f2	POU domain, class 2, transcription factor 2(Pou2f2)	-1.8	0.004
Ctss	cathepsin S(Ctss)	-1.9	0.07
Pdia4	protein disulfide isomerase associated 4(Pdia4)	-1.9	0.004
C1qc	complement component 1, q subcomponent, C chain(C1qc)	-2.0	0.09
Lgals3	lectin, galactose binding, soluble 3(Lgals3)	-2.0	0.06
Lyz2	lysozyme 2(Lyz2)	-2.0	0.02
Anpep	alanyl (membrane) aminopeptidase(Anpep)	-2.0	0.03
Tyrobp	TYRO protein tyrosine kinase binding protein(Tyrobp)	-2.5	0.10
Spp1	secreted phosphoprotein 1(Spp1)	-2.7	0.004
Gpnmb	glycoprotein (transmembrane) nmb(Gpnmb)	-2.8	0.004
Mirlet7b	microRNA let7b(Mirlet7b)	-3.1	0.02
Mmp12	matrix metallopeptidase 12(Mmp12)	-3.4	0.01
Cd36	CD36 antigen(Cd36)	-3.6	0.02
2010111I01Rik	RIKEN cDNA 2010111I01 gene(2010111I01Rik)	-3.9	0.004
Tlcd1	TLC domain containing 1(Tlcd1)	-5.1	0.004
Ttr	transthyretin(Ttr)	-5.5	0.08
Xist	inactive X specific transcripts(Xist)	-6.2	0.004

Table 5.5 shows the gene names, descriptions, $\log_2$(FC) where "FC" refers to fold-change, and q-values for the differentially-expressed genes in the S133A/WT comparison at 8 hours. The upregulated and downregulated genes are highlighted in pink and green respectively. q-values (FDR) < 0.1.

Table 5.6 PANTHER gene ontology classifications for S133A/WT upregulated genes

Gene ontology	Gene names	FE	P-value
Nervous system development	*Mfge8, Wnt7b, Fat1, Olig2, Ntng2, Hap1, Olig1, Plxnd1, Dach2*	18	9.1×10^{-8}
Ectoderm development	*Fat1, Olig2, Hap1, Olig1, Dach2*	17	2.2×10^{-3}
System development	*Fn1, Rttn, Pdgfra, Mfge8, Wnt7b, Cebpb, Fat1, Olig2, Ntng2, Hap1, Olig1, Plxnd1, Dach2*	12	2.5×10^{-7}
Developmental process	*Fn1, Rttn, Pdgfra, Mfge8, Wnt7b, Cebpb, Fat1, Olig2, Ntng2, Hap1, Olig1, Plxnd1, Dach2*	6.5	9.8×10^{-5}

Table 5.6 displays the major GO slim biological processes PANTHER gene ontology categories that are enriched in the S133A/WT upregulated genes at 8 hours. P-values are Bonferroni corrected. FE= fold enrichment.

Table 5.7 PANTHER gene ontology classifications for S133A/WT downregulated genes

Gene ontology	#	Gene names	FE	P-value
Extracellular matrix organization	7	*Ctss, Col19a1, Lgals3, Acan, Mkx, Grem1, Npnt*	23	2.0×10^{-4}
Extracellular structure organization	7	*Ctss, Col19a1, Lgals3, Acan, Mkx, Grem1, Npnt*	23	2.1×10^{-4}
Cell differentiation	18	*Spp1, Btg2, Col19a1, Lgals3, Zeb2, Bcap29, B2m, Acan, Eya4, Mkx, Rgs2, Txnip, Sema3c, Grem1, Anpep, Pou2f2, Maf, Npnt*	3.0	4.4×10^{-2}

Table 5.7 displays the major GO biological processes PANTHER gene ontology categories that are enriched in the S133A/WT downregulated genes at 8 hours. P-values are Bonferroni corrected. FE= fold enrichment.

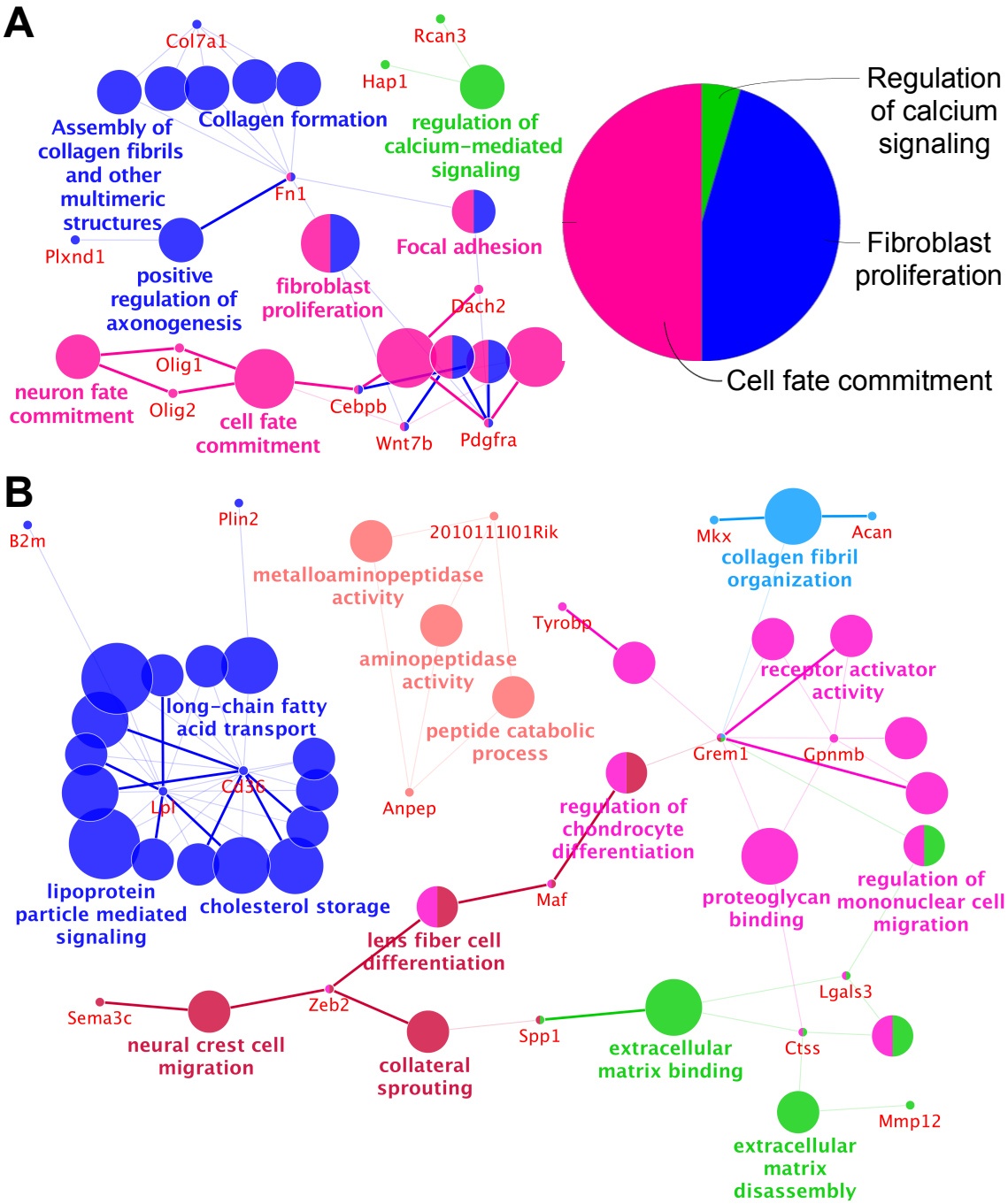


Figure 5.13 Cytoscape gene ontology annotations for the differentially-expressed genes in the S133A/WT comparison at 8 hours. (A) The upregulated genes are enriched for genes in involved in cell fate commitment and nervous and other system development. (B) The downregulated genes were enriched for lipid localization, extracellular matrix organization, neuronal development, and differentiation. The genes that belong to and are shared by these gene ontology categories are also shown. The gene annotations were created performed using the GO Biological Processes and Molecular Functions databases with the

ClueGO extension in Cytoscape. The significance for all listed categories are Bonferroni step down corrected $P < 0.015$.

5.7 The S40A-S133A double mutant affects nervous system development and the regulation of lipid localization

We next explored the DE genes from the glycosylation- and phosphorylation-deficient mutant (the S40A-S133A/WT comparison) (Table 5.8). In particular, genes related to nervous system development ($P = 4.7\text{x}10^{-5}$) were upregulated in the S40A-S133A DE genes, including fatty acid-binding protein, brain *(Fabp7,* 2.1-fold), *Olig1* (1.8-fold), *Nntg2* (2.0-fold), *Hap1* (1.7-fold), *Wnt7b* (1.8-fold), and *Dach2* (3.3-fold) (Figure 5.14A). In addition to their upregulation in the double mutant, *Olig1*, *Nntg2, Hap1, Wnt7b,* and *Dach1* were also upregulated in S133A, indicating that ablation of phosphorylation at serine 133 influences nervous system development independent of the CREB glycosylation state. The downregulated genes in the double mutant were enriched for genes involved in the positive regulation of lipid localization ($P = 3.0\text{x}10^{-2}$). Downregulated lipid localization genes included long-chain-fatty-acid-CoA ligase 5 (*Acsl5*, 2.1-fold), lipoprotein lipase (*Lpl*, 2.2-fold), and platelet glycoprotein IV (*Cd36*, 8.0-fold) (Figure 5.14B).

Table 5.8 List of differentially-expressed S40A-S133A/WT genes at 8 hours

Gene names	Gene description	S40A-S133A/ WT log2(FC)	q-value
Loxl2	lysyl oxidase-like 2(Loxl2)	7.4	0.004
Actr5	ARP5 actin-related protein 5(Actr5)	6.8	0.004
Parp1	poly (ADP-ribose) polymerase family, member 1(Parp1)	3.2	0.004
Dach2	dachshund 2 (Drosophila)(Dach2)	1.7	0.03
Tnc	tenascin C(Tnc)	1.5	0.004
Mfge8	milk fat globule-EGF factor 8 protein(Mfge8)	1.3	0.004
Fabp7	fatty acid binding protein 7, brain(Fabp7)	1.1	0.004
Apba3	amyloid beta (A4) precursor protein-binding, family A, member 3(Apba3)	1.1	0.05

Ntng2	netrin G2(Ntng2)	1.0	0.04
Fn1	fibronectin 1(Fn1)	1.0	0.004
Aldoc	aldolase C, fructose-bisphosphate(Aldoc)	1.0	0.04
Aldh1l1	aldehyde dehydrogenase 1 family, member L1(Aldh1l1)	1.0	0.06
Olig1	oligodendrocyte transcription factor 1(Olig1)	0.85	0.04
Kcnc4	potassium voltage gated channel, Shaw-related subfamily, member 4(Kcnc4)	0.83	0.09
Wnt7b	wingless-type MMTV integration site family, member 7B(Wnt7b)	0.82	0.06
Hap1	huntingtin-associated protein 1(Hap1)	0.78	0.04
Txnip	thioredoxin interacting protein(Txnip)	-0.90	0.09
Hmgcll1	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase-like 1(Hmgcll1)	-1.0	0.09
Acsl5	acyl-CoA synthetase long-chain family member 5(Acsl5)	-1.0	0.07
Akt1s1	AKT1 substrate 1 (proline-rich)(Akt1s1)	-1.0	0.004
Zfp612	zinc finger protein 612(Zfp612)	-1.1	0.10
Lpl	lipoprotein lipase(Lpl)	-1.1	0.004
Btg2	B cell translocation gene 2, anti-proliferative(Btg2)	-1.2	0.01
Mbp	myelin basic protein(Mbp)	-1.3	0.02
Nts	neurotensin(Nts)	-1.3	0.02
Prss35	protease, serine 35(Prss35)	-1.3	0.04
Zeb2	zinc finger E-box binding homeobox 2(Zeb2)	-1.5	0.004
Plin2	perilipin 2(Plin2)	-1.5	0.03
Grem1	gremlin 1, DAN family BMP antagonist(Grem1)	-1.6	0.05
Plxnb3	plexin B3(Plxnb3)	-1.7	0.06
Acan	aggrecan(Acan)	-1.7	0.004
Anpep	alanyl (membrane) aminopeptidase(Anpep)	-1.9	0.04
C1qc	complement component 1, q subcomponent, C chain(C1qc)	-1.9	0.08
H19	H19, imprinted maternally expressed transcript(H19)	-2.0	0.004
Lyz2	lysozyme 2(Lyz2)	-2.0	0.01
Lgals3	lectin, galactose binding, soluble 3(Lgals3)	-2.0	0.04
Gpnmb	glycoprotein (transmembrane) nmb(Gpnmb)	-2.5	0.004
Spp1	secreted phosphoprotein 1(Spp1)	-2.5	0.004
Cd36	CD36 antigen(Cd36)	-3.0	0.01
Itgb2	integrin beta 2(Itgb2)	-3.2	0.09
Mir344g	microRNA 344g(Mir344g)	-3.2	0.004
Enpp6	ectonucleotide pyrophosphatase/phosphodiesterase 6(Enpp6)	-3.4	0.04
Kif1c	kinesin family member 1C(Kif1c)	-3.8	0.004
2010111I01Rik	RIKEN cDNA 2010111I01 gene(2010111I01Rik)	-3.8	0.004
Mmp12	matrix metallopeptidase 12(Mmp12)	-4.1	0.007

A330023F24Rik	RIKEN cDNA A330023F24 gene(A330023F24Rik)	-5.3	0.004
Xist	inactive X specific transcripts(Xist)	-5.7	0.004

Table 5.8 Shown here are the gene names, descriptions, $\log_2$(FC) where "FC" refers to fold-change, and q-values for the differentially-expressed genes in the S40A-S133A/WT comparison at 8 hours. The upregulated genes are highlighted in pink while the downregulated genes are highlighted in green. q-values (FDR) < 0.1.

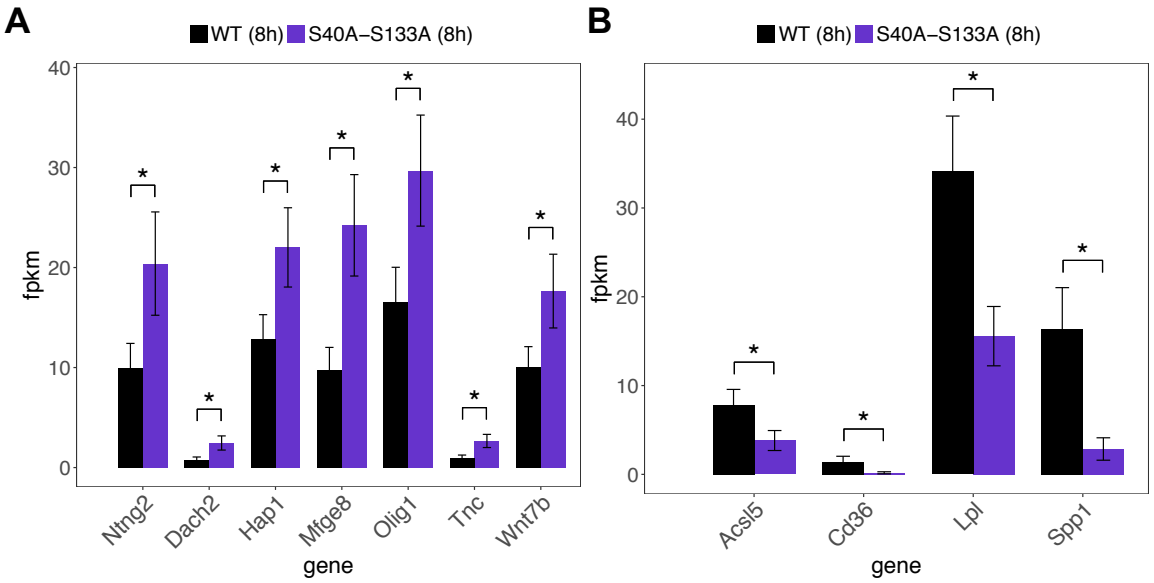


Figure 5.14 Differentially-expressed genes in S40A-S133A/WT at 8 hours are enriched for nervous system development and lipid localization genes. Genes involved in the (A) nervous system development and (B) positive regulation of lipid localization are upregulated in the S40A-S133A-CREB condition over the CREB condition. *FDR < 0.1.

5.8 CREB glycosylation and phosphorylation regulate different gene networks with the double mutant similar to the phosphorylation-deficient mutant

In addition to the three pairwise comparisons discussed earlier, there were 28 possible pairwise comparisons, which are represented in venn diagrams showing the number of common genes between the downregulated and upregulated genes for each pairwise comparison (Figures 5.15, 5.16). Overall, the largest differences could be observed when comparing 4 hours and 8 hours of expression, which is consistent with downstream second "wave" of transcription and late-phase LTP (L-LTP), the type of LTP that requires CREB-mediated transcription followed by translation.[2,20]

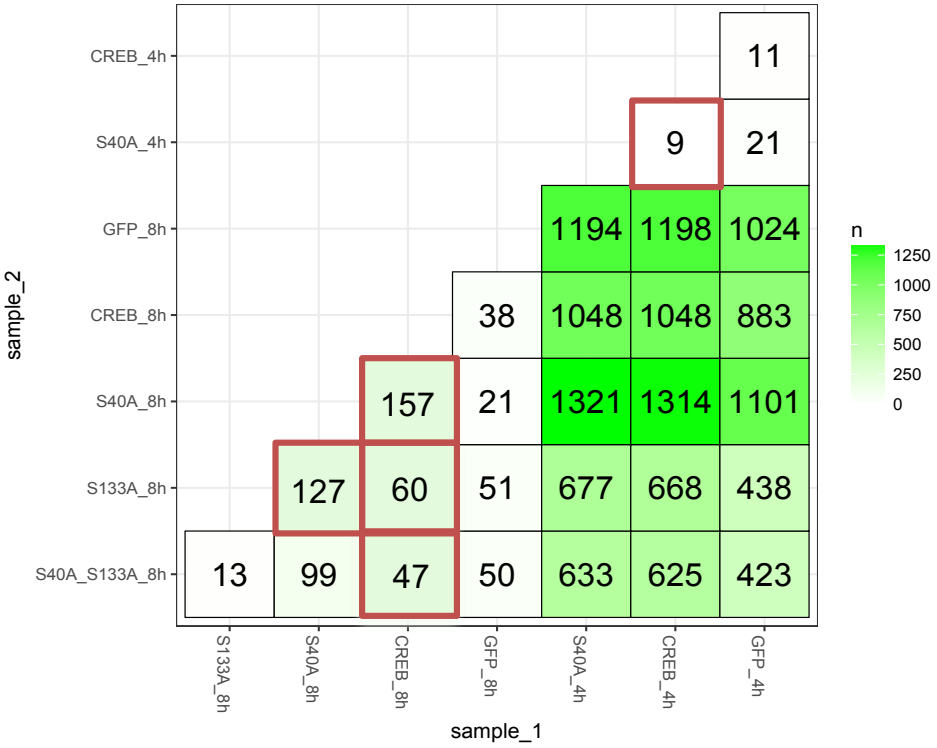
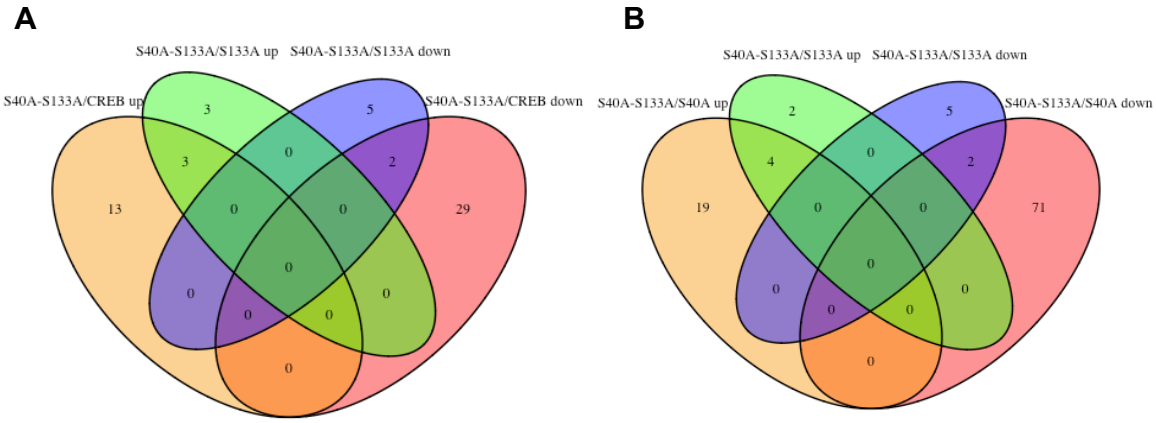


Figure 5.15 Differentially-expressed genes from the pairwise comparisons between different HSV treatment conditions. Following RNA-Seq and differential expression analysis using Cuffdiff, we obtained 28 different possible pairwise comparisons including the 4 hour and 8 hour time points. In the body of the paper, we focus on 5 of these comparisons: CREB_4h vs. S40A_4h, CREB_8h vs. S40A_8h, CREB_8h vs. S133A_8h, CREB_8h vs. S40A_S133A_8h, and S40A_8h vs. S133A_8h. These conditions that are explored are boxed in red. The q-value (FDR) cut-off used is 0.1 unless otherwise noted.



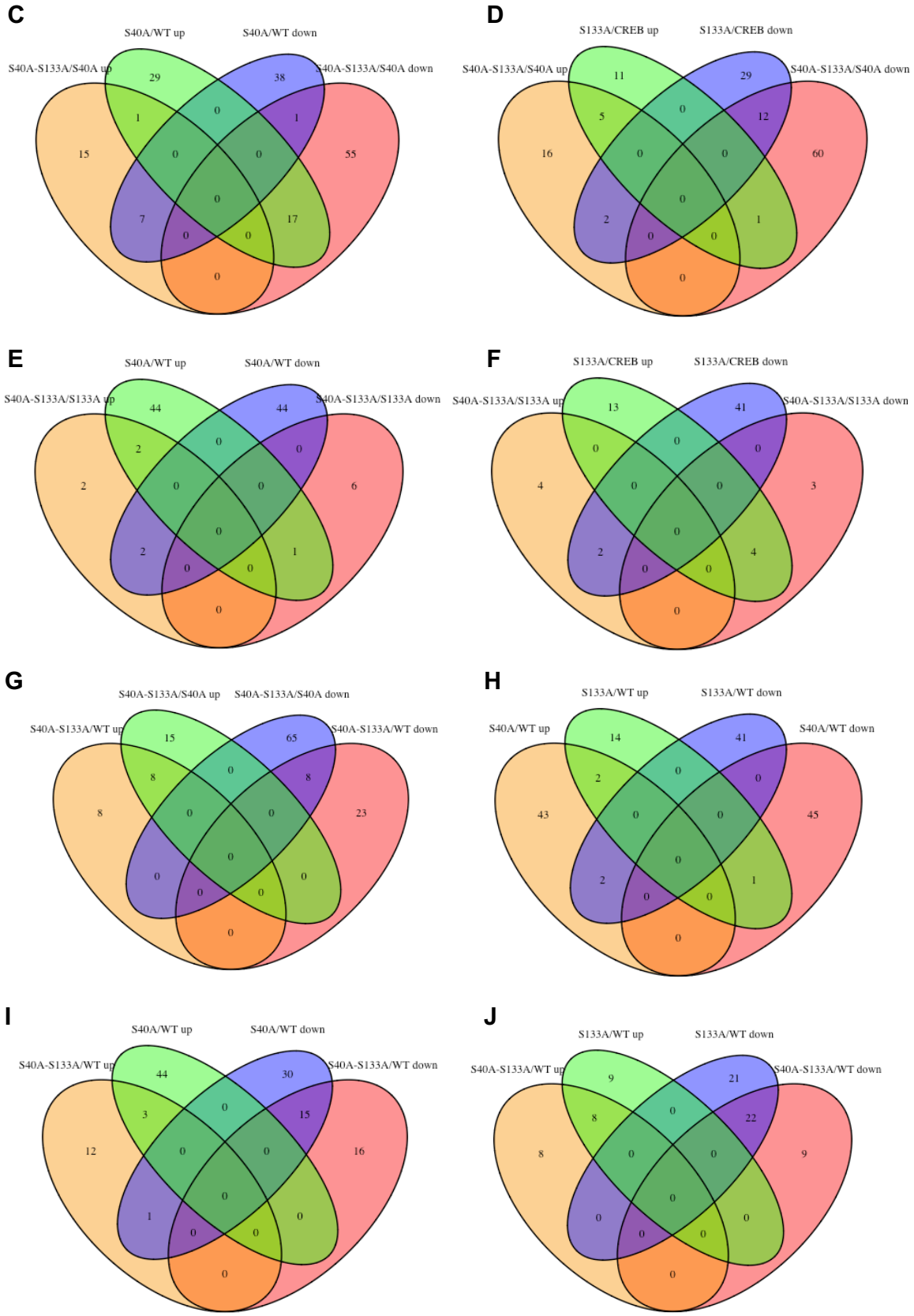

C
S40A/WT up
S40A/WT down
S40A-S133A/S40A up
S40A-S133A/S40A down
29
38
1
1
0
15
0
0
55
0
7
17
0
0
0
D
S133A/CREB up
S133A/CREB down
S40A-S133A/S40A up
S40A-S133A/S40A down
11
29
5
12
0
16
0
0
60
0
2
1
0
0
0
E
S40A/WT up
S40A/WT down
S40A-S133A/S133A up
S40A-S133A/S133A down
44
44
2
0
0
2
0
0
6
0
2
1
0
0
0
F
S133A/CREB up
S133A/CREB down
S40A-S133A/S133A up
S40A-S133A/S133A down
13
41
0
0
0
4
0
0
3
0
2
4
0
0
0
G
S40A-S133A/S40A up
S40A-S133A/S40A down
S40A-S133A/WT up
S40A-S133A/WT down
15
65
8
0
8
8
0
0
23
0
0
0
0
0
0
H
S133A/WT up
S133A/WT down
S40A/WT up
S40A/WT down
14
41
2
0
0
43
0
0
45
0
2
1
0
0
0
I
S40A/WT up
S40A/WT down
S40A-S133A/WT up
S40A-S133A/WT down
44
30
3
0
15
12
0
0
16
0
1
0
0
0
0
J
S133A/WT up
S133A/WT down
S40A-S133A/WT up
S40A-S133A/WT down
9
21
8
0
22
8
0
0
9
0
0
0
0
0
0

Figure 5.16 Venn diagrams showing pairwise CREB mutant comparisons at 8 hours. Shown here are the extra pairwise comparisons for the different CREB mutants separating upregulated and downregulated genes. (A) S40A-S133A/WT vs. S40A-S133A/S133A, (B) S40A-S133A/S40A vs. S40A-S133A/S133A, (C) S40A-S133A/S40A vs. S40A/WT, (D) S40A-S133A/S40A vs. S133A/WT, (E) S40A-S133A/S133A vs. S40A/WT, (F) S40A-S133A/S133A vs. S133A/WT, (G) S40A-S133A/WT vs. S40A-S133A/S40A, (H) S40A/WT vs. S133A/WT, (I) S40A-S133A/WT vs. S40A/WT, (J) S40A-S133A/WT vs. S133A/WT.

We also compared the DE genes in the S40A/WT, S133A/WT, and S40A-S133A/WT comparisons (Figure 5.17). Interestingly, there was no overlap between upregulated genes of the three mutant conditions, but there was some overlap between the downregulated genes. The downregulated genes that were common to all three pairwise comparisons were enriched for secreted proteins ($P = 2.3\text{x}10^{-2}$). The genes that were only downregulated in the S40A/WT comparison were enriched for immune response ($P = 3.2\text{x}10^{-3}$) genes. Neuronal excitability genes were exclusively upregulated in the S40A/WT condition and not in the S40A-S133A/WT and S133A/WT conditions. The pairwise comparisons between the S40A and S133A conditions show that S133A-CREB expression leads to enhanced expression of differentiation-related genes while S40A-CREB increases expression of calcium signaling, ion channel activity, and synaptic plasticity (Figures 5.9, 5.13). Finally, the S40A-S133A condition showed increased expression of nervous system development genes ($P = 4.7\text{x}10^{-5}$) similar to the S133A condition and decreased expression of genes involved in lipid localization ($P = 3.1\text{x}10^{-2}$) (Figures 5.13, 5.14). After comparing the double mutant with the single mutants, it appears that the double mutant displays more similarities with the S133A condition while the S40A condition alone enhances neuronal excitability. The DE genes suggest that glycosylation at serine 40 modulates distinct gene networks from phosphorylation at serine 133.

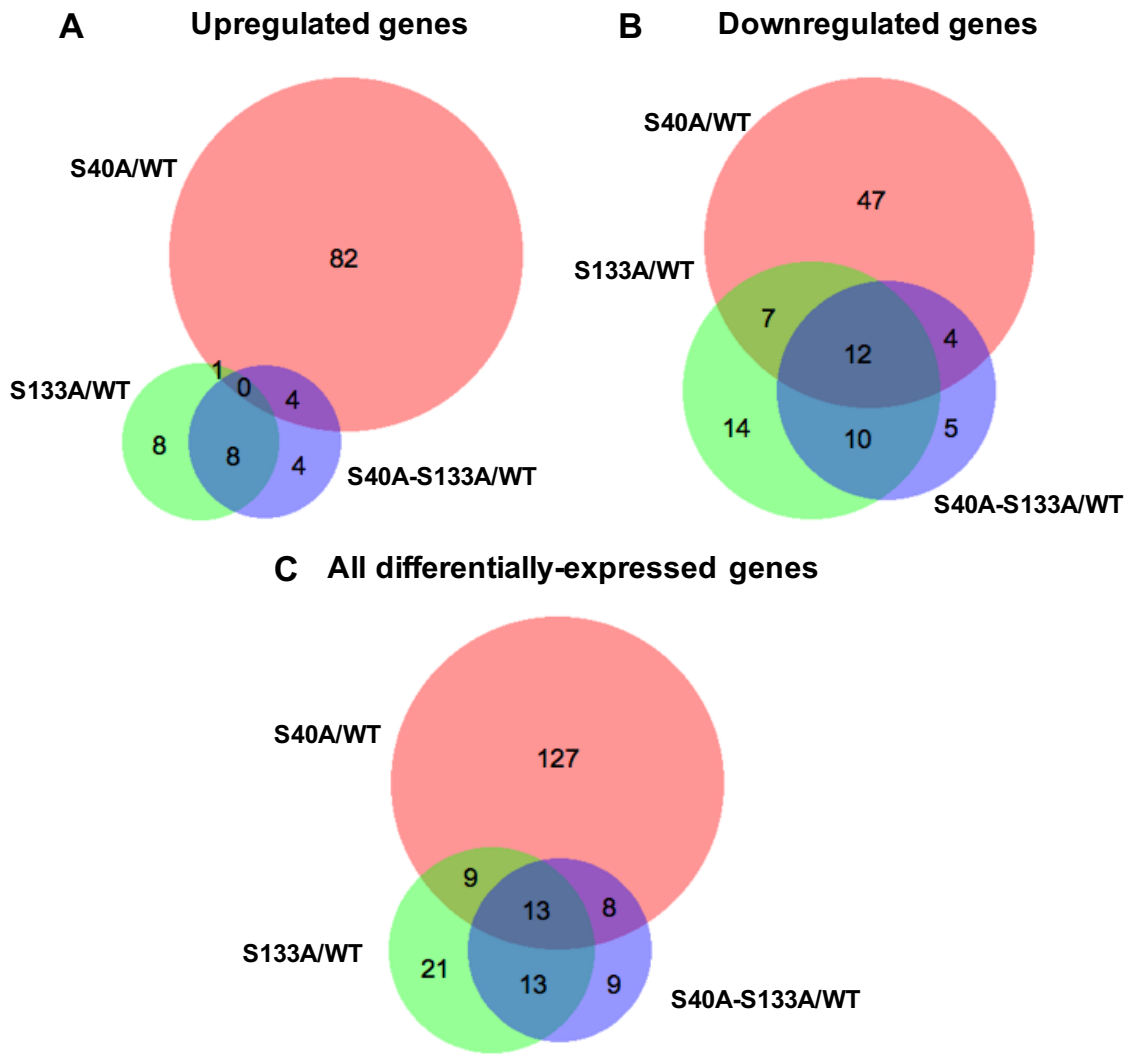


Figure 5.17 Venn diagrams showing overlap DE genes between various CREB mutants at 8 hours. These venn diagrams summarize the overlap between the DE genes in the following three conditions: S40A/WT CREB, S133A/WT CREB, and S40A-S133A/WT CREB. The venn diagrams show the overlap for (A) the upregulated DE genes, (B) the downregulated DE genes, and (C) all the DE genes. These venn diagrams show that there are some common genes downregulated by phosphorylation at S133 and glycosylation at S40, but no overlap between the upregulated genes. The Venn diagrams were generated using the BioVenn website.

5.9 CREB and its co-activators bind directly to DE gene promoters

We investigated whether CREB directly or indirectly regulated the DE genes identified from the S40A/WT, S133A/WT, and S40A-S133A/WT comparisons. CREB binds to full CRE sites (TGACGTCA) with high affinity or half CRE sites (TGACG/CGTCA) with slightly lower affinity.[21] Importantly, we found that 77-95% of

all the DE genes contained half CRE sites, suggesting that they are likely direct CREB targets (Table 5.9, Figure 5.18). Previous studies found no difference between the DNA binding capabilities of WT CREB and the S40A and S133A mutants, so we explored the WT CREB bound to these DE gene promoters in previous CREB ChIP-Seq studies.[3,22,23]

When we compared all our DE genes (S40A/WT, S133A/WT, S40A-S133A/WT) to CREB ChIP-Seq studies, we observed widely variable CREB ChIP-Seq enrichment on our genes across studies (0-48%). The CREB ChIP-Seq study performed in E16.5 mouse cortical neurons using the same conditions as ours (Kim *et al.* data set) yielded very low CREB ChIP-Seq binding to our DE gene promoters (0-6%).[5] A CREB ChIP-Seq study in mouse liver showed that 13-24% of DE genes were occupied by CREB while a CREB ChIP-Seq study using rat hippocampi had significantly higher overlap of 29-48%.[22,24] In the rat hippocampal CREB ChIP-Seq study, Lesiak and colleagues reanalyzed the mouse cortical study's CREB ChIP-Seq data (Kim *et al*. data set) using the same methods that they used for their ChIP-Seq analysis.[24] With this reanalyzed Kim *et al*. CREB ChIP-Seq peak data set (with FDR < 0.001), we found that about 8-17% of the DE genes were bound by the CREB. This coverage is closer to that observed in the CREB ChIP-Seq study from the mouse liver although the specific genes only show ~40% concordance for our DE genes when you compare the CREB occupancy from the E16.5 cortical neurons and mouse liver CREB ChIP-Seq studies (Figure 5.19). It is also important to note that we could have also used the rat hippocampal data set to compare our mouse data set as well since this has been done previously.[25] However, we decided to use the reanalyzed Kim data set at FDR < 0.001 because (1) we endeavored to utilize the ChIP-Seq data sets that were most similar to our conditions in terms of cell type and treatment and (2) the

FDR was sufficiently low to provide statistical confidence (FDR < 0.001). Thus, of all the DE genes, only 8-17% were occupied by CREB suggesting that many of the DE genes are downstream indirect targets of CREB potentially from the second "wave" of transcription following CREB activation (Table 5.10).[5,24]

Table 5.9 Full and half CRE sites on DE genes at 8 hours

Comparison	Up/Down	Full CRE	Half CRE	% CRE
S133A/WT	Up	1	20	80%
S133A/WT	Down	3	38	78%
S40A/WT	Up	1	37	77%
S40A/WT	Down	3	39	87%
S40A-S133A/WT	Up	1	20	95%
S40A-S133A/WT	Down	2	29	94%

Table 5.9 We used cruzdb to search 5000 bases upstream and 500 bases downstream of the transcription start sites in order to identify the total number of full and half CRE sites present for the DE genes. The reported %CRE gives the percent of genes that have either full or half CRE sites present. The number of full CRE sites includes the number of half CRE sites.

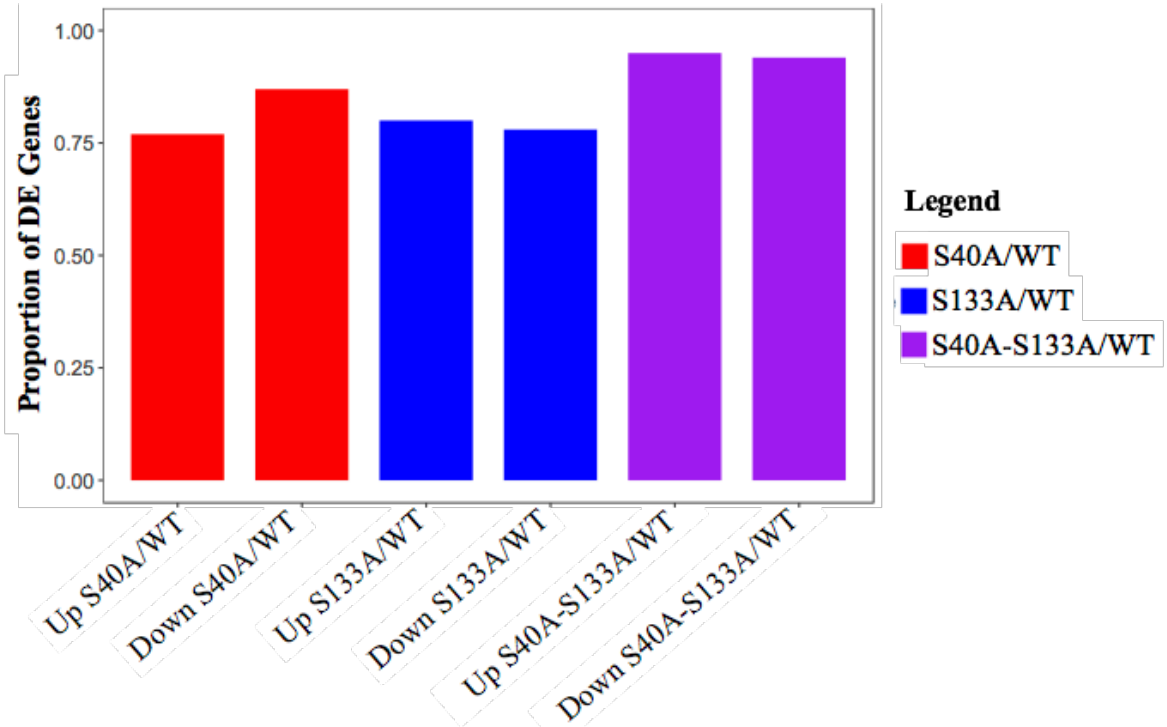


Figure 5.18 Barplot of half CRE sites on DE genes at 8 hours. The proportion of DE genes containing half CRE sites is shown in barplot format.

Table 5.10 CREB-regulated DE genes at 8 hours

	S133A/WT	S40A/WT	S40A-S133A/WT

Condition	% Up	% Down	% Up	% Down	% Up	% Down	Cell type
CBP_KCl[5]	52%	61%	60%	42%	57%	55%	CNs
CBP_noKCl[5]	4%	8%	2%	4%	10%	3%	CNs
CBP_KCl_dependent[5]	48%	53%	58%	38%	48%	52%	CNs
CREB[5,24]	8%	10%	17%	16%	14%	16%	CNs
CREB[22]	24%	18%	13%	13%	14%	19%	Liver
p300[26]	4%	10%	6%	13%	19%	3%	FB/MB
CRTC2[27]	8%	12%	8%	7%	10%	6%	3T3-L1
Up S133A/WT[28]	4%	4%	0%	0%	5%	5%	NA
Down S133A/WT[28]	0%	0%	2%	2%	0%	0%	NA
Up VP16-CREB/VP16[29]	8%	14%	42%	9%	10%	19%	HIPN
Down VP16-CREB/VP16[29]	12%	8%	35%	11%	5%	10%	HIPN
Up VP16-CREB/WT[16]	0%	8%	2%	11%	5%	10%	HIP
DownVP16-CREB/WT[16]	0%	0%	0%	0%	0%	0%	HIP

Table 5.10 shows an overview of the percent of differentially-expressed genes that are regulated by CREB and its co-activators. We compared our DE genes to ChIP-Seq studies (shown in blue) and microarray experiments (shown in purple). CN = E16.5 cortical neurons (DIV7 with KCl or without KCl), Liver = liver from mice fasted and then re-fed (males, age 8-12 weeks) FB/MB = E11.5 forebrain and midbrain, 3T3-L1 = 3T3-L1 mouse adipocytes treated with 10 μM forskolin, NA = overexpressing S133A or WT in nucleus accumbens (age 8 weeks), HIPN = overexpressing VP16-CREB or VP16 for 6 days in hippocampal neuronal cultures (10DIV), HIP = overexpressing VP16 or WT in hippocampus (age 8 weeks)

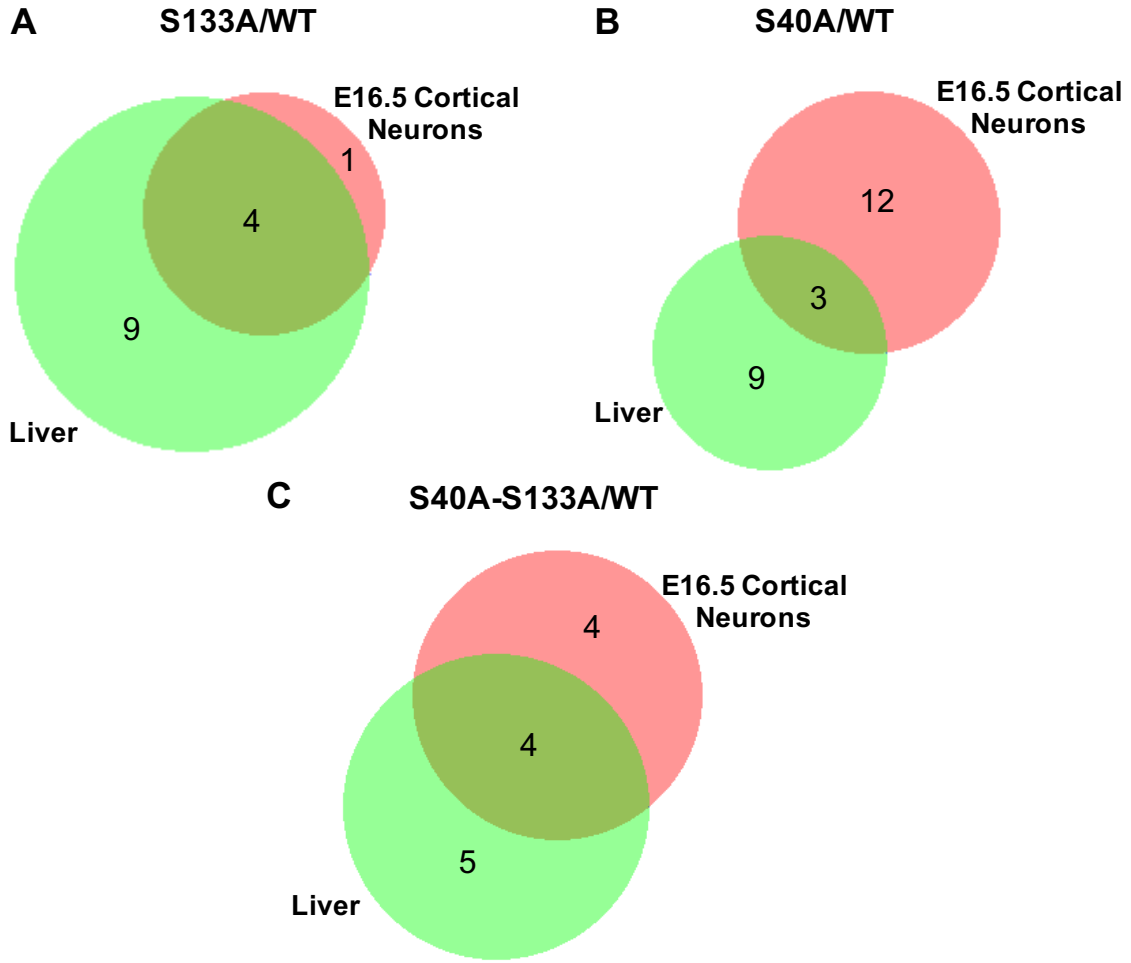


Figure 5.19 Overlap between CREB ChIP-Seq studies. Shown here are the overlapping DE genes that are occupied by CREB in E16.5 cortical neurons and 8-12 week livers (fasted and re-fed).

We next looked at whether other CREB-related proteins occupied the DE gene promoters.[5] Half of the DE gene promoters were occupied by the CREB coactivator CBP, but only under depolarization (not basal) conditions, indicating an activity-inducible CBP binding for these promoters (Table 5.10). The related CREB coactivator, p300, was found on fewer (3-19%) DE gene promoters in E11.5 brains.[26] For both p300 and CBP, CREB binding is mediated and encouraged by CREB phosphorylation at serine 133. However, we did not observe an overrepresentation of DE gene promoters occupied by p300 or CBP in the S133A/WT or S40A-S133A/WT conditions when compared to the S40A/WT condition. Given that CBP/p300 are coactivators for many TFs other than CREB and that the timing of our studies allows for secondary indirect CREB effects, the change in CBP or p300 binding cannot be solely attributed to direct CREB binding changes.[30] Finally, CRTC2, the major CREB glycosylation-dependent coactivator, was bound to 6-12% of DE gene promoters in 3T3-L1 mouse adipocytes.[27] We did not find an enrichment CRTC2 on the promoters of the DE genes in the S40A or S40A-S133A conditions when compared to the S133A condition likely for the same aforementioned reasons that we did not see differential binding of CBP and p300 in S133A mutant conditions. In summary, CREB and its coactivators were present on many of the S40A/WT, S133A/WT, and S40A-S133A/WT DE gene promoters.

5.10 Our study shows neuronal activity genes are upregulated by both VP16-CREB and S40A-CREB and minimal overlap between S133A and other studies exploring S133A-CREB gene changes

When we compared our DE genes to microarray and RNA-Seq studies expressing S133A mutant CREB, we found very little accordance except for a few key genes. First, McClung and colleagues showed that S133A-CREB overexpression results in opposing effects on transcription when compared to WT CREB in the nucleus accumbens of mice with a total of 24 differentially-expressed genes between the S133A-CREB and WT CREB conditions.[28] We first compared our DE genes to the 24 DE genes from a microarray study where McClung and colleagues overexpressed WT and S133A-CREB in the nucleus accumbens (NA).[28] Only one common gene was upregulated in both our and their S133A/WT comparisons: huntington-associated protein 1 (*Hap1*) (note: *Hap1* was not statistically significantly upregulated in the NA S133A/WT study).[28] Hap1 is associated with neuronal differentiation, signaling, and morphogenesis.[31] There was also one neuronal growth gene, *Bdnf*, shared between the downregulated S133A/WT NA genes and the S40A/WT upregulated genes.[28] Finally, *Fabp7* (fatty acid brain protein 7), a gene important for neuronal differentiation, was shared between the downregulated S133A/WT NA comparison and the upregulated S40A-S133A/WT comparison. This is supportive of the role of S133A in regulating expression of neuronal differentiation genes.

Overall, we did not observe significant overlap between the McClung and colleagues microarray study; potential reasons for the differences could be (1) that the study involved overexpression of S133A-CREB and WT CREB in the nucleus accumbens over course of 8 weeks while our study used a Creb1$^{\alpha\delta}$ background and explored considerably shorter time periods (4 and 8 hours) in a E16.5 cortical neuronal

population and (2) the differing sensitivity in detection for microarray studies when compared to RNA-Seq experiments, especially with low abundance transcripts.[32]

Another RNA-Seq study by Briand and coworkers showed that a phosphorylation-deficient S133A mutant mouse had no DE genes in the hippocampus when compared with WT mice.[23] In this study, breeding of the heterozygous S133A-CREB mice resulted in lower numbers than the expected Mendelian frequencies (only 11% homozygotes), leading the authors to suggest that while S133 phosphorylation may not affect transcription and memory-related behavior, it may affect development.[23] Prior studies have shown that CREB plays a vital role in neuronal development and differentiation.[2,33,34] Our results corroborate the paramount role of phosphorylation at S133 in arbitrating neuronal development and differentiation.

Next, we compared our DE genes to another microarray study that explored the effects of constitutively active VP16-CREB overexpression in the hippocampus of mice. The VP16-CREB mutant heightened neuronal activity, but lead to eventual excitotoxicity.[16] VP16-CREB-expressing mice displayed increased the expression of immune response genes and impaired spatial memory retrieval despite exhibiting enhanced LTP.[2,16,35] We found that *Bdnf* was upregulated in our S40A/WT and their VP16-CREB/WT comparisons, which is supportive of more active CREB leading to neuronal growth gene expression.[16] We found that the immune response related genes, *Spp1* and *C1qa,* were both upregulated in the VP16-CREB/WT comparison and downregulated all of the comparisons S133A/WT, S40A/WT, and S40A-S133A/WT. This suggests that the glycosylation and phosphorylation mutants may have impaired transcription of immune response genes while the VP16-CREB mouse might have

enhanced expression likely due to the presence of excitotoxic effects in the VP16-CREBexpressing mouse. Altogether, the comparison with the VP16-CREB mouse shows that loss of glycosylation at serine 40 leads to similar enhancements in neuronal activity-related genes such as *Bdnf* and dissimilar decreases in immune response-related gene expression like *Spp1* and *C1qa*. Importantly, the VP16-CREB overexpression analysis only yielded 4 differentially-expressed genes total across all time points explored (1-5 weeks expression), and the only gene discussed here that was differentially-expressed in the VP16-CREB mice was *Bdnf*.

Finally, another study conducted by Benito and colleagues explored the effects of overexpressing constitutively active (ca) VP16-TF fusions (CREB, EGR1, FOS, SRF) in regulating transcription in cultured mouse hippocampal neurons. Briefly, wildtype or hippocampal neurons were treated with (1) lentivirus expressing GFP and caCREB, caEGR1, caFOS, caSRF, or VP16 (control) for 6 days or (2) various stimuli including forskolin (fors), bicuculline (bic), and BDNF.[29] The DE genes for each of these conditions and our mutant CREB conditions are provided in Table 5.11 (percentage of total DE genes are reported in Table 5.10). Through a comparison of the caCREB/VP16 and S40A/WT conditions, we sought to determine which DE genes are either expressed (1) similarly in response to heightened CREB activity (DE genes shared by both conditions) (type 1) or (2) in opposite directions in the caCREB and S40A-CREB conditions (DE genes that show contrasting expression in the two conditions) (type 2). Critically, the VP16-CREB fusion facilitates the association with the coactivator, CBP, whereas S40A-CREB facilitates the association with the coactivator, CRTC, and potentially CBP as well (because S40A-CREB can be phosphorylated at S133).

Therefore, any differences between VP16-CREB- and S40A-CREB-mediated expression may reflect the different genes that are regulated by CRTC specifically.

First, we looked at the overlap between the S40A/WT and caCREB/VP16 conditions in order to see which genes were differentially expressed in response to elevated CREB activity (type 1 genes). The upregulated genes in the caCREB/VP16 and the S40A/WT conditions were enriched for genes involved in chemical synaptic transmission ($P < 6.3\text{x}10^{-3}$) including *Bdnf*, *Cartpt* (cocaine- and amphetamine-related transcript protein), *Ache* (acetylcholinesterase), *Crh* (corticotropin releasing hormone), *Lynx1* (Ly6/neurotoxin 1), and *Cplx1* (complexin 1). This implicates enhanced CREB activity in elevating the expression of synaptic activity genes, which is consistent with the well-established role of CREB in mediating synaptic and neuronal activity.[20] Interestingly, many of these upregulated genes are also upregulated in response to forskolin, suggesting that increased cAMP signaling can induce the expression of some of the same neuronal activity genes that are induced in response to caCREB and S40A-CREB. The genes that were downregulated in the S40A/WT and caCREB/VP16 conditions were involved in neuronal axonal growth and motility (*Nefm*, *Islr2*, and *Gpc2*) and oxidative stress response (*Carhsp1* and *Txnip*). Altogether, these common DE genes likely reflect the genes regulated by CREB and both of the coactivators, CBP and CRTC.

We next explored the DE genes that displayed opposing expression in the S40A/WT and caCREB/VP16 comparisons (type 2 genes). The S40A/WT condition showed elevated expression of ion and solute transport genes ($P < 7.5\text{x}10^{-3}$) such as *Slc6a1*, *Cacna1i*, *Kcns2*, *Slc6a7*, and *Kcnh3*, while the caCREB/VP16 condition revealed downregulation of these genes. Specifically, GAT1 (*Slc6a1*) is the most prevalent brain-

specific GABA (gamma-aminobutyric acid) transporter that removes GABA from synapses thereby enhancing excitatory signaling.[36] These results indicate that S40A-CREB specifically strengthens whereas caCREB diminishes neuronal excitability through ion channel expression. Furthermore, genes involved in the regulation of cell adhesion ($P < 6.1\text{x}10^{-3}$) and cytokine production ($P < 3.6\text{x}10^{-2}$) were downregulated in S40A/WT, but upregulated in caCREB/VP16 conditions. This suggests that immune response and cell adhesion gene expression was downregulated in the S40A-CREB condition and upregulated in the caCREB/VP16 condition. Altogether, this comparative analysis suggested that enhanced CREB activity (through either S40A-CREB or caCREB expression) facilitates the expression of genes involved in neuronal activity. Unlike caCREB, ablation of CREB glycosylation specifically upregulated genes that enhance neuronal excitability and downregulated genes involved in immune response and cell adhesion likely through specific enhancement of CRTC binding capabilities.

In order to check to see if these DE genes were mediated by CBP or CRTC, we explored the ChIP-Seq data sets for the genes. Overall, CBP could be found on the promoters of 15 of the similarly expressed (type 1) genes and 13 of the differently expressed (type 2) genes in mouse cortical neuronal cultures.[5] In contrast, CRTC1 was solely found on 8 of the differently expressed (type 2) genes and none of the similarly expressed (type 1) genes in drosophila mushroom bodies in response to learning.[37] This supports the hypothesis that the differently expressed genes between the S40A-CREB and VP-CREB reflect genes that are regulated by CRTC rather than CBP. However, ChIP-Seq in mouse neurons rather than drosophila would more firmly establish this connection.

Table 5.11 Differentially-expressed genes in the Benito study and our study

Gene	caCREB	caFOS	caEGR1	caSRF	bic	fors	BDNF	S133A	S40A	S40A-S133A
Spock3	2.86								1.55	
Ptprn	2.49					2.16			1.41	
Bcap29	2.35							-2.19		
Scg2	2.31	1.8			1.49	2.44			1.43	
Crh	2.26								2.14	
Kcnc4	2.26								1.57	1.78
Gpnmb	2.18							-6.92	-6.10	-5.68
Ache	2.15								1.42	
Bdnf	2.12	2.08			1.66	1.75			1.58	
Gng4	2.03		-2.18			1.3			1.66	
Fn1	2.02					1.73		2.19	-1.71	2.04
Lynx1	1.93								1.80	
Cntnap1	1.9								1.67	
Btg2	1.9				6.01			-1.94		-2.34
Gda	1.81								1.39	
Nrsn1	1.78								1.69	
Prss35	1.69							-2.50		-2.51
Rcan2	1.54								1.56	
Car4	1.5								3.87	
Kcnh1	1.5	-1.4							1.47	
Acan	1.49							-3.17		-3.28
Nckap1l	1.49								-2.93	
Ngef	1.45								1.51	
Zfp612	1.43							-2.30		-2.12
Cplx1	1.43								1.41	
Acsl5	1.41									-2.05
Cebpb	1.4							1.80		
Eya4	1.39							-2.08		
Cartpt	1.39					1.77			4.43	
*Shb**	1.36	1.68							-1.43	
Zbtb4	1.34								1.43	
Shh	1.27								1.54	
Parp1	-1.29					1.39			2.22	9.10
*Osbpl1a**	-1.31								1.46	
*Kcns2**	-1.32								1.79	
2310003H01Rik	-1.34							1.41		
Cacna1i	-1.35								1.57	
Bai2	-1.42							1.16		
Fxyd6	-1.45		-1.47					1.62		
*Crhr1**	-1.46	1.63							1.46	
Fam81a	-1.47								1.46	
D3Bwg0562e	-1.47								1.42	
*Negr1**	-1.47								1.38	
Nefm	-1.48								-1.46	
Ildr2	-1.59	1.43							1.49	

Creg2	-1.68					-1.26			2.02	
*Itpka**	-1.72					-1.25			2.35	
Zeb2	-1.75							-2.85		-2.78
Slc6a7	-1.75								1.65	
Gpc2	-1.75								-1.54	
Phyhip	-1.76								1.51	
*Kcnh3**	-1.81								1.61	
A830018L16Rik	-1.83								1.46	
Txnip	-1.91					-2.48		-2.08	-1.93	-1.87
Islr2	-1.91								-1.42	
Sema3c	-1.94		-2.69					-1.75		
Carhsp1	-2.02								-1.64	
Camk2a	-2.04								1.58	
*Slc6a1**	-2.14								1.46	
Lpl	-2.73	-2.52		-2.53			-1.72	-2.25		-2.19
Ntng2							-1.86	2.17		2.04
Mfge8						-2.36		1.81		2.49
Plxnd1						1.21		1.69		
Rgs2					2.59			-1.70		
Npnt		-1.66				1.2		-1.82		
Cckbr		1.55							2.69	
Rasal1			1.98	1.56					1.67	
Ncald						-1.16			1.51	
Nrip3							1.82		1.48	
AI593442						-1.49			1.45	
Grik3				1.83					-1.39	
Bcl2l11					1.25				-1.45	
Sox11		-1.82			1.27				-1.46	
Fst		-1.36							-1.49	
Tnc						1.69				2.75
Aldoc						-2.04				2.03
Akt1s1						1.28				-2.06
Itgb2		1.41							-5.66	-9.13
Mmp12				1.57				-10.5	-24.8	-16.9
Nab2			5.53	1.58					1.75	

Table 5.11 indicates the fold changes for various conditions in the Benito and colleagues study and for our CREB mutants in our study. The results are color-coded based on whether the genes are upregulated (red) or downregulated (green). The hippocampal neurons were treated with (1) lentivirus expressing GFP and constitutively active (ca) caCREB, caEGR1, caFOS, and caSRF for 6 days (normalized to lentivirus expressing the VP16 domain and GFP) or (2) various stimulation conditions including forskolin (fors), bicuculline (bic), and BDNF (normalized to vehicle). In our conditions, embryonic cortical neurons were treated for 8 hours with S40A, S133A, and S40A-S133A mutant CREB (normalized to WT CREB). *genes that are bound by CRTC1 in response to learning in drosophila (Hirano and colleagues, 2016).

5.11 DE gene promoters are occupied by activating histone modifications

After identifying several CREB-related activators were present across the DE gene promoters, we assessed the overall chromatin accessibility across all of the S40A/WT, S133A/WT, and S40A-S133A/WT DE gene promoters at 8 hours (both upregulated and downregulated). We found that the promoters of these DE genes were mostly associated with activating histone modifications such as H3K4me$_1$, H3K4me$_3$, and H3K27Ac and less associated with the repressive histone modification H3K27me$_3$ in E13.5 basal ganglia (BG).[38] For the activating histone modifications, we found H3K4me$_1$, H3K4me$_3$, and H3K27Ac on 42-57%, 37-44%, and 15-38% of the DE gene promoters respectively (Table 5.12, Figure 5.20).[38] In contrast, the repressive histone modification H3K27me$_3$ occupied 6-20% of the DE gene promoters across all comparisons.[38] Few of our DE genes have bivalent promoters, which are modified by both the activating H3K4me$_3$ and repressive H3K27me$_3$ modifications (0-6%). Bivalent promoters occur during differentiation and indicate that the gene is poised to change to either an activating or repressive state.[39] Under our conditions, most of our DE gene promoters contain univalent modifications and are therefore committed to either repressive or activating cell fates. Upon the DE gene promoters, we also observed the presence of activating gene-body associated H3K79me$_2$ and H3K36me$_3$ histone modifications, which are associated with cell cycle (highest at G2 phase) and actively transcribed exons respectively.[40,41] Based on the predominance of activating histone modifications, both the upregulated and downregulated DE genes are generally found in euchromatic and more transcriptionally active regions of the genome.

Table 5.12 Histone code for DE gene promoters at 8 hours

	S133A/WT		**S40A/WT**		**S40A-S133A/WT**		
Condition	**% Up**	**% Down**	**% Up**	**% Down**	**% Up**	**% Down**	**Cell type**

H3K4me$_3$[42]	64%	63%	83%	56%	71%	58%	ESCs
H3K4me$_3$[43]	60%	45%	31%	27%	57%	23%	dNPCs
H3K27me$_3$[43]	0%	8%	15%	2%	5%	6%	dNPCs
Both H3K4me$_3$ and H3K27me$_3$[43]	0%	4%	2%	0%	5%	6%	dNPCs
H3K4Me$_1$[5]	48%	57%	56%	42%	57%	52%	CNs
Increasing H3K27Ac with KCl[44]	16%	6%	6%	0%	10%	3%	CNs
Decreasing H3K27Ac with KCl[44]	4%	10%	6%	7%	10%	10%	CNs
Constant H3K27Ac with KCl[44]	8%	4%	8%	7%	19%	6%	CNs
Total H3K27Ac[44]	28%	20%	21%	13%	38%	19%	CNs
H2B-S112-O-GlcNAc[45]	44%	20%	33%	42%	52%	29%	ESCs
H3K4me3[46]	32%	45%	56%	56%	48%	48%	BM
H3K27me$_3$ gain with neuronal differentiation[47]	8%	0%	10%	2%	0%	0%	CNs/NPCs
H3K27me$_3$ loss with neuronal differentiation[47]	12%	8%	8%	4%	10%	6%	CNs/NPCs
H3K27me$_3$[38]	12%	16%	15%	20%	14%	19%	BG
H3K4me$_1$[38]	48%	35%	38%	33%	33%	39%	BG
H3K27Ac[38]	36%	37%	40%	40%	33%	35%	BG
H3K4me$_3$[38]	44%	37%	40%	42%	38%	39%	BG
Tet2, OGT, H2B-S112-O-GlcNAc[45]	16%	4%	17%	18%	19%	6%	ESCs
Tet2, O-GlcNAc, H3K4me$_3$[46]	0%	8%	0%	2%	5%	6%	BM
H3K79me$_2$[48]	36%	18%	4%	20%	24%	16%	ESCs
H3K36me$_3$[48]	20%	14%	2%	16%	14%	13%	ESCs

Table 5.12 shows an overview of the percent of the promoters of the differentially-expressed genes that are regulated by various histone modifications identified in ChIP-Seq experiments. We identified which of our DE gene promoters were occupied by activating (pink), repressive (green), and bivalent (blue) histone modifications. ESCs = v6.5 embryonic stem cells (E3.5), dNPCs = v6.5 embryonic stem cells differentiated into neural progenitor cells, CN = E16.5 cortical neurons (DIV7 with KCl or without KCl), BM = bone marrow, CNs/NPCs = FACS sorted E15.5 neural progenitor cells and neurons, BG = E13.5 basal ganglia

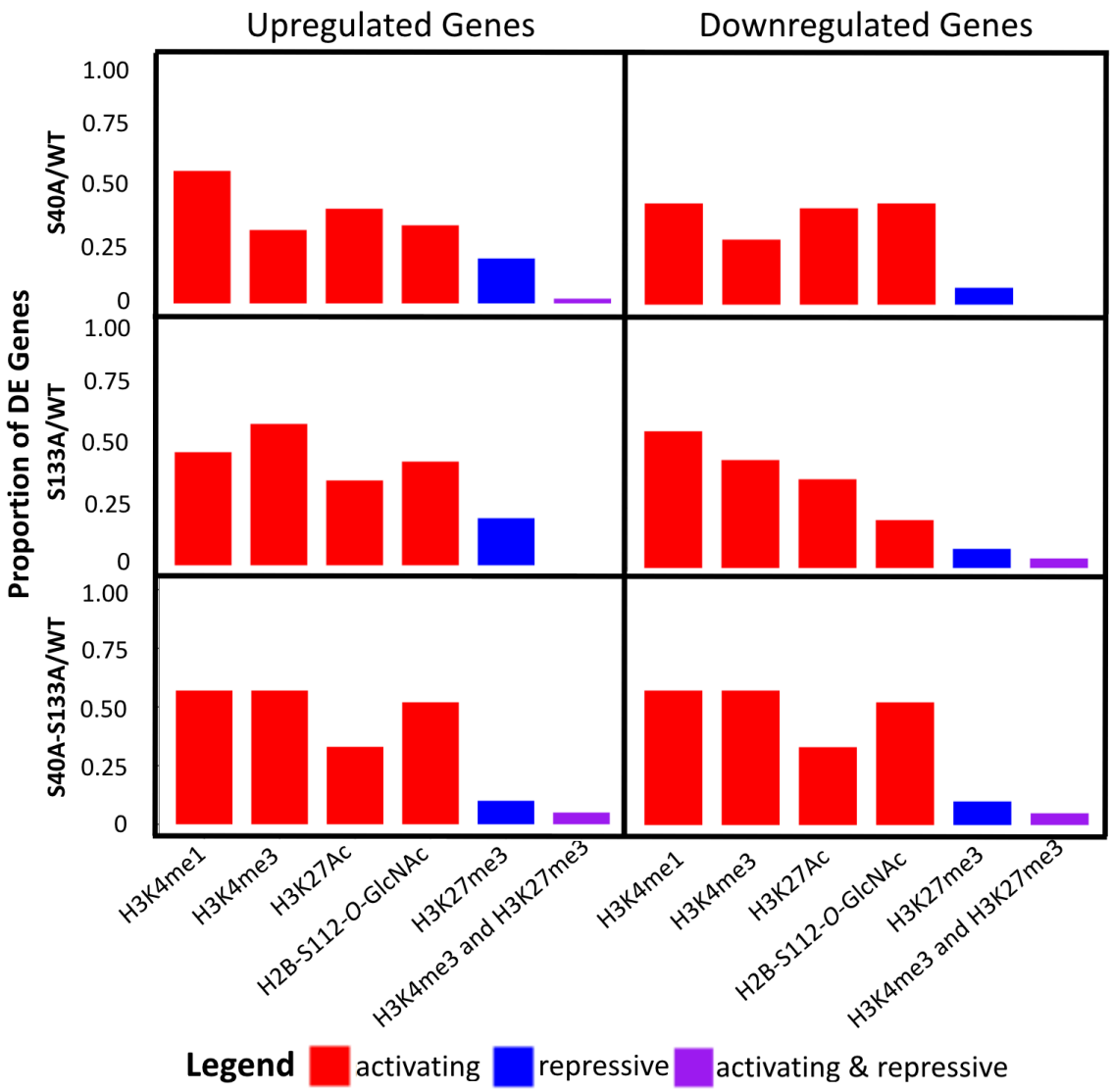


Figure 5.20 Histone modifications associated with the promoters of S40A/WT, S133A/WT, and S40A-S133A/WT DE genes. Shown here are the proportion of the promoters of DE genes occupied by either activating, repressive, or activating and repressive histone modifications including H3K4me$_1$[5], H3K4me$_3$[43], H3K27Ac[38], H2B-S112-*O*-GlcNAc[45], H3K27me$_3$[47], and bivalent modified genes (occupied by both H2K4me$_3$ and H3K27me$_3$)[43]. Overall, we observed that activating (red) more so than repressive (blue) histone modifications are associated with the 8-hour DE gene promoters. Few DE genes were occupied by bivalent, poised histone modifications (purple).

5.12 OGT, *O*-GlcNAc, and OGT-associated proteins and DNA modifications regulate the S40A/WT and S133A/WT DE genes at 8 hours

We then focused on the glycosylation-deficient mutant DE genes and found that the S40A/WT DE genes at 8 hours were regulated by *O*-GlcNAc and OGT directly. In a ChIP-seq study in mouse embryonic stem cells (ESCs), we found many of these gene

promoters were bound by OGT (65% of upregulated (up); 38% of downregulated (down) genes) as well as a variety of other TFs known to associate with OGT including the mSin3a corepressor (40% up; 44% down) and polycomb repressor complex 2 (PRC2) proteins EZH2 (46% up; 9% down) and Suz12 (31% up; 7% down) (Table 5.13).[42,45,46,49] The function and activity of the polycomb repressor complex 2 has been shown to be necessary for regulating *O*-GlcNAcylation throughout cells.[50,51] Furthermore, *O*-GlcNAcylation of RING1B, a major subunit of PRC1, increases the targeting of RING1B to neuronal genes and decreases the association with cell cycle genes in ESCs.[52] We observed 6-24% of DE gene promoters are occupied by RING1B in ESCs (Table 5.13). OGT also *O*-GlcNAc modifies EZH2 and the pluripotency regulating TFs, Sox2 and Oct4, and mediates their function.[53-55] In mouse bone marrow cells, *O*-GlcNAc (13% of upregulated and 16% of downregulated genes) was found on these S40A/WT DE gene promoters as well (Table 5.13).[46]

Table 5.13 OGT-related protein bound to DE gene promoters at 8 hours

	S133A/WT		**S40A/WT**		**S40A-S133A/WT**		
Condition	**% Up**	**% Down**	**% Up**	**% Down**	**% Up**	**% Down**	**Cell type**
O-GlcNAc[46]	12%	16%	13%	16%	19%	16%	BM
OGT[45]	32%	24%	33%	33%	38%	35%	ESCs
OGT[42]	72%	41%	65%	38%	48%	32%	ESCs
Tet1[42]	72%	71%	98%	76%	86%	71%	ESCs
Tet2[45]	32%	47%	46%	42%	29%	39%	ESCs
Tet2[46]	4%	20%	0%	11%	14%	16%	BM
Tet3[56]	4%	0%	2%	0%	5%	3%	CNs/ NPCs
5hmC coverage in NPC[47]	32%	18%	21%	24%	29%	13%	CNs
5hmC coverage in neurons[47]	44%	14%	27%	22%	38%	16%	NPCs
5hmC loss with neuronal differentiation[47]	4%	4%	4%	2%	5%	0%	CNs/ NPCs
5hmC gain with neuronal	20%	4%	10%	9%	19%	6%	CNs/ NPCs

differentiation[47]							
5hmC in either neurons or NPCs[47]	48%	20%	38%	29%	38%	16%	CNs/ NPCs
mSin3a[42]	48%	33%	75%	44%	57%	35%	ESCs
mSin3a[57]	36%	37%	40%	44%	43%	39%	pre-iPSCs
Ezh2[49]	32%	31%	46%	9%	24%	16%	ESCs
Suz12[48]	24%	22%	31%	9%	14%	10%	ESCs
Suz12[49]	24%	24%	31%	7%	14%	13%	ESCs
Ring1B[49]	24%	14%	17%	7%	14%	6%	ESCs
Oct4[48]	16%	16%	15%	27%	10%	16%	ESCs
Sox2[48]	16%	14%	15%	24%	19%	13%	ESCs
Nanog[48]	24%	8%	10%	16%	19%	10%	ESCs

Table 5.13 shows is an overview of the percent of differentially-expressed gene promoters that are bound by different OGT-related proteins and modifications and OGT and *O*-GlcNAc itself. We found the percentage of our DE gene promoters bound or modified by OGT and *O*-GlcNAc (green), Tet proteins or their products (orange), known OGT interactors including polycomb repressor complex 2 proteins (pink), polycomb repressor 1 complex proteins (purple), and pluripotency-related proteins (blue). All results come from ChIP-Seq experiments except for the 5hmC results, which are derived from hMeDIP experiments. BM = bone marrow, ESCs = v6.5 embryonic stem cells (E3.5), CNs/NPCs = FACS sorted E15.5 neural progenitor cells and neurons, pre-iPSCs = neural stem cell (NSC)-derived pre-induced pluripotent stem cells

In particular, we discovered that the promoters of the upregulated S40A/WT genes were overwhelmingly occupied by the OGT-associated Tet family of proteins [Tet1 (98%) in ESCs; Tet2 (46%) in ESCs; Tet3 (2%) in CNs/NPCs (E15.5 cortical neurons and neural progenitor cells)] (Table 5.13, Figure 5.21).[45,47,56] The Tet (ten-eleven translocation methylcytosine dioxygenase) family catalyzes the conversion of 5-methylcytosine (5mC) to 5-hydroxy-methylcytosine (5hmC), which are the first steps in the removal of 5-methylcytosine, a modification associated with gene silencing.[47] The 5hmC modification and Tet activity have been shown to be important for regulating pluripotency, neuronal activity, and gene activity.[45,46] When compared to a hMeDIP experiment that measured the occupancy of 5hmC, we found that the S40A/WT upregulated genes were 5hmC modified (38%) in CNs/NPCs (Table 5.13).[47] The presence Tet1 and 5hmC at gene promoters is associated with neuronal activity and

higher transcription.[58] The co-regulation of our S40A/WT upregulated genes with Tet1 and 5hmC underscores the impact of CREB glycosylation in regulating genes important for neuronal activity and growth.

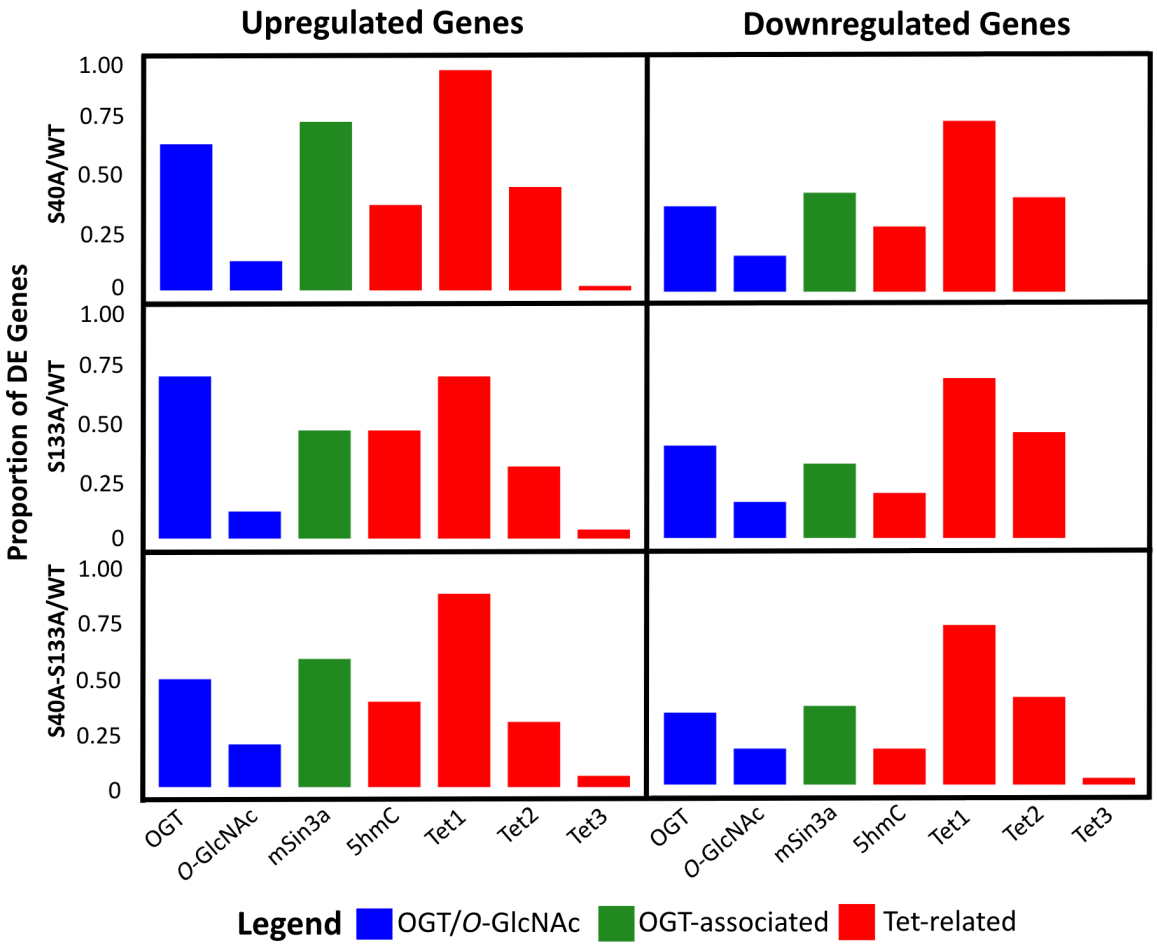


Figure 5.21 OGT-related proteins and modifications association with the S40A/WT, S133A/WT, and S40A-S133A/WT DE genes. Shown here are the proportion of genes occupied by OGT-related proteins and modifications including OGT[42], *O*-GlcNAc[46], mSin3a[42], 5hmC[47], Tet1[42], Tet2[45], and Tet3[56]. Overall, we observed that glycosylation-mediated genes are enriched for association with these different proteins and DNA modifications.

Previous studies have shown 5hmC and Tet1 association along active genes associated with neuronal differentiation.[47] Indeed, we see a particular increase in 5hmC with neuronal differentiation on the upregulated S133A/WT gene promoters: 20% of these promoters were associated with 5hmC gain with neuronal differentiation and only

4% of these promoters were associated 5hmC loss with differentiation (Table 5.12).[47] In addition, the TF Nanog was bound to the promoters of the upregulated S133A/WT genes (24%) more than the promoters of the downregulated S133A/WT genes (6%) indicating that these genes are important for pluripotency and self-renewal.[57] This supports the role of phosphorylation of CREB at serine 133 in adjudicating neuronal differentiation.

5.13 S133A and S40A-S133A is associated with neuronal differentiation and energy metabolism

To complement the pairwise comparisons of specific genes moderated by S133 phosphorylation and S40 glycosylation ablation, we next determined the gene networks through a weighted gene network analysis (WGCNA). WGCNA extracts biological meaningful gene networks from gene expression data using an unsupervised clustering approach based on gene coexpression (Figure 5.22).[59] The S133A and S40A-S133A conditions were anti-correlated ($cor_{S133A} = cor_{S40A\text{-}S133A} = -0.32$, $P = 4.1\text{x}10^{-7}$) with the cyan module enriched for genes involved in the regulation of NPC proliferation ($P = 3.6\text{x}10^{-2}$) and integration of energy metabolism ($P = 2.0\text{x}10^{-2}$) (Figures 5.23, 5.24). The differentiation and metabolism module was neither correlated with WT CREB (cor = 0.032, $P = 0.62$) nor S40A-CREB (cor = 0.024, $P = 0.71$) (Figure 5.24). The anti-correlation of the phosphorylation-deficient mutants with a module enriched for NPC proliferation genes supports the differential expression analysis revealing an important role of CREB phosphorylation at serine 133 in neuronal differentiation. In addition, the anti-correlation of both the single and double phosphorylation with the cyan module show the similarity in the genes and gene networks regulated by the S133A and S40A-S133A mutants. Finally, we expand upon our DE gene analysis by discovering a role in

CREB phosphorylation in the regulation of several metabolic pathways: fatty acid metabolism (P = 2.0x10^{-2}), amino acid metabolism (P = 2.0x10^{-2}), insulin secretion (P = 2.0x10^{-2}), and mitochondrial targeting (P = 2.1x10^{-2}). CREB governs the metabolism of glucose, mitochondria, insulin, lipids, and fatty acids.[60] Furthermore, CREB phosphorylation at serine 133 increased the transcription of gluconeogenic genes.[61] Our study demonstrates a key role of CREB serine 133 phosphorylation in controlling metabolic and differentiation gene expression.

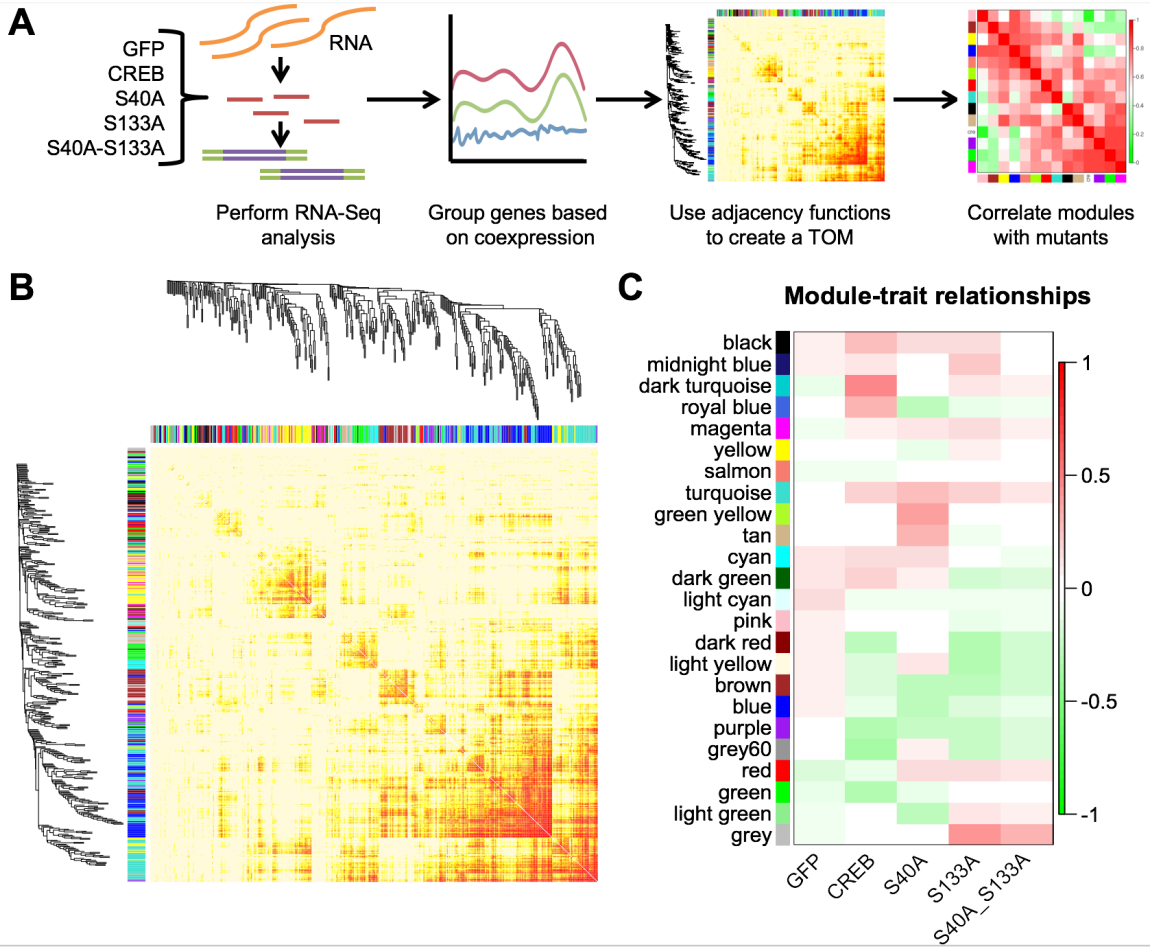


Figure 5.22 WGCNA overview. (A) Here is an overview of the procedure for WGCNA. Briefly, after RNA-Seq processing of various CREB-expressing neurons, the genes are grouped by co-expression. Then, using a weighted adjacency function, genes are clustered into biologically meaningful gene networks. Finally, networks are correlated with the different CREB mutants. (B) We generated gene dendrogram and network heatmap using our RNA-Seq data. Using a power of 10, the genes were divided into 24 different gene networks. (C) This heatmap indicates the module-trait relationships− specifically, the gene networks that are correlated with the different CREB mutants. The colors are arbitrarily assigned.

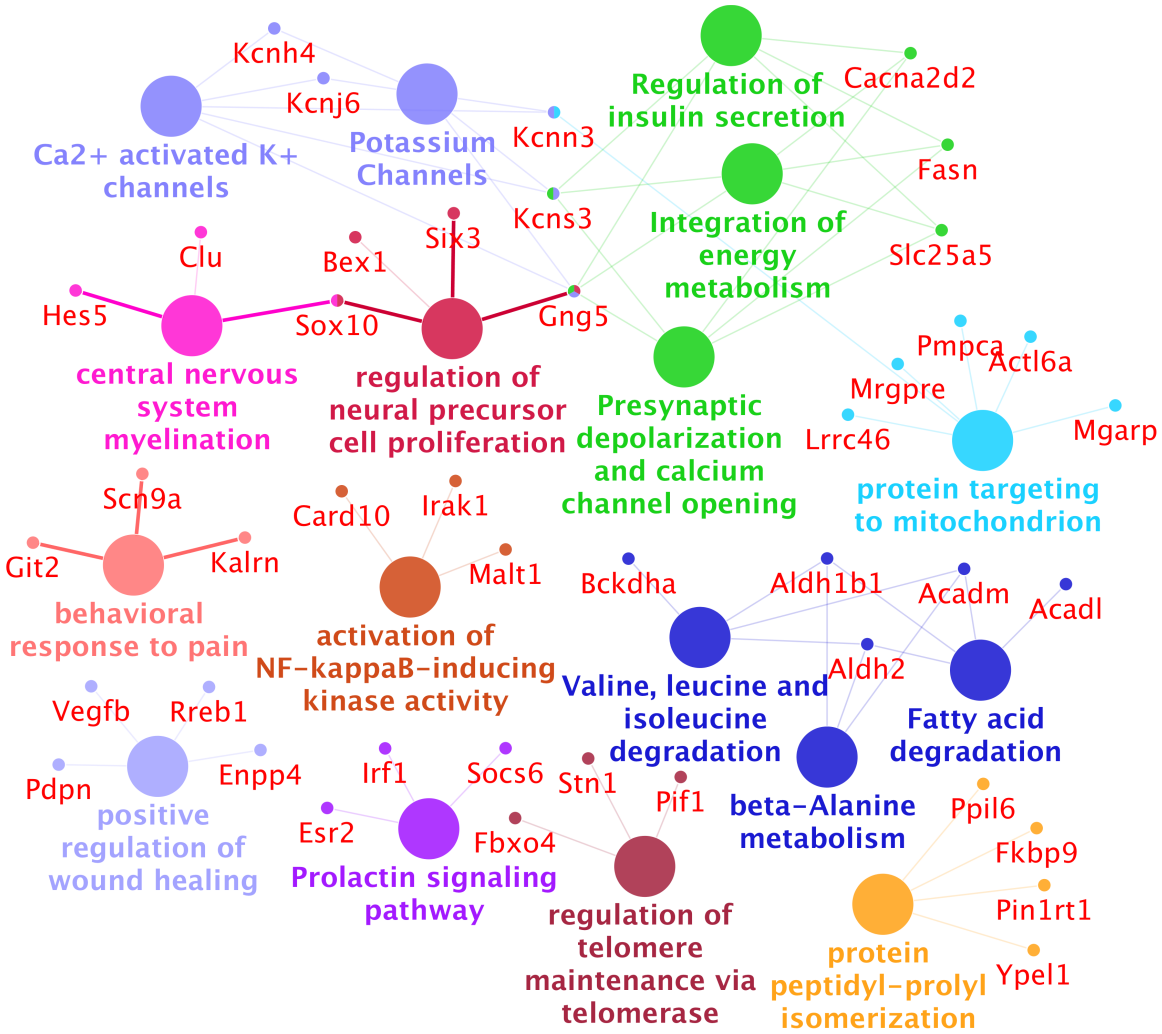
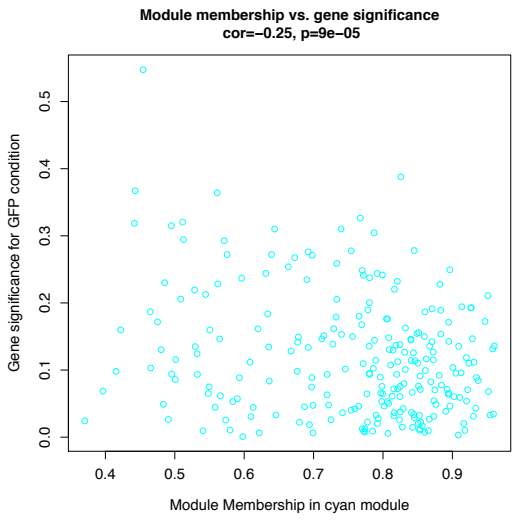
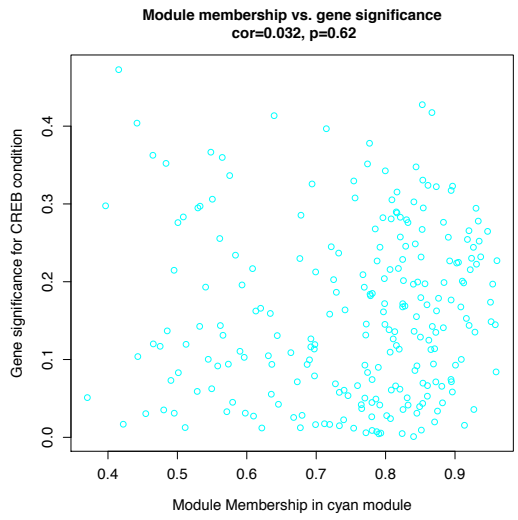
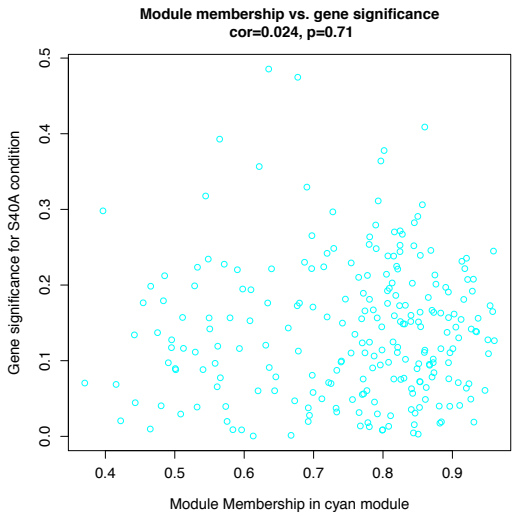
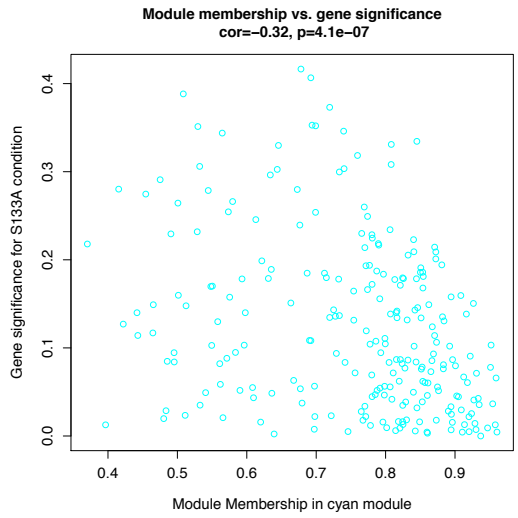


Figure 5.23 Gene ontology annotations for the S133A and S40A-S133A associated NPC proliferation- and metabolism-related module. Shown here are the top gene ontology categories, which include neural precursor cell proliferation ($P < 0.036$) and integration of energy metabolism ($P < 0.02$). The genes that belong to and are shared by these gene ontology categories are also shown. The module was annotated using the GO Biological Process, KEGG Pathways, and Reactome Pathways through the ClueGO extension in Cytoscape. The significance for all listed categories are Bonferroni step down corrected $P < 0.05$.


Module membership vs. gene significance
cor=−0.25, p=9e−05
Gene significance for GFP condition
Module Membership in cyan module


Module membership vs. gene significance
cor=0.032, p=0.62
Gene significance for CREB condition
Module Membership in cyan module


Module membership vs. gene significance
cor=0.024, p=0.71
Gene significance for S40A condition
Module Membership in cyan module


Module membership vs. gene significance
cor=−0.32, p=4.1e−07
Gene significance for S133A condition
Module Membership in cyan module

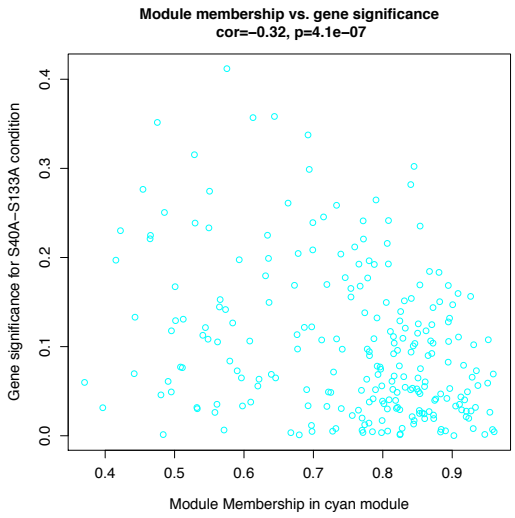

Module membership vs. gene significance
cor=−0.32, p=4.1e−07
Gene significance for S40A−S133A condition
Module Membership in cyan module

Figure 5.24 A differentiation- and metabolism-related cyan module is anti-correlated with GFP, S133A and S40A-S133A, and not correlated with WT and S40A. Module membership for the cyan module is anti-correlated with GFP (cor = -0.25, $P = 9.0 \times 10^{-5}$), S133A-CREB (cor = -0.32, $P = 4.1 \times 10^{-7}$), and S40A-S133A-CREB (cor = -0.32, $P = 4.1 \times 10^{-7}$), but not correlated with WT CREB (cor = 0.032, P = 0.62) and S40A-CREB (cor = 0.024, P = 0.71).

In addition, genes with the highest and strongest intramodular connections (called "hub genes") likely play a critical role in regulating their modules.[62] The top five hub genes for the cyan module are solute carrier family 29 member 1 and family 25 member 37 (*Slc29a1, Slc25a37*), helicase like transcription factor (*Hltf*, HLTF), caspase recruitment domain family member 10 (*Card10*, CARD10), and arylsulfatase B (*N*-acetylgalactosamine-4-sulfatase, *Arsb,* Arsb) (Figure 5.25, Table 5.14). *Slc25a37* (mitoferrin 1) and *Slc29a1* both encode transporters that localize to the mitochondrial membrane (and plasma membrane in the case of *Slc29a1*) and regulate the uptake of heme and nucleosides respectively, which in turn dictates heme and nucleotide biosynthesis and metabolism.[63,64] HLTF is a translocase that is critical for coordinating and mediating DNA replication fork progression and DNA damage response.[65-68] CARD10 is a caspase recruitment domain that participates in apoptosis and mediates NF-κB signaling.[69] Arylsulfatase B catalyzes the removal of 4-sulfate groups from chondroitin-4-sulfate and dermatan sulfate and has been shown to play a critical role in neurite outgrowth and aerobic metabolism.[70,71] Together, *Slc25a37, Slc29a1, Hltf, Card10,* and *Arsb* are the hub genes that could be critical mediators of NPC proliferation and metabolic energy processes.

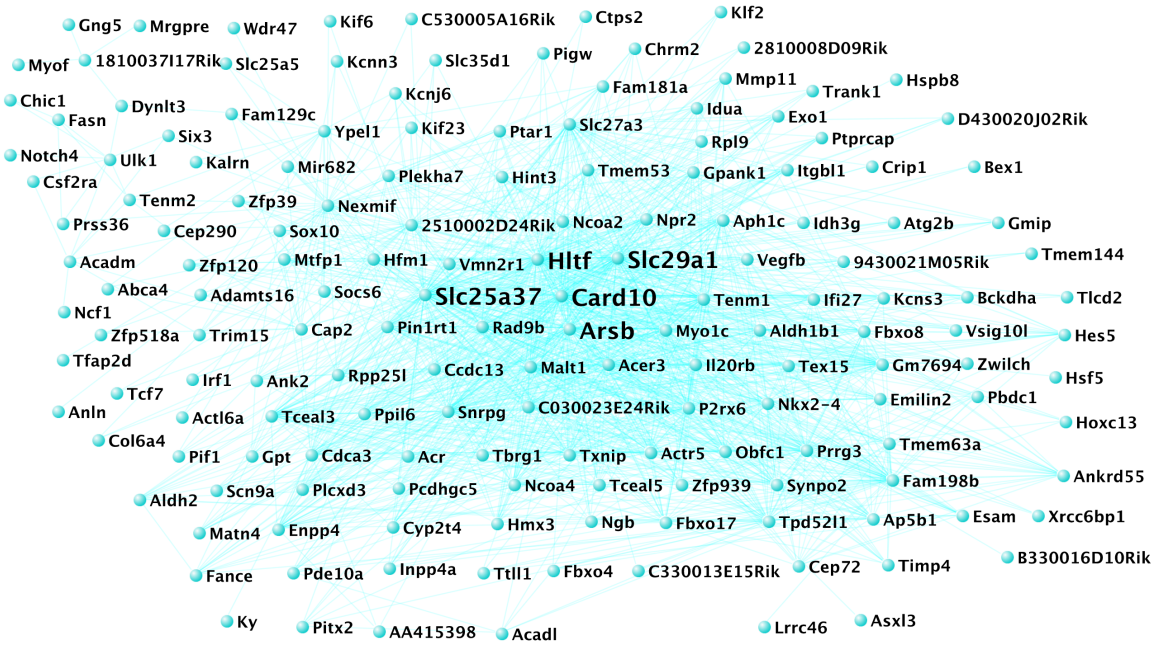


Figure 5.25 The cyan module is enriched for NPC proliferation- and metabolism-related genes. The cyan module gene network is enriched for genes involved in NPC proliferation (P = 3.6x10^{-2}) and integration of energy metabolism (P = 2.0x10^{-2}). The gene network image was generated using VisANT (edge weight cutoff > 0.1). The top five hub genes *Slc25a37, Slc29a1, Hltf, Card10*, and *Arsb* are enlarged.

Table 5.14 Cyan module hub gene connectivity

Genes	Description	Interactors	weight	k_{Within}	k_{Out}
Slc25a37	solute carrier family 25, member 37	111	0.134	55	450
Hltf	helicase-like transcription factor	109	0.141	44	661
Card10	caspase recruitment domain family, member 10	98	0.136	52	438
Arsb	arylsulfatase B	89	0.128	51	416
Slc29a1	solute carrier family 29 (nucleoside transporters), member 1	89	0.133	44	537
Slc27a3	solute carrier family 27 (fatty acid transporter), member 3	65	0.127	36	608
Fam198b	family with sequence similarity 198, member B	58	0.130	31	624
Malt1	mucosa associated lymphoid tissue lymphoma translocation gene 1	53	0.122	48	332
Tpd52l1	tumor protein D52-like 1	52	0.117	35	502
Prrg3	proline rich Gla (G-carboxyglutamic acid) 3 (transmembrane)	50	0.126	38	426
P2rx6	purinergic receptor P2X, ligand-gated ion channel, 6	49	0.123	51	259
Tceal3	transcription elongation factor A (SII)-like 3	49	0.119	46	351
2510002D24Rik	RIKEN cDNA 2510002 D24 gene	49	0.115	42	429
Rad9b	RAD9 homolog B	48	0.116	45	381
Aph1c	anterior pharynx defective 1c homolog	46	0.119	44	382

Gm7694	predicted gene 7694	44	0.119	47	353
Actr5	ARP5 actin-related protein 5 homolog	41	0.131	51	195
Ncoa4	nuclear receptor coactivator 4	40	0.121	46	286
Myo1c	myosin IC	39	0.120	50	292
C030023E24Rik	RIKEN cDNA C030023E24 gene	39	0.118	39	395
Synpo2	synaptopodin 2	35	0.129	28	87
Acer3	alkaline ceramidase 3	35	0.114	35	462
Nkx2-4	NK2 transcription factor related, locus 4	35	0.119	33	466
Txnip	thioredoxin interacting protein	32	0.129	46	195
Cdca3	cell division cycle associated 3	31	0.118	48	219

Table 5.14 shows the top 25 hub genes in the cyan module including their gene names, descriptions, the number of interactors within the module, the average connection weight, the connectivity within the module (k_{Within}), and the connectivity outside of the network (k_{Out}). The number of interactors only includes interactions with a connectivity strength > 0.1.

5.14 S40A is associated with gene networks involved in neuronal activity and excitotoxicity

Consistent with the differential expression analysis, the S40A-CREB condition was correlated with a synaptic activity module (cor = 0.47, $P = 2.3\text{x}10^{-19}$) that was enriched for genes associated with neurotransmitter release genes ($P = 2.7\text{x}10^{-3}$) and genes involved in transmission across chemical synapses ($P = 2.7\text{x}10^{-3}$) (Figures 5.26, 5.27). Interestingly, the synaptic activity module was anti-correlated with all other conditions including S133A-CREB (cor = -0.44, $P = 4.8\text{x}10^{-17}$) (Figure 5.27). This is consistent with a previous study in our lab that found that ablation of CREB glycosylation at serine 40 lead to increased transcription of neuronal activity genes.[3]

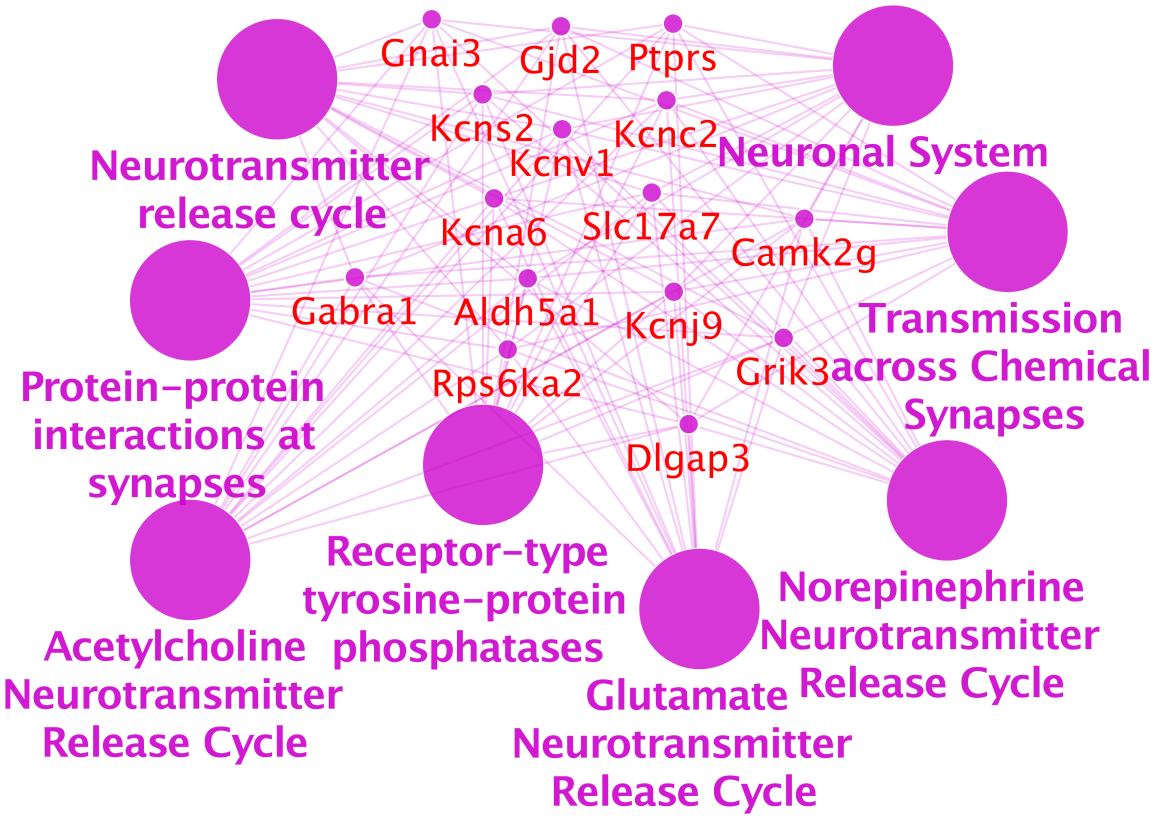
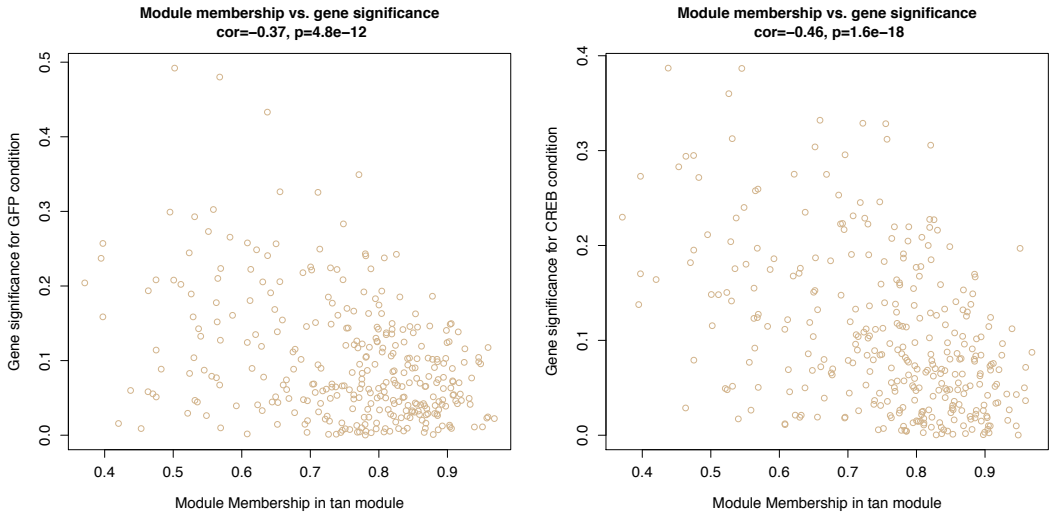


Figure 5.26 Gene ontology annotations for S40A-associated neuronal activity-related module. Shown here are the top gene ontology categories, which include neurotransmitter release cycle as well as synaptic activity. The neuronal activity module is enriched for genes involved in transcriptional activation (P = 3.0×10^{-2}) and the synapse (P = 2.4×10^{-2}). The genes that belong to and are shared by these gene ontology categories are also shown. The module was annotated using the Reactome pathways through the ClueGO extension in Cytoscape. The significance for all listed categories are Bonferroni corrected $P < 0.00027$.



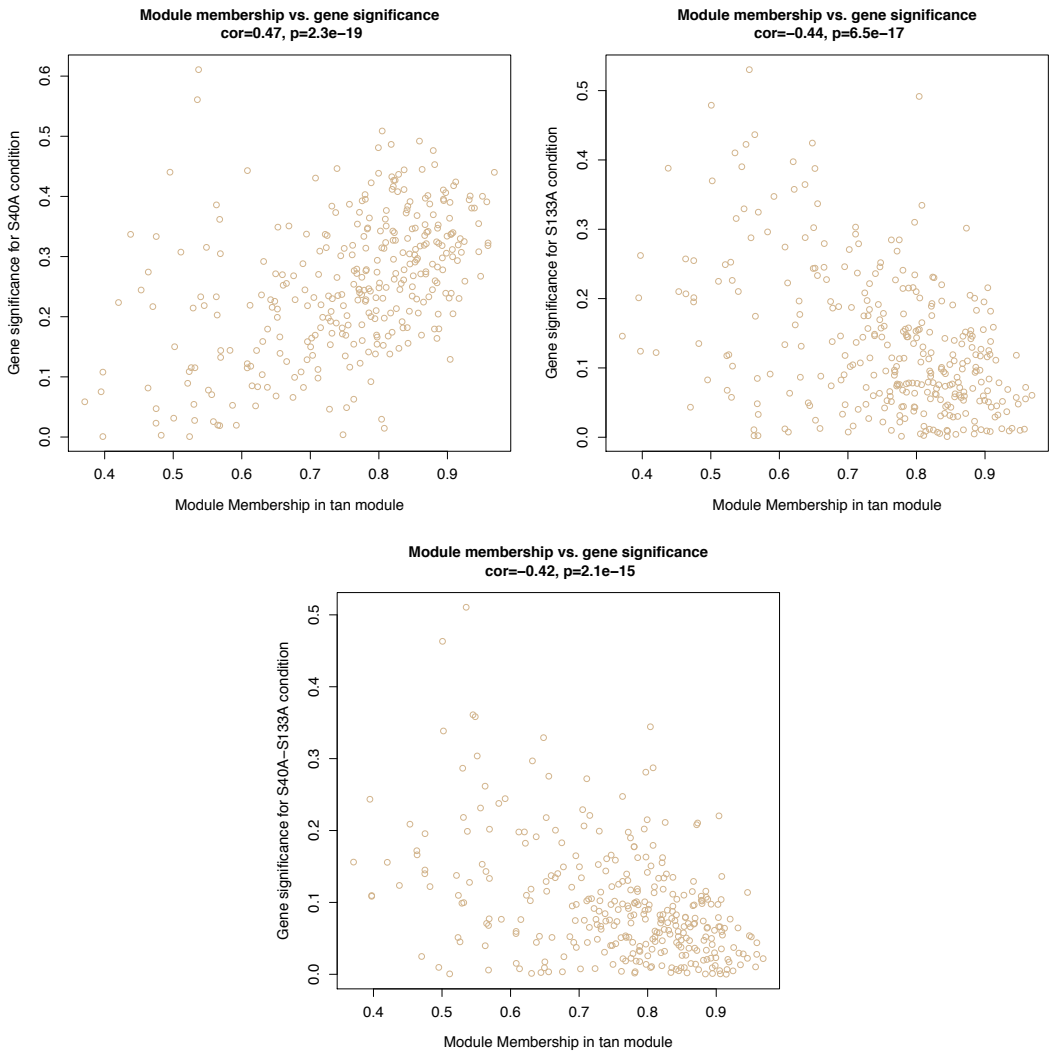


Figure 5.27 Synaptic activity-related module is correlated with S40A and anti-correlated with all other conditions. Module membership for the tan module is anti-correlated with GFP (cor = -0.37, P = 4.8x10^{-12}), WT CREB (cor = -0.46, P = 1.6x10^{-18}), S133A-CREB (cor = -0.44, P = 6.5x10^{-17}), S40A-S133A-CREB (cor = -0.42, P = 2.1x10^{-15}) and correlated with S40A-CREB (cor = 0.47, P = 2.3x10^{-19}).

The hub genes for the tan module included G protein-activated inward rectifier potassium channel 3 (Girk3, *Kcnj9*), pre-B-cell leukemia transcription factor (PBX1, *Pbx1*), SRY-box containing gene 11 (SOX11, *Sox11*), RNA binding motif protein X chromosome (RBMX, *Rbmx*), and oxysterol binding protein 2 (OSBP2, *Osbp2*) (Figure 5.28, Table 5.15). Girk3 is an inhibitory potassium channel that is important for mediating the excitability of neurons and has been shown to regulate dopaminergic

addiction pathways.[72,73] PBX1 is a transcription factor that has been shown to be important for neuronal patterning throughout the cortex, and disruption of PBX1 signaling has been implicated in Parkinson's disease.[74,75] SOX11 is another transcription factor responsible for neuronal development and differentiation.[76] RBMX is an RNA-binding protein that mediates alternative splicing and transcription and has been shown to necessary during neural development.[77-79] OSBP2 links sterol levels to sphingolipid metabolism and can regulate mitogenic signaling.[80] Together, these hub genes are poised to be critical regulators of the synaptic activity-related tan module.

Figure 5.28 The tan module is enriched for synaptic activity-related genes. The tan module gene network is enriched for neurotransmitter release genes ($P = 2.7\text{x}10^{-3}$) and genes involved in transmission across chemical synapses ($P = 2.7\text{x}10^{-3}$). The gene network image was generated using VisANT (weight cutoff > 0.1). The top five hub genes *Kcnj9, Pbx1, Sox11, Rbmx,* and *Osbp2* are enlarged.

Table 5.15 Tan module hub gene connectivity

Genes	Description	Interactors	weight	k_{Within}	k_{Out}
Kcnj9	potassium inwardly-rectifying channel, subfamily J, member 9	182	0.164	64	1135
Pbx1	preB-cell leukemia transcription factor 1	180	0.149	70	824
Sox11	SRY-box containing gene 11	174	0.166	59	1166
Rbmx	RNA binding motif protein, X chromosome	168	0.165	56	1171
Osbp2	oxysterol binding protein 2	164	0.147	53	1010

Dnajc27	DnaJ (Hsp40) homolog, subfamily C, member 27	146	0.131	67	740
Rfxank	regulatory factor X-associated ankyrin-containing protein	144	0.130	74	683
Mul1	mitochondrial ubiquitin ligase activator of NF-κB 1	129	0.146	51	982
Cds1	CDP-diacylglycerol synthase 1	126	0.129	67	688
Lsm11	U7 snRNP-specific Sm-like protein LSM11	124	0.137	49	967
Lphn2	latrophilin 2	119	0.135	57	711
Snph	syntaphilin	118	0.160	43	1237
Gpbp1	GC-rich promoter binding protein 1	116	0.153	41	1094
2900026A02Rik	RIKEN cDNA 2900026A02 gene	114	0.129	59	775
Smarcc1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	112	0.128	66	574
Ift27	intraflagellar transport 27 homolog	108	0.136	43	1007
Ssx2ip	synovial sarcoma, X break point 2 interacting protein	108	0.132	44	966
Tes	testis derived transcript	108	0.128	56	652
Lamb2	laminin, beta 2	107	0.132	55	736
Dgka	diacylglycerol kinase, alpha	105	0.124	68	558
Recql5	RecQ protein-like 5	98	0.135	48	890
Ly6g5b	lymphocyte antigen 6 complex, locus G5B	92	0.119	56	599
Carnmt1	carnosine *N*-methyltransferase 1	89	0.125	61	648
Slc17a7	solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member7	89	0.128	47	818
Gm996	predicted gene 996	83	0.127	53	768

Table 5.15 shows the top 25 hub genes in the tan module including their gene names, descriptions, the number of interactors within the module, the average connection weight, the connectivity within the module (k_{Within}), and the connectivity outside of the network (k_{Out}). The number of interactors only includes interactions with a connectivity strength > 0.1.

Intriguingly, the S40A-CREB condition was correlated with a neuronal excitotoxicity module (cor = 0.46, $P = 1.8\text{x}10^{-19}$) that was enriched for genes involved in neurodegeneration (Alzheimer's ($P = 6.5\text{x}10^{-8}$), Parkinson's ($P = 1.2\text{x}10^{-7}$) and Huntington's diseases ($P = 1.4\text{x}10^{-7}$)), mitochondrial activity ($P = 1.2\text{x}10^{-7}$), and in ion channels ($P = 1.0\text{x}10^{-4}$) (Figures 5.29 and 5.30).[81] This green yellow module was anti-correlated with all other conditions (cor = -0.27 to -0.32, $P < 9.8\text{x}10^{-7}$) (Figure 5.30). Our results suggest that glycosylation at serine 40 acts as a deterrent against aberrant neuronal

activity and excitotoxicity. The major hub genes for this module regulate synaptic activity [cholinergic receptor muscarinic 3 (*Chrm3*), calcium voltage-gated channel auxiliary subunit γ 2 (*Cacng2*), and γ-aminobutyric acid type A receptor δ subunit (*Gabrd*)] and cell growth and survival [nephronectin (*Npnt*) and inhibitor of DNA binding, HLH protein (*Id1*)] (Figure 5.31, Table 5.16). The WGCNA results support and extend the DE gene data by uncovering the association of S40A with synaptic activity and excitotoxicity modules.

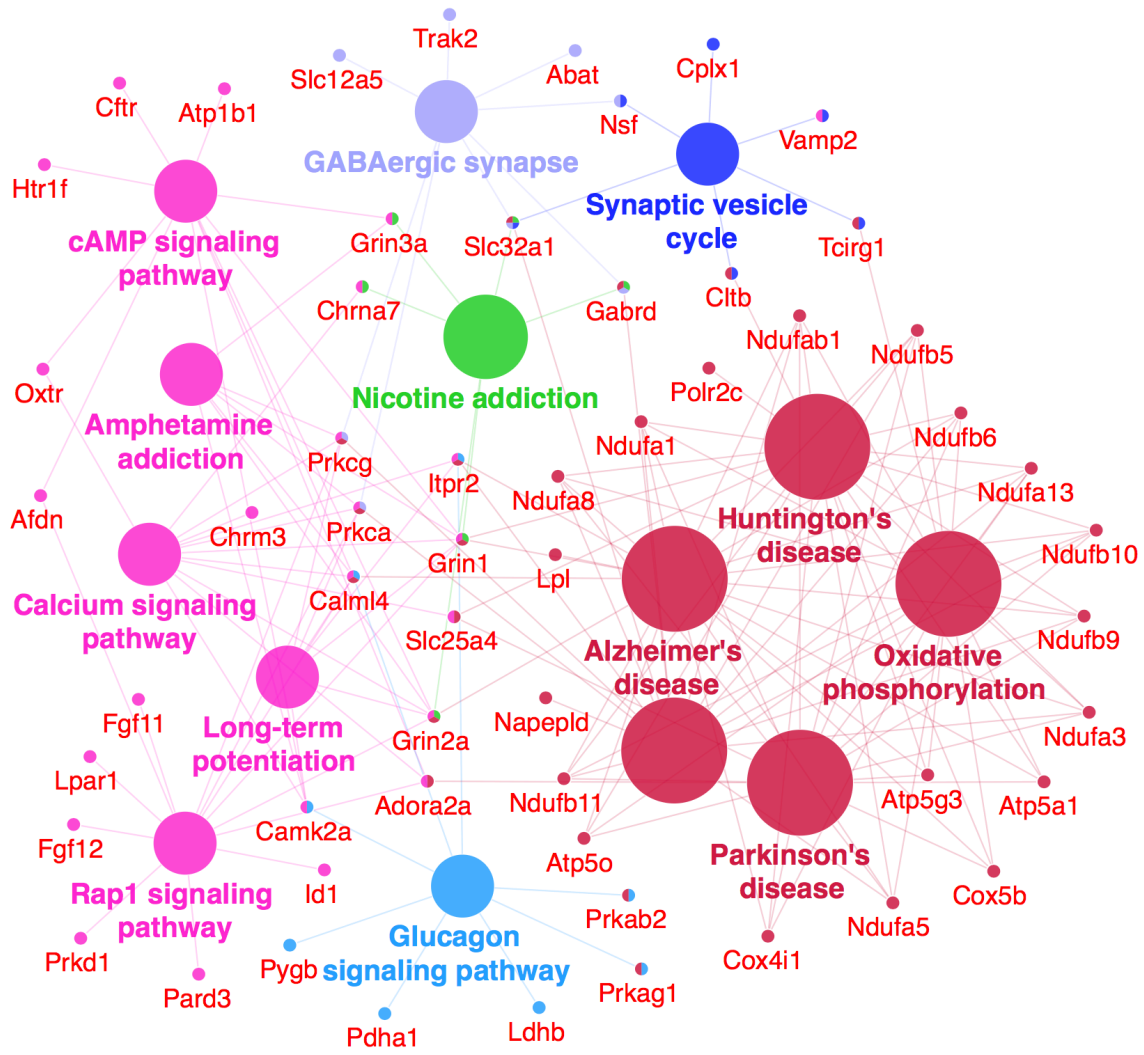
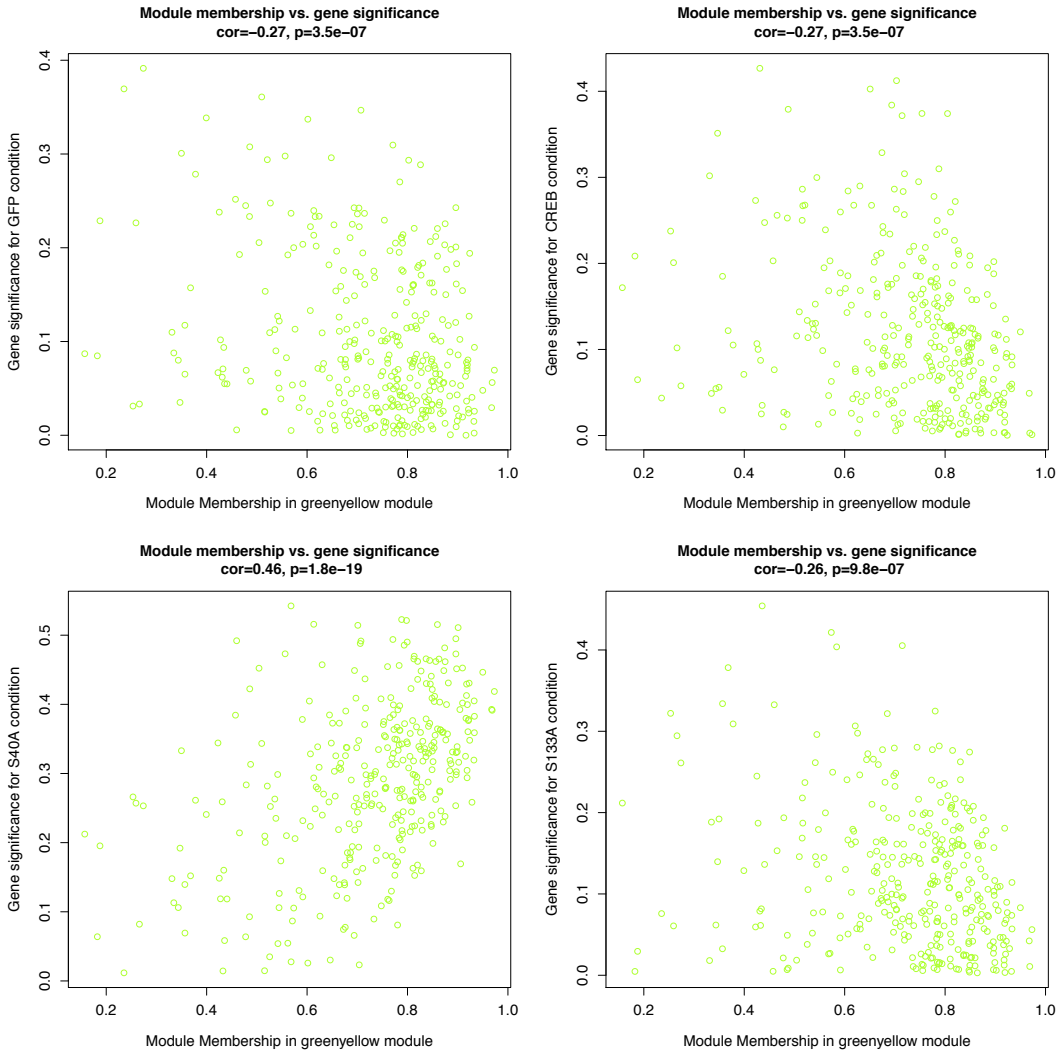


Figure 5.29 Gene ontology annotations for the neuronal activity- and excitotoxicity-related module correlated with S40A. Shown here are the top gene ontology categories, which include oxidative phosphorylation and neurodegenerative diseases in addition to long-term potentiation and calcium

signaling. The green yellow module is enriched for genes involved in neurodegeneration (Alzheimer's (P = 6.5×10^{-8}), Parkinson's (P = 1.2×10^{-7}) and Huntington's diseases (P = 1.4×10^{-7})), mitochondrial activity (P = 1.2×10^{-7}), and ion channels (P = 1.0×10^{-4}). The genes that belong to and are shared by these gene ontology categories are listed. This excitotoxicity module was annotated using the KEGG pathways through the ClueGO extension in Cytoscape. The significance for all listed categories are Bonferroni step down corrected $P < 0.002$.



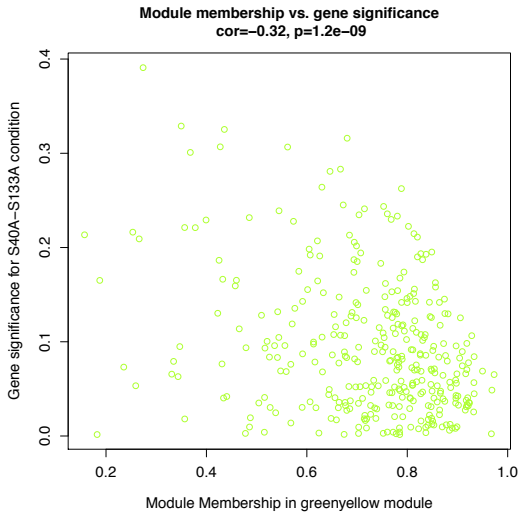


Figure 5.30 Green yellow module is positively correlated with S40A and anti-correlated with all other conditions. Module membership for the green yellow module is anti-correlated with GFP (cor = -0.27, $P = 3.5\text{x}10^{-7}$), WT CREB (cor = -0.27, $P = 3.5\text{x}10^{-7}$), S133A-CREB (cor = -0.26, $P = 9.8\text{x}10^{-7}$), S40A-S133A-CREB (cor = -0.32, $P = 1.2\text{x}10^{-9}$) and correlated with S40A-CREB (cor = 0.46, $P = 1.8\text{x}10^{-19}$).

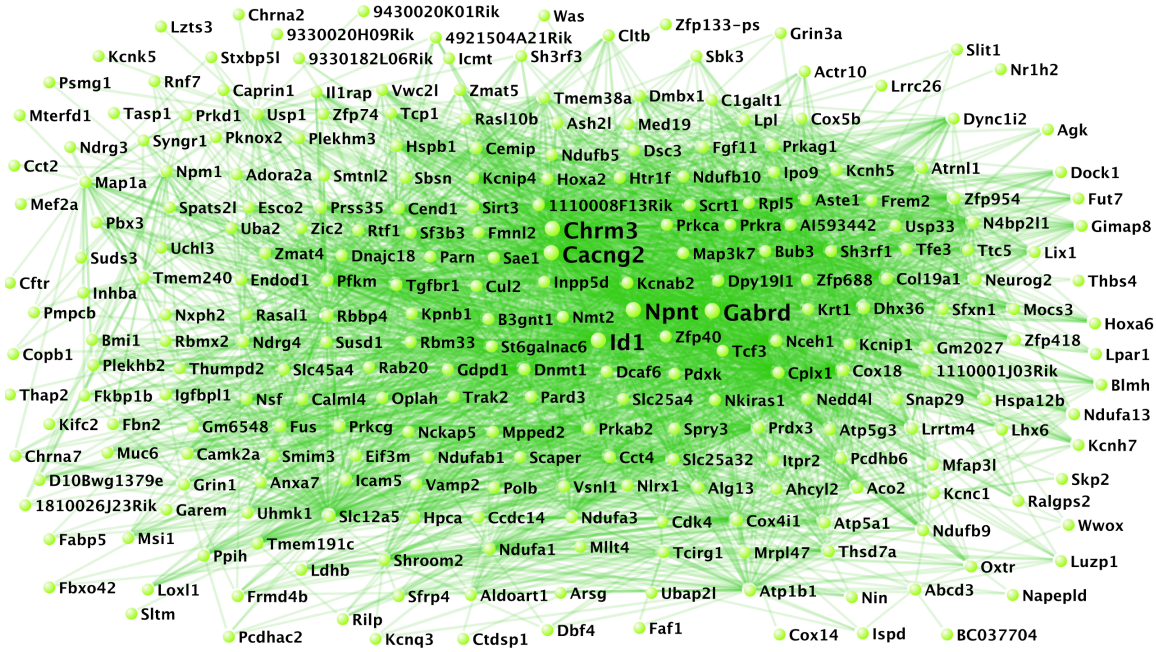


Figure 5.31 The green yellow module is enriched for neuronal activity and excitotoxicity genes. The tan module gene network is enriched for genes involved in neurodegeneration (Alzheimer's ($P = 6.5\text{x}10^{-8}$), Parkinson's ($P = 1.2\text{x}10^{-7}$) and Huntington's diseases ($P = 1.4\text{x}10^{-7}$)), mitochondrial activity ($P = 1.2\text{x}10^{-7}$), and in ion channels ($P = 1.0\text{x}10^{-4}$). The gene network image was generated using VisANT (weight cutoff > 0.1). The top five hub genes *Chrm3*, *Cacng2, Gabrd, Npnt,* and *Id1* are enlarged.

Table 5.16 Green yellow module hub gene connectivity

Genes	Description	Interactors	weight	k_{Within}	k_{Out}
Cacng2	calcium channel, voltage-dependent, γ subunit 2	150	0.143	51	1050
Npnt	nephronectin	147	0.138	65	834

Gabrd	γ-aminobutyric acid (GABA) A receptor, subunit δ	146	0.149	51	1069
Id1	inhibitor of DNA binding 1	146	0.144	43	1104
Chrm3	cholinergic receptor, muscarinic 3, cardiac	143	0.136	46	1063
Kcnab2	potassium voltage-gated channel, shaker-related subfamily, β member 2	138	0.127	64	747
Cplx1	complexin 1	130	0.127	73	660
Cox18	COX18 cytochrome c oxidase assembly homolog	128	0.135	35	1089
Dnmt1	DNA methyltransferase (cytosine-5) 1	124	0.124	59	703
Dpy19l1	dpy-19-like 1	124	0.132	56	852
Pfkm	phosphofructokinase, muscle	121	0.126	69	707
Tgfbr1	transforming growth factor, β receptor I	121	0.126	60	762
Aste1	asteroid homolog 1	103	0.140	49	858
Slc12a5	solute carrier family 12, member 5	102	0.130	47	784
Dhx36	DEAH(Asp-Glu-Ala-His) box polypeptide 36	100	0.127	42	916
Col19a1	collagen, type XIX, α1	98	0.128	59	720
Nceh1	aryl acetamide deacetylase-like 1	98	0.124	56	698
Sh3rf1	SH3 domain containing ring finger 1	98	0.140	40	976
Rbbp4	retinoblastoma binding protein 4	96	0.129	43	877
1110008F13Rik	RIKEN cDNA 1110008F13 gene	94	0.134	33	1081
AI593442	expressed sequence AI593442	89	0.129	62	642
Bub3	budding uninhibited by benzimidazoles 3 homolog	82	0.124	45	769
St6galnac6	ST6 (α-*N*-acetyl-neuraminyl-2,3-β-galactosyl-1,3)-*N*-acetylgalactosaminide α -2,6-sialyltransferase 6	80	0.117	43	807
Cox4i1	cytochrome c oxidase subunit IV isoform 1	74	0.126	46	759
Cend1	cell cycle exit and neuronal differentiation 1	72	0.119	55	677

Table 5.16 shows the top 25 hub genes in the green yellow module including their gene names, descriptions, the number of interactors within the module, the average connection weight, the connectivity within the module (k_{Within}), and the connectivity outside of the network (k_{Out}). The number of interactors only includes interactions with a connectivity strength > 0.1.

5.15 Discussion

Overall, our results show that regulation of CREB through glycosylation and phosphorylation is critical for negotiating distinct transcriptional repertoires. In particular, we show that removal of glycosylation at S40 of CREB leads to the expression of neuronal excitability genes while abrogation of S133 phosphorylation affects the transcription of genes involved in neuronal development and differentiation. Minimal

overlap between the S40A and S133A conditions especially in the upregulated genes reveals that CREB glycosylation and phosphorylation mediate different sets of genes. Furthermore, we found that some of the genes differentially regulated by CREB phosphorylation and glycosylation were (1) known targets of CREB and its coactivators, CBP and p300 and (2) associated with euchromatic regions of the genome. Furthermore, OGT, *O*-GlcNAc, mSin3a, and Tet1 bind directly to the genes upregulated in the glycosylation-deficient CREB. Finally, we found that phosphorylation of CREB is important for mediating neuronal differentiation and several metabolic pathways while glycosylation of CREB is a central regulator of synaptic activity- and excitotoxicity-related gene networks. Together, CREB *O*-GlcNAcylation at serine 40 and phosphorylation at serine 133 are critical regulators of CREB transcription and mediators of neuronal activity and homeostasis.

Previous studies had implicated CREB phosphorylation at S133 as enhancing CREB transcription of specific genes through the enhancement of CREB's interaction with CBP and p300.[2,82] More recent studies have begun to explore the role of S133 phosphorylation on CREB-mediated transcription globally. In a microarray study, McClung and colleagues showed that S133A-CREB overexpression results in opposing effects on transcription when compared to WT CREB in the nucleus accumbens of mice with a total of 24 differentially-expressed genes between the S133A-CREB and WT CREB conditions.[28] We did not observe overlap between the microarray study; potential reasons for the differences could be (1) that the study involved overexpression of S133A-CREB and WT CREB in the nucleus accumbens over course of 8 weeks while our study used a Creb1$^{\alpha\delta}$ background and explored considerably shorter time periods (4 and 8

hours) in a E16.5 cortical neuronal population and (2) the differing sensitivity in detection for microarray studies when compared to RNA-Seq experiments, especially with low abundance transcripts.[32]

Another RNA-Seq study by Briand and coworkers showed that a phosphorylation-deficient S133A mutant mouse had no DE genes in the hippocampus when compared with WT mice.[23] In this study, breeding of the heterozygous S133A-CREB mice resulted in lower numbers than the expected Mendelian frequencies (only 11% homozygotes), leading the authors to suggest that while S133 phosphorylation may not affect transcription and memory-related behavior, it may affect development.[23] Prior studies have shown that CREB plays a vital role in neuronal development and differentiation.[2,33,34] Moreover, we found that phosphorylation of CREB modulates key metabolic pathways including fatty acid, mitochondrial, and amino acid metabolism as well as insulin secretion. CREB governs the metabolism of glucose, mitochondria, insulin, lipids, and fatty acids.[60] Furthermore, CREB phosphorylation at serine 133 increased the transcription of gluconeogenic genes.[61] Our study demonstrates a key role of CREB serine 133 phosphorylation in controlling metabolic gene expression and arbitrating neuronal development and differentiation.

Our work demonstrates that glycosylation of CREB, rather than just acting as a brake to overall CREB-mediated transcription, specifically mitigates the expression of neuronal activity and excitability genes. Removal of the glycosylation at serine 40 upregulated the expression of neuronal activity genes important for every part of the neuron from the dendrite to the axon. This is consistent with a previous study in our lab that found that ablation of CREB glycosylation at serine 40 led to increased transcription

of neuronal activity genes, enhanced neurite outgrowth, and accelerated memory formation when compared to WT CREB.[3] Elevated CREB activity through the expression of more active VP16, Y134F, and DIEDML CREB mutants led to heightened transcription of neuronal excitability genes.[16,17] We show that glycosylation of CREB at serine 40 is critical for fine-tuning the transcription of neuronal activity and excitability genes.

In addition, our S40A-CREB displayed key transcriptional similarities and differences with the constitutively active CREB mutant, VP16-CREB, which enhances the association of CREB with its coactivators, CBP and p300; similar to VP16-CREB, S40A-CREB upregulated genes associated with synaptic activity, but unlike VP16-CREB, S40A-CREB also increased neuronal excitation gene expression (Table 5.11). Interestingly, these neuronal excitability genes were known targets of CRTC1 in drosophila in response to learning (Table 5.11).[37] In contrast, the similarly regulated genes in S40A-CREB and VP16-CREB were not known targets of CRTC1, but instead were targets of CBP. This suggests that CREB glycosylation at serine 40 specifically affects genes directly regulated by CRTC1, consistent with our previous study showing that S40A-CREB enhanced binding to CRTC.[3] Indeed, recent studies have shown that CRTC is more critical for long-term memory formation and neuronal activity, whereas CBP may be important for the expression of genes early on in memory formation.[37,83-85]

Finally, our WGCNA showed that ablation of CREB serine 40 glycosylation correlated with excitotoxicity gene expression. Our results suggest that glycosylation at serine 40 acts as a significant deterrent against aberrant neuronal activity and excitotoxicity. This excitotoxicity association was also seen in a constitutively active

CREB mutant, VP16-CREB; mice expressing VP16-CREB across the brain displayed increased expression of immune response genes and impaired spatial memory retrieval despite exhibiting enhanced LTP.[2,16,35] Indeed, neurons with higher CREB activation appear to be more prone to excitotoxic effects.[86] Consistent with this finding, researchers discovered elevated CRE-mediated transcription in Huntington's disease model mice suggesting that elevated CREB transcription may lead to an excitotoxicity-induced neurodegenerative phenotype.[87] In contrast, Alzheimer's disease, Rubenstein-Taybi syndrome, and Coffin-Lowry syndrome are associated with impaired CREB-mediated transcription.[88]

These studies highlight the necessity of a judicious PTM rheostat to avoid neuronal apoptosis through CREB under-activation or excitotoxicity-induced neurodegeneration through CREB over-activation. Through a global analysis of the gene networks regulated by phosphorylation, glycosylation, and both, we have begun to decipher the PTM transcription code of CREB that ordains neuronal fate. Control of CREB activity through *O*-GlcNAcylation at serine 40 attenuates neuronal activity and excitotoxicity while phosphorylation at serine 133 moderates neuronal development. Together, CREB *O*-GlcNAcylation at serine 40 and phosphorylation at serine 133 are critical for the maintenance of neuronal homeostasis and could therefore be therapeutic targets for neurodegenerative diseases.

5.16 Methods

5.16.1 Breeding and genotyping Creb1$^{\alpha_\delta}$ mice

All animal procedures were performed in accordance with Institutional Animal Care and Use Committee (IACUC) guidelines of the California Institute of Technology.

CREB (α and δ) knock-out mice heterozygous mice (129S2/SvPasCrl background and C57BL/6 background) were group-housed with ad libitum access to food and water. Mice were bred and genotyped as previously described.[89,90] The mice were maintained with outbreeding with their respective genetic backgrounds- either 129/SVJ mice or C57BL/6 mice.[91,92] The heterozygotic 129S2/SvPasCrl CREB+/- (129.Creb1$^{\alpha_{\delta}+/-}$) mice were mated with the heterozygotic C57BL/6 CREB+/- (B6.Creb1$^{\alpha_{\delta}+/-}$) mice in order to yield homozygous, heterozygous, and wild-type embryos (Figure 5.32). We encountered some pregnancy detection and breeding issues that were rectified by breeding B6.Creb1$^{\alpha_{\delta}+/-}$ females with 129.Creb1$^{\alpha_{\delta}+/-}$ males and extra maternal support (lower rack placement and cage rearrangement to allow mothers to descend from upper play area easier) respectively.

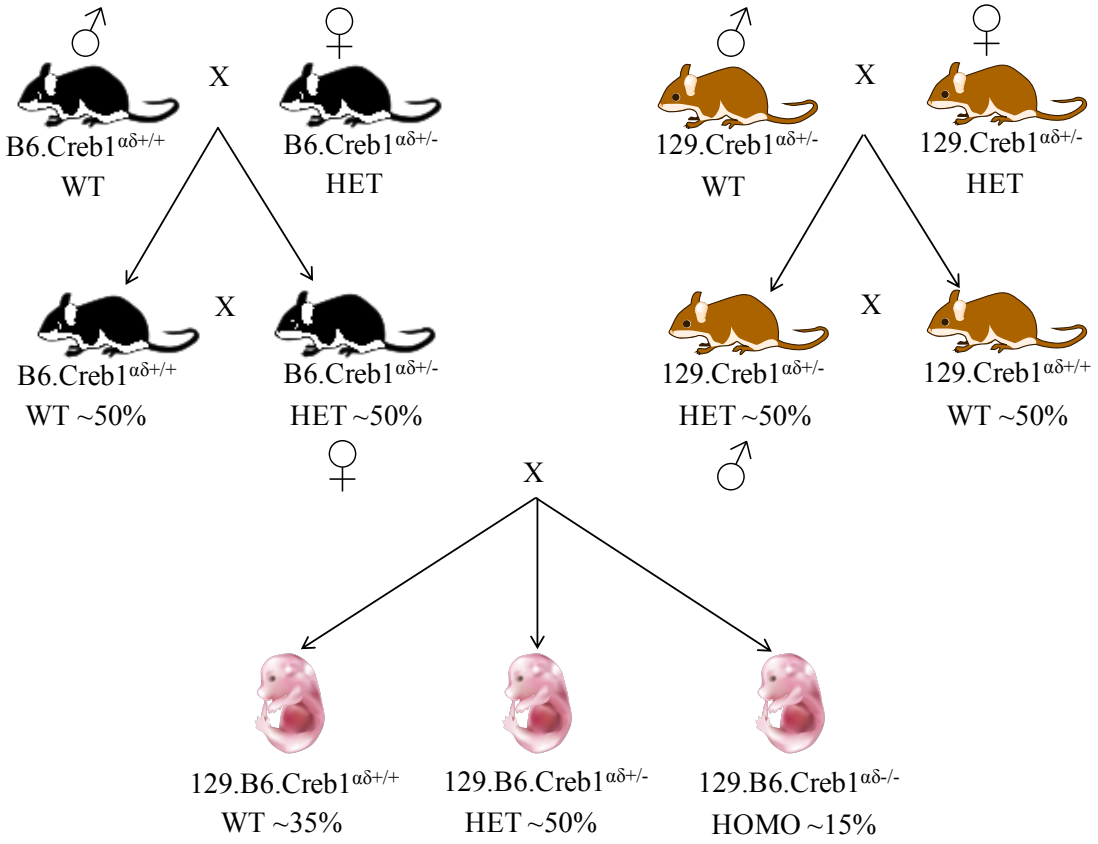


Figure 5.32 Breeding scheme for Creb1$^{\alpha_\delta}$ homozygous knockout mice. Shown here is the overall schematic for breeding the C57BL/6 Creb1$^{\alpha_\delta}$ heterozygous mice (denoted B6.Creb1$^{\alpha_\delta+/-}$) and 129S2/SvPasCrl Creb1$^{\alpha_\delta}$ (denoted 129.Creb1$^{\alpha_\delta+/-}$) in order to obtain 129.B6.Creb1$^{\alpha_\delta-/-}$ embryos. Mouse image from ChemDraw Professional 16.0 and mouse embryo image from the DataBase Center for Life Sciences (DBCLS).

Only homozygous neuronal cultures were used for subsequent experiments. Genotyping was performed by PCR as described by previous studies.[91] The DNA samples were isolated from tail tips using the standard procedure from the DNeasy Blood & Tissue Kits (Qiagen, 69504). Then, PCR amplification was performed on 50 ng of template genomic DNA using the Q5 Hot Start High-Fidelity DNA polymerase (New England BioLabs, Inc.) [98°C for 30 sec; [98°C for 10 sec; 51.8°C for 20 sec; 72°C for 30 sec (34 cycles)], and then 72°C for 2 min] to identify the wild-type (150 bp band), heterozygous (150 + 350 bp band), and homozygous (350 bp) using DNA agarose electrophoresis (1.5% agarose gel in TAE buffer). The primers used were as follows: oIMR3081 (Mutant): 5'-TGATGGATACTTTCTCGGCA-3', oIMR3082 (Common): 5'-ATGTATTTTTATACCTGGGC-3'; and oIMR3083 (Wild type): 5'-TATTGTAGGTAACTAAATGA-3'. We used Laragen's genotyping services to corroborate our genotyping results.

5.16.2 Creb1$^{\alpha_\delta}$ E16.5 cortical dissections

Mouse cortical neurons were prepared as described previously [5]. Briefly, we dissected E16.5 neurons by separating embryos and then using 1X TrypLE (ThermoFisher, 12605028) to trypsinize individual cortices. Following 20 minutes of trypsinization, 1X Defined Trypsin Inhibitor (ThermoFisher, R007100) was added to the neurons to quench trypsinization. Next, DMEM, high glucose, GlutaMAX Supplement (ThermoFisher, 10566024) with 10% fetal bovine serum, heat-inactivated

(ThermoFisher, 10082147) was added to the mixture followed by trituration using a flame-tipped glass pipette with washing in clean 1X HBSS, no calcium, no magnesium, phenol red (ThermoFisher, 14170161). Then, neurons were plated at approximately $2x10^6$ cells in 6 cm dishes. The dishes were coated for at least 2 hours in poly-D-lysine hydrobromide MW 70,000-150,000 (ThermoFisher, P6407-5MG) in 0.2 μm filter-sterilized PBS (10 mM PO_4, 137 mM NaCl, 2.7 mM KCl), pH 7.4 at 37°C and 5% CO_2 and then washed twice with sterile-filtered water and allowed to dry. The neurons were cultured in 5 ml of neuronal medium (Neurobasal Medium (ThermoFisher, 21103049) supplemented with GS21 Neural Supplement (Amsbio, GSM-3100), Penicillin-Streptomycin (ThermoFisher, 15070063), and GlutaMAX Supplement (ThermoFisher, 35050061)). Two hours after plating all the media was removed and replaced with warmed media. After 4 days *in vitro* (DIV), half of the media was replaced with fresh media. Following 6DIV, half of the media was replaced with fresh media supplemented with a final concentration of 1 μM tetrodotoxin (TTX) (Tocris Biosciences, 1078) and 100 μM D-(-)-2-amino-5-phosphonopentanoic acid (D-AP5) (Tocris Biosciences, 0106-1mg). At 12-16 hours following silencing, we added HSV for 4 hours or 8 hours and then depolarized neurons for 2 hours using 55 mM KCl. For the axonal and dendritic growth assays, neurons were plated and at 1DIV, HSV was added to the neurons. At 3DIV, the neurons were fixed and then prepared for ICC as described below.

5.16.3 Herpes simplex virus (HSV) transduction and immunocytochemistry (ICC)

We cloned FLAG-tagged rat Creb1 gene into the Gateway pENTR 1A entry dual selection vector (ThermoFisher, A10462). Then, we used the Q5 Site-Directed Mutagenesis Kit (New England BioLabs, E0554S) in order to mutate serine 40 and/or

serine 133 to alanines. The MIT Viral Gene Transfer Core (Dr. R. Neve) then recombined the various CREB mutants into the Gateway HSV-Syn -IRES-GFP construct (Syn refers to the synapsin promoter) and packaged this in high titer herpes simplex virus (HSV). Viruses were diluted in sterile 25 mM HEPES, pH 7.3 (ThermoFisher, 1530080) and stored at -80°C until addition to cells. The viral titer was tested initially in primary cortical neuronal cultures using immunocytochemistry (ICC) until ~100% transduction was achieved as measured by GFP readout. For the RNA-Seq experiments, the viral levels were titered to ensure that the amount of Creb1 expression was controlled across samples through qPCR (see qPCR section in Chapter 2.14.8).

For ICC experiments, 15mm coverslips were coated as described above with poly-D-lysine hydrobromide in 12-well plates. Then, E16.5 cortical neurons were seeded at 100,000 neurons/well in 1 ml of neuronal medium. On 5DIV (or 3DIV for the neuronal growth assays), the neurons were treated with HSV-Syn-IRES-GFP (GFP), HSV-Syn-CREB-IRES-GFP (WT), HSV-Syn-S40A-CREB-IRES-GFP (S40A), HSV-Syn-S133A-CREB-IRES-GFP (S133A), or HSV-Syn-S40A-S133A-CREB-IRES-GFP (S40A-S133A) for the indicated amount of time (2, 4, 6, 8, 10, 12, or 48 hours). After HSV treatment, the media was removed and the cells were rinsed once with ice cold PBS and then fixed with 4% paraformaldehyde for 20 min. Following fixation, the coverslips were rinsed 3X with PBS then permeabilized and blocked with 0.3% Triton X-100 in 10% normal goat serum (NGS) (ThermoFisher, 16310064) in PBS for 30 min. Next, the cells were incubated for 2 hours at room temperature in the following primary antibodies: 1:200 anti-M2 FLAG (Sigma, F1804-1MG), 1:200 anti-Tau-1 (Millipore, clone PC1C6, MAB3420), and 1:1,000 anti-microtubule-associated protein 2 (MAP2) (Millipore,

AB5622) in dilution buffer (2% NGS in PBS). The coverslips were rinsed 3X with PBS and incubated in the following secondary antibodies: 1:1,000 Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 546, ThermoFisher, A-11030 or Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 405, ThermoFisher, A-31556 in PBS for 2 hours at room temperature and rinsed 3X in PBS. Finally, the coverslips were mounted on slides in Vectashield mounting medium with DAPI (Vector Labs, H-1200) or without DAPI (Vector Labs, H-1000). The slides were visualized using a Zeiss LSM700 confocal microscope with 20X magnification with and without 5x5 tiling.

5.16.4 RNA extraction, qPCR, and RNA-Seq

The neuronal cultures were lysed and total RNA was extracted using an RNeasy Plus Mini Kit per the manufacturer's instructions (Qiagen, 74134). Following RNA extraction, qPCR analysis was performed as described in Chapter 2 with primers listed in Appendix III. For RNA-Seq, the total RNA samples were processed by Igor Antoshechkin and Vijaya Kumar at the Millard and Muriel Jacobs Genomics Laboratory (Caltech) using the mRNA-Seq Sample Preparation Kit following the manufacturer's instructions (Illumina). Briefly, the samples were enriched for poly-A mRNA using Sera-Mag magnetic Oligo(dT) beads and then subjected to divalent cation-catalyzed fragmentation to an average of ~300-350 bp. The enriched mRNA was converted to double-stranded cDNA through reverse transcription using random primers and SuperScript™ II Reverse Transcriptase (ThermoFisher) followed by degradation of mRNA and synthesis of the second strand of cDNA using DNA polymerase I. Then, cDNA purification, end repair, adaptor ligation, DNA purification, and several rounds of

PCR amplification were performed sequentially. Finally, the amplified cDNA was run on a short read the Illumina HiSeq2500. Sample quality and library validation were assessed using the Agilent 2100 Bioanalyzer and Nanodrop ND-1000 spectrophotometer (ThermoFisher).

Data quality processing and analysis were performed on the usegalaxy.org interface.[7] First, we took the raw fastq files and preprocessed by FastQC and low quality data was trimmed using FASTQ Trimmer. Following quality control preprocessing, the data was aligned to the mouse genome (*mm10*) using the TopHat and Cufflinks with bias correction.[93] The TopHat mapping results are displayed in Table 5.17.

Table 5.17 Summary of TopHat alignments to mm10 genome using Galaxy

Condition	Rep.	Input	Mapped	% input	MA	% input	>20 MA
GFP_4h	1	34195901	26456732	77.4%	3255290	12.3%	38
GFP_4h	2	27460980	24016164	87.5%	2959035	12.3%	2106
WT_4h	1	33535091	27341653	81.5%	3417555	12.5%	22
WT_4h	2	26656082	23080495	86.6%	3032739	13.1%	50
S40A_4h	1	36484023	29348405	80.4%	3582952	12.2%	36
S40A_4h	2	26409455	22548830	85.4%	2876216	12.8%	38
GFP_8h	1	42732370	34177292	80.0%	4494319	13.2%	69
GFP_8h	2	47061738	35357771	75.1%	4301622	12.2%	87
WT_8h	1	36321945	29029453	79.9%	3648703	12.6%	42
WT_8h	2	29035772	20981169	72.3%	2629444	12.5%	42
S40A_8h	1	44277390	34987651	79.0%	4275386	12.2%	77
S40A_8h	2	37784130	27882309	73.8%	3478354	12.5%	54
GFP_8h	1	33562997	23993641	71.5%	2949809	12.3%	40
GFP_8h	2	31777703	21733281	68.4%	2933432	13.5%	38
WT_8h	1	34856413	27226805	78.1%	3217959	11.8%	19
WT_8h	2	36621109	21798551	59.5%	2941624	13.5%	35
S40A_8h	1	35949409	30959749	86.1%	3668572	11.8%	45
S40A_8h	2	33798531	24387908	72.2%	3317849	13.6%	31
S133A_8h	1	34109018	29721320	87.1%	3514517	11.8%	34
S133A_8h	2	36072611	27299133	75.7%	3541779	13.0%	38
S40A-S133A_8h	1	33474368	28320199	84.6%	3300468	11.7%	22
S40A-S133A_8h	2	30656808	22663665	73.9%	2961173	13.1%	24

Table 5.17 shows an overview of the alignments of all the different RNA-Seq samples using the TopHat program in Galaxy. The first column denotes the HSV treatment (GFP, WT, S40A, S133A, S40A-S133A) along with the length of time (4 hours or 8 hours of expression). The second column shows the replicate number. The third and fourth columns show the number of input reads and the number of reads mapped reads after TopHat mapping to the mm10 genome respectively. The % of input represents the percent of the input reads that were mapped uniquely by TopHat. The next Multiple Alignments (MA) and % of input columns represent the total number of reads and percentage of reads that did not map uniquely to the mm10 mouse genome. The final column shows the number of non-unique alignments that had over 20 multiple alignments.

Next, the transcript assemblies were merged using Cuffmerge. The differentially expressed genes were determined using Cuffdiff with a cutoff of $\alpha = 0.1$. The differential expression results were then indexed and visualized using CummeRbund (version 2.16.0) and R package version 3.3.2. Venn diagrams of differentially expressed genes were generated using the VennDiagram R program (version 1.6.17) or BioVenn.[94,95] Reported false discovery rates for the RNA-Seq from this study were calculated q-values.[96] Gene ontology analysis is described in the "WGCNA and Gene Ontology Analysis" section.

5.16.5 ChIP-Seq, RNA-Seq, and microarray comparative analysis

Identifying the frequency of CRE sites in promoters

We used cruzdb in Python (version 2.7) in order to find the number of half and full CRE sites (TGACG/CGTCA or TGACGTCA) present 5000 bases upstream and 500 bases downstream from the transcription start site of all the differentially-expressed genes of interest.[97] The regions were mapped onto the *mm9* reference genome.

Identifying ChIP-Seq peaks

For several ChIP-Seq data sets, a list of chromosomes and start and stop sites for peaks were provided in Excel format instead of a list of nearest genes. In order to annotate the peaks with the nearest gene, we used cruzdb in Python (version 2.7) to identify the nearest gene such that the peak is upstream of the gene of interest transcription start site in the correct orientation of the gene.[97] We used this procedure to annotate peaks for the

following ChIP-Seq data sets: CREB;[5] p300;[26] H3K4me3, mSin3a, TetC, TetN, OGT;[42] H3K4me3, *O*-GlcNAc, Tet2.[46]

For the histone modification data set from Sandberg and colleagues, the peak data was obtained from GEO database as ChIP vs. Input peaks with detection $p < 0.00001$ (GSM2281997-GSM2282000).[38] The data was converted to the BED format and then aligned to the *mm9* mouse RefSeq gene coordinates using Galaxy.[7] This procedure was used for the H3K27me3, H3K4me1, H3K27Ac, and H3K4me3 ChIP-Seq data from Sandberg and colleagues.[38] Finally, for the CREB ChIP-Seq data set originally generated by Kim and colleagues, we obtained the reanalyzed peak data set from Lesiak and coworkers with FDR < 0.001.[5,24] Then, we used cruzdb in Python to annotate the new peaks as described above.

Microarray annotation and other RNA-Seq data set comparison

Several microarray data sets reported the data sets with microarray probe IDs. These data sets were annotated in R using the correct Affymetrix array annotation programs. For the McClung and Nestler *Nature Neuroscience* (2003) paper and Barco, *et al Neuron* (2005), we used the mgu74av2.db (version 3.2.3) and mgu74a.db (version 3.2.3) respectively.[16,28,98,99] For the Qiu, *et al eLife* (2016) RNA-Seq data set comparison, we used a cutoff of alpha=0.1 for differentially-expressed genes (KCl depolarized/ not depolarized) from the mouse DIV4 or DIV10 data sets.[100]

For the following data sets, we used the published annotated data without any pre-processing:

- H3K4me1, CBP, Npas4, Pol2, and SRF ChIP-Seq; RNA-Seq (KCl/no KCl)[5]
- 5hmC and H3K27me3 ChIP-Seq; RNA-Seq (NPC/neuron)[47]

- CREB ChIP-Seq[22]
- H3K27Ac ChIP-Seq[44]
- Tet2, OGT, and H2B-S112-*O*-GlcNAc ChIP-Seq[45]
- Tet3 ChIP-Seq[56]
- H3K4me3, H3K27me3 ChIP-Seq[43]

Program code for analysis is available upon request.

5.16.6 WGCNA and gene ontology analysis

Prior to performing WGCNA, we took the TopHat output from Galaxy and processed the data using HTSeq from the samtools program in order to determine the total counts for all the different samples.[101] This entailed sorting and indexing the BAM files and finally creating SAM files in command line for downstream analysis.[102] Next, the sorted SAM files were processed by HTSeq's counting program in order to get the final count for all the genes. Finally, we performed a data normalization using the variance stabilization transformation from DESeq2 (version 1.14.1) in R.[102-104]

Using the preprocessed RNA-Seq data, we performed WGCNA (version 1.51) in R on all detected genes using the protocols previously described.[59,105] Following hierarchical clustering and module assignment, gene ontology enrichment analysis was performed using the Bioconductor R packages AnnotationDbi (version 1.36.2), GO.db (version 3.4.0), and org.Mm.eg.db (version 3.4.0).[106-108] In addition, the Database for Visualization and Integrated Discovery (DAVID), PANTHER, and Enrichr as previously described.[109-111] The package Cytoscape version 3.5.1 was used to visualize the differentially-expressed gene ontology annotations using ClueGO version 2.3.3 and CluePedia version 1.3.3.[112-114] The package VisANT version 5.0 (weight cutoff of 0.1

(cor > 0.1)), Cytoscape, ClueGO, and CluePedia was used to visualize the WGCNA gene networks.[112-115] The top "hub" genes were identified as the genes with the most intramodular interactors with an edge weight cutoff of 0.1. The total number of interactors is listed in the Tables 5.13-5.15 above (all hub genes had at least 30 interactors).

5.17 References

1. 1 Khidekel, N. & Hsieh-Wilson, L.C. A 'molecular switchboard'--covalent modifications to proteins and their impact on transcription. *Org. Biomol. Chem.* (2004) **2**, 1:1-7.
2. 2 Sakamoto, K., Karelina, K. & Obrietan, K. CREB: a multifaceted regulator of neuronal plasticity and protection. *J. Neurochem.* (2011) **116**, 1:1-9.
3. 3 Rexach, J.E. *et al.* Dynamic O-GlcNAc modification regulates CREB-mediated gene expression and memory formation. *Nat. Chem. Biol.* (2012) **8**, 3:253-261.
4. 4 Blendy, J.A., Kaestner, K.H., Schmid, W., Gass, P. & Schutz, G. Targeting of the CREB gene leads to up-regulation of a novel CREB mRNA isoform. *EMBO J.* (1996) **15**, 5:1098-1106.
5. 5 Kim, T.K. *et al.* Widespread transcription at neuronal activity-regulated enhancers. *Nature* (2010) **465**, 7295:182-187.
6. 6 Trapnell, C. *et al.* Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotech* (2013) **31**, 1:46-53.
7. 7 Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* (2016) **44**, W1:W3-W10.
8. 8 Theisen, U., Straube, E. & Straube, A. Directional persistence of migrating cells requires Kif1C-mediated stabilization of trailing adhesions. *Dev. Cell* (2012) **23**, 6:1153-1166.
9. 9 Meng, J. *et al.* Myrf ER-Bound Transcription Factors Drive C. elegans Synaptic Plasticity via Cleavage-Dependent Nuclear Translocation. *Dev. Cell* (2017) **41**, 2:180-194 e187.
10. 10 Zheng, X. *et al.* CNOT3-Dependent mRNA Deadenylation Safeguards the Pluripotent State. *Stem Cell Reports* (2016) **7**, 5:897-910.
11. 11 Kuo, T.F. *et al.* Protein disulfide isomerase a4 acts as a novel regulator of cancer growth through the procaspase pathway. *Oncogene* (2017).
12. 12 Jensen, L.E. & Whitehead, A.S. IRAK1b, a novel alternative splice variant of interleukin-1 receptor-associated kinase (IRAK), mediates interleukin-1 signaling and has prolonged stability. *J. Biol. Chem.* (2001) **276**, 31:29037-29044.
13. 13 Diaz-Perales, A. *et al.* Identification of human aminopeptidase O, a novel metalloprotease with structural similarity to aminopeptidase B and leukotriene A4 hydrolase. *J. Biol. Chem.* (2005) **280**, 14:14310-14317.
14. 14 Simms, B.A. & Zamponi, G.W. Neuronal voltage-gated calcium channels: structure, function, and dysfunction. *Neuron* (2014) **82**, 1:24-45.
15. 15 Flavell, S.W. & Greenberg, M.E. Signaling mechanisms linking neuronal activity to gene expression and plasticity of the nervous system. *Annu. Rev. Neurosci.* (2008) **31**, 1:563-590.
16. 16 Barco, A. *et al.* Gene expression profiling of facilitated L-LTP in VP16-CREB mice reveals that BDNF is critical for the maintenance of LTP and its synaptic capture. *Neuron* (2005) **48**, 1:123-137.
17. 17 Suzuki, A. *et al.* Upregulation of CREB-mediated transcription enhances both short- and long-term memory. *J. Neurosci.* (2011) **31**, 24:8786-8802.
18. 18 Stephan, A.H., Barres, B.A. & Stevens, B. The complement system: an unexpected role in synaptic pruning during development and disease. *Annu. Rev. Neurosci.* (2012) **35**, 369-389.

19. 19 Cachat, J., Deffert, C., Hugues, S. & Krause, K.H. Phagocyte NADPH oxidase and specific immunity. *Clin. Sci. (Lond.)* (2015) **128**, 10:635-648.
20. 20 Benito, E. & Barco, A. CREB's control of intrinsic and synaptic plasticity: implications for CREB-dependent memory models. *Trends Neurosci.* (2010) **33**, 5:230-240.
21. 21 Richards, J.P., Bachinger, H.P., Goodman, R.H. & Brennan, R.G. Analysis of the structural properties of cAMP-responsive element-binding protein (CREB) and phosphorylated CREB. *J. Biol. Chem.* (1996) **271**, 23:13716-13723.
22. 22 Everett, L.J. *et al.* Integrative genomic analysis of CREB defines a critical role for transcription factor networks in mediating the fed/fasted switch in liver. *BMC Genomics* (2013) **14**, 1:337.
23. 23 Briand, L.A., Lee, B.G., Lelay, J., Kaestner, K.H. & Blendy, J.A. Serine 133 phosphorylation is not required for hippocampal CREB-mediated transcription and behavior. *Learn. Mem.* (2015) **22**, 2:109-115.
24. 24 Lesiak, A. *et al.* A genome-wide screen of CREB occupancy identifies the RhoA inhibitors Par6C and Rnd3 as regulators of BDNF-induced synaptogenesis. *PLoS One* (2013) **8**, 6:e64658.
25. 25 Hansen, K.F., Sakamoto, K., Pelz, C., Impey, S. & Obrietan, K. Profiling status epilepticus-induced changes in hippocampal RNA expression using high-throughput RNA sequencing. *Sci. Rep.* (2014) **4**, 6930.
26. 26 Visel, A. *et al.* ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* (2009) **457**, 7231:854-858.
27. 27 Hallenborg, P. *et al.* MDM2 facilitates adipocyte differentiation through CRTC-mediated activation of STAT3. *Cell Death Dis.* (2016) **7**, 6:e2289.
28. 28 McClung, C.A. & Nestler, E.J. Regulation of gene expression and cocaine reward by CREB and DeltaFosB. *Nat. Neurosci.* (2003) **6**, 11:1208-1215.
29. 29 Benito, E., Valor, L.M., Jimenez-Minchan, M., Huber, W. & Barco, A. cAMP response element-binding protein is a primary hub of activity-driven neuronal gene expression. *J. Neurosci.* (2011) **31**, 50:18237.
30. 30 Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* (1996) **87**, 5:953-959.
31. 31 Mejia, L.A. *et al.* A novel Hap1-Tsc1 interaction regulates neuronal mTORC1 signaling and morphogenesis in the brain. *J. Neurosci.* (2013) **33**, 46:18015-18021.
32. 32 Zhao, S., Fung-Leung, W.P., Bittner, A., Ngo, K. & Liu, X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One* (2014) **9**, 1:e78644.
33. 33 Zhang, J. *et al.* CREB-mediated synaptogenesis and neurogenesis is crucial for the role of 5-HT1a receptors in modulating anxiety behaviors. *Sci. Rep.* (2016) **6**, 29551.
34. 34 Mantamadiotis, T. *et al.* Disruption of CREB function in brain leads to neurodegeneration. *Nat. Genet.* (2002) **31**, 1:47-54.
35. 35 Viosca, J. *et al.* Chronic enhancement of CREB activity in the hippocampus interferes with the retrieval of spatial information. *Learn. Mem.* (2009) **16**, 3:198-209.
36. 36 Pramod, A.B., Foster, J., Carvelli, L. & Henry, L.K. SLC6 transporters: structure, function, regulation, disease association and therapeutics. *Mol. Aspects Med.* (2013) **34**, 2-3:197-219.
37. 37 Hirano, Y. *et al.* Shifting transcriptional machinery is required for long-term memory maintenance and modification in Drosophila mushroom bodies. *Nature Communications* (2016) **7**, 13471.
38. 38 Sandberg, M. *et al.* Transcriptional Networks Controlled by NKX2-1 in the Development of Forebrain GABAergic Neurons. *Neuron* (2016) **91**, 6:1260-1275.
39. 39 Bernhart, S.H. *et al.* Changes of bivalent chromatin coincide with increased expression of developmental genes in cancer. *Sci. Rep.* (2016) **6**, 37393.
40. 40 Kolasinska-Zwierz, P. *et al.* Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat. Genet.* (2009) **41**, 3:376-381.
41. 41 Nguyen, A.T. & Zhang, Y. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev.* (2011) **25**, 13:1345-1358.
42. 42 Vella, P. *et al.* Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells. *Mol. Cell* (2013) **49**, 4:645-656.
43. 43 Mikkelsen, T.S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* (2007) **448**, 7153:553-560.

44. 44 Malik, A.N. *et al.* Genome-wide identification and characterization of functional neuronal activity-dependent enhancers. *Nat. Neurosci.* (2014) **17**, 10:1330-1339.
45. 45 Chen, Q., Chen, Y., Bian, C., Fujiki, R. & Yu, X. TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature* (2013) **493**, 7433:561-564.
46. 46 Deplus, R. *et al.* TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J.* (2013) **32**, 5:645-655.
47. 47 Hahn, M.A. *et al.* Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis. *Cell Rep* (2013) **3**, 2:291-300.
48. 48 Marson, A. *et al.* Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* (2008) **134**, 3:521-533.
49. 49 Ku, M. *et al.* Genome-wide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* (2008) **4**, 10:e1000242.
50. 50 Myers, S.A., Panning, B. & Burlingame, A.L. Polycomb repressive complex 2 is necessary for the normal site-specific O-GlcNAc distribution in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* (2011) **108**, 23:9490-9495.
51. 51 Gambetta, M.C., Oktaba, K. & Muller, J. Essential role of the glycosyltransferase sxc/Ogt in polycomb repression. *Science* (2009) **325**, 5936:93-96.
52. 52 Maury, J.J. *et al.* RING1B O-GlcNAcylation regulates gene targeting of polycomb repressive complex 1 in human embryonic stem cells. *Stem Cell Res* (2015) **15**, 1:182-189.
53. 53 Jang, H. *et al.* O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. *Cell Stem Cell* (2012) **11**, 1:62-74.
54. 54 Myers, S.A. *et al.* SOX2 O-GlcNAcylation alters its protein-protein interactions and genomic occupancy to modulate gene expression in pluripotent cells. *Elife* (2016) **5**, e10647.
55. 55 Chu, C.S. *et al.* O-GlcNAcylation regulates EZH2 protein stability and function. *Proc. Natl. Acad. Sci. U. S. A.* (2014) **111**, 4:1355-1360.
56. 56 Jin, S.G. *et al.* Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. *Cell Rep* (2016) **14**, 3:493-505.
57. 57 Saunders, A. *et al.* The SIN3A/HDAC Corepressor Complex Functionally Cooperates with NANOG to Promote Pluripotency. *Cell Rep* (2017) **18**, 7:1713-1726.
58. 58 Rudenko, A. *et al.* Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. *Neuron* (2013) **79**, 6:1109-1122.
59. 59 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* (2008) **9**, 559.
60. 60 Altarejos, J.Y. & Montminy, M. CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. *Nat. Rev. Mol. Cell Biol.* (2011) **12**, 3:141-151.
61. 61 Herzig, S. *et al.* CREB regulates hepatic gluconeogenesis through the coactivator PGC-1. *Nature* (2001) **413**, 6852:179-183.
62. 62 Langfelder, P., Mischel, P.S. & Horvath, S. When is hub gene selection better than standard meta-analysis? *PLoS One* (2013) **8**, 4:e61505.
63. 63 Rouault, T.A. Iron metabolism in the CNS: implications for neurodegenerative diseases. *Nat Rev Neurosci* (2013) **14**, 8:551-564.
64. 64 Choi, D.S. *et al.* The type 1 equilibrative nucleoside transporter regulates ethanol intoxication and preference. *Nat Neurosci* (2004) **7**, 8:855-861.
65. 65 Lin, J.R., Zeman, M.K., Chen, J.Y., Yee, M.C. & Cimprich, K.A. SHPRH and HLTF act in a damage-specific manner to coordinate different forms of postreplication repair and prevent mutagenesis. *Mol Cell* (2011) **42**, 2:237-249.
66. 66 Motegi, A. *et al.* Polyubiquitination of proliferating cell nuclear antigen by HLTF and SHPRH prevents genomic instability from stalled replication forks. *Proc Natl Acad Sci U S A* (2008) **105**, 34:12411-12416.
67. 67 Unk, I. *et al.* Human HLTF functions as a ubiquitin ligase for proliferating cell nuclear antigen polyubiquitination. *Proceedings of the National Academy of Sciences of the United States of America* (2008) **105**, 10:3768-3773.
68. 68 Kile, A.C. *et al.* HLTF's Ancient HIRAN Domain Binds 3' DNA Ends to Drive Replication Fork Reversal. *Mol Cell* (2015) **58**, 6:1090-1100.

69. 69 Wang, L. *et al.* Card10 is a novel caspase recruitment domain/membrane-associated guanylate kinase family member that interacts with BCL10 and activates NF-kappa B. *J Biol Chem* (2001) **276**, 24:21405-21409.
70. 70 Zhang, X. *et al.* Arylsulfatase B modulates neurite outgrowth via astrocyte chondroitin-4-sulfate: dysregulation by ethanol. *Glia* (2014) **62**, 2:259-271.
71. 71 Bhattacharyya, S., Feferman, L. & Tobacman, J.K. Restriction of Aerobic Metabolism by Acquired or Innate Arylsulfatase B Deficiency: A New Approach to the Warburg Effect. *Sci Rep* (2016) **6**, 32885.
72. 72 Lujan, R., Marron Fernandez de Velasco, E., Aguado, C. & Wickman, K. New insights into the therapeutic potential of Girk channels. *Trends Neurosci* (2014) **37**, 1:20-29.
73. 73 Munoz, M.B. *et al.* A role for the GIRK3 subunit in methamphetamine-induced attenuation of GABAB receptor-activated GIRK currents in VTA dopamine neurons. *J Neurosci* (2016) **36**, 11:3106-3114.
74. 74 Golonzhka, O. *et al.* Pbx regulates patterning of the cerebral cortex in progenitors and postmitotic neurons. *Neuron* (2015) **88**, 6:1192-1207.
75. 75 Villaescusa, J.C. *et al.* A PBX1 transcriptional network controls dopaminergic neuron development and is impaired in Parkinson's disease. *The EMBO Journal* (2016) **35**, 18:1963.
76. 76 Muralidharan, B. *et al.* LHX2 Interacts with the NuRD Complex and Regulates Cortical Neuron Subtype Determinants Fezf2 and Sox11. *J Neurosci* (2017) **37**, 1:194-203.
77. 77 Yasushi, O. *et al.* SAFB1, an RBMX-binding protein, is a newly identified regulator of hepatic SREBP-1c gene. *BMB Rep.* (2009) **42**, 4:232-237.
78. 78 Tsend-Ayush, E. *et al.* RBMX gene is essential for brain development in zebrafish. *Dev Dyn* (2005) **234**, 3:682-688.
79. 79 Shashi, V. *et al.* The RBMX gene as a candidate for the Shashi X-linked intellectual disability syndrome. *Clin Genet* (2015) **88**, 4:386-390.
80. 80 Olkkonen, V.M. & Li, S. Oxysterol-binding proteins: sterol and phosphoinositide sensors coordinating transport, signaling and metabolism. *Prog Lipid Res* (2013) **52**, 4:529-538.
81. 81 Prentice, H., Modi, J.P. & Wu, J.Y. Mechanisms of Neuronal Protection against Excitotoxicity, Endoplasmic Reticulum Stress, and Mitochondrial Dysfunction in Stroke and Neurodegenerative Diseases. *Oxid. Med. Cell. Longev.* (2015) **2015**, 964518.
82. 82 Shaywitz, A.J., Dove, S.L., Kornhauser, J.M., Hochschild, A. & Greenberg, M.E. Magnitude of the CREB-dependent transcriptional response is determined by the strength of the interaction between the kinase-inducible domain of CREB and the KIX domain of CREB-binding protein. *Mol. Cell. Biol.* (2000) **20**, 24:9409-9422.
83. 83 Hirano, Y. *et al.* Fasting launches CRTC to facilitate long-term memory formation in Drosophila. *Science* (2013) **339**, 6118:443-446.
84. 84 Uchida, S. *et al.* CRTC1 Nuclear Translocation Following Learning Modulates Memory Strength via Exchange of Chromatin Remodeling Complexes on the Fgf1 Gene. *Cell Rep* (2017) **18**, 2:352-366.
85. 85 Nonaka, M. *et al.* Region-specific activation of CRTC1-CREB signaling mediates long-term fear memory. *Neuron* (2014) **84**, 1:92-106.
86. 86 Lopez de Armentia, M. *et al.* cAMP response element-binding protein-mediated gene expression increases the intrinsic excitability of CA1 pyramidal neurons. *J. Neurosci.* (2007) **27**, 50:13909-13918.
87. 87 Obrietan, K. & Hoyt, K.R. CRE-mediated transcription is increased in Huntington's disease transgenic mice. *J. Neurosci.* (2004) **24**, 4:791-796.
88. 88 Saura, C. & Valero, J. in *Rev. Neurosci.* Vol. 22 153 (2011).
89. 89 Briand, L.A. & Blendy, J.A. Not all stress is equal: CREB is not necessary for restraint stress reinstatement of cocaine-conditioned reward. *Behav. Brain Res.* (2013) **246**, 63-68.
90. 90 Pandey, S.C., Roy, A., Zhang, H. & Xu, T. Partial deletion of the cAMP response element-binding protein gene promotes alcohol-drinking behaviors. *J. Neurosci.* (2004) **24**, 21:5022-5030.
91. 91 Bourtchuladze, R. *et al.* Deficient long-term memory in mice with a targeted mutation of the cAMP-responsive element-binding protein. *Cell* (1994) **79**, 1:59-68.
92. 92 Hummler, E. *et al.* Targeted mutation of the CREB gene: compensation within the CREB/ATF family of transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* (1994) **91**, 12:5647-5651.

93. 93 Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* (2012) **7**, 3:562-578.
94. 94 Chen, H. & Boutros, P.C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* (2011) **12**, 1:35.
95. 95 Hulsen, T., de Vlieg, J. & Alkema, W. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* (2008) **9**, 488.
96. 96 Storey, J.D., Taylor, J.E. & Siegmund, D. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. *Journal of the Royal Statistical Society Series B-Statistical Methodology* (2004) **66**, 1:187-205.
97. 97 Pedersen, B.S., Yang, I.V. & De, S. CruzDB: software for annotation of genomic intervals with UCSC genome-browser database. *Bioinformatics* (2013) **29**, 23:3003-3006.
98. 98 mgu74av2.db: Affymetrix Murine Genome U74v2 annotation data (chip mgu74av2) v. Version 3.2.3 (2016).
99. 99 mgu74a.db: Affymetrix Murine Genome U74v2 annotation data (chip mgu74a) v. Version 3.2.3 (2016).
100. 100 Qiu, J. *et al.* Evidence for evolutionary divergence of activity-dependent gene expression in developing neurons. *Elife* (2016) **5**, e20337.
101. 101 Anders, S., Pyl, P.T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* (2015) **31**, 2:166-169.
102. 102 Anders, S. *et al.* Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc.* (2013) **8**, 9:1765-1786.
103. 103 McCarthy, D.J., Chen, Y. & Smyth, G.K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* (2012) **40**, 10:4288-4297.
104. 104 Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* (2014) **15**, 12:550.
105. 105 Langfelder, P. & Horvath, S. Fast R Functions for Robust Correlations and Hierarchical Clustering. *J Stat Softw* (2012) **46**, 11:17.
106. 106 GO.db: A set of annotation maps describing the entire Gene Ontology v. Version 3.4.1 (2017).
107. 107 AnnotationDbi: Annotation Database Interface. v. 1.38.1. (2017).
108. 108 org.Mm.eg.db: Genome wide annotation for Mouse v. Version 3.4.1 (2017).
109. 109 Mi, H., Muruganujan, A., Casagrande, J.T. & Thomas, P.D. Large-scale gene function analysis with the PANTHER classification system. *Nat. Protoc.* (2013) **8**, 8:1551-1566.
110. 110 Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* (2009) **4**, 1:44-57.
111. 111 Kuleshov, M.V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* (2016) **44**, W1:W90-97.
112. 112 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* (2003) **13**, 11:2498-2504.
113. 113 Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* (2009) **25**, 8:1091-1093.
114. 114 Bindea, G., Galon, J. & Mlecnik, B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics* (2013) **29**, 5:661-663.
115. 115 Hu, Z. *et al.* VisANT 4.0: Integrative network platform to connect genes, drugs, diseases and therapies. *Nucleic Acids Res.* (2013) **41**, Web Server issue:W225-231.