# Improving Outcomes in First Episode Psychosis: The Role of Genetic Variation of *NPAS3*

by

Leiah Marie Luoma

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Medical Sciences - Medical Genetics
University of Alberta

# Abstract

*NPAS3* encodes a transcription factor of the bHLH-PAS family, which has been robustly associated with neurodevelopmental and psychiatric disorders with intellectual disability as a common feature. *NPAS3* was initially discovered as a schizophrenia-associated gene in a family with a translocation breaking only *NPAS3*. Since then, deletions encompassing *NPAS3* have been associated with neurodevelopmental disorders such as holoprosencephaly, holoprosencephaly microforme and Sotos syndrome. SNPs linked to, and within, *NPAS3* have been associated with neuropsychiatric disorders, including schizophrenia and bipolar disorder, as well as antipsychotic responsiveness. A common feature of all of these disorders is intellectual disability. Deletion of *NPAS3* in mice results in marked behavioural and neuroanatomical phenotypes reminiscent of neuropsychiatric disorders, including increased anxiety, altered behavior and deficits in learning and memory. Further, deficits in neurodevelopment *in utero* resulting in reduced interneuron number, smaller cortical volume, increased ventricles and altered cerebellar foliation, as well as deficits in adult neurogenesis in the dentate gyrus of the hippocampus, have been observed. Despite the robustness of the evidence that *NPAS3* is associated with neuropsychological function and neuropsychiatric dysfunction, its molecular function is not well characterized.

Here, I set out to functionally characterize NPAS3 as a transcription factor of the bHLH-PAS family. I assessed its ability to interact with its presumed obligate heterodimeric partner, ARNT, as well as its ability to directly bind and regulate genes it has been found to genetically regulate. I have determined that NPAS3 does functionally interact with ARNT to regulate two target genes directly: *VGF* and *TXNIP,* both of which

have functions relevant to cell survival and neuronal functions. Through the assays developed to characterize NPAS3 function, I have characterized the function of its predicted domains and found them to participate in DNA binding, protein::protein interaction and transactivation, consistent with the categorization of NPAS3 as a bHLH-PAS transcription factor. Assessment of NPAS3 function at the cellular level was performed and *NPAS3* expression at the mRNA level was found to respond to soluble factors, as well as cell contact, suggesting a role of NPAS3 in environmental response.

As the disorders with which *NPAS3* has been associated share a common feature of intellectual disability, a collaborative study was undertaken to assess the contribution of *NPAS3* variation to normal variation in neurocognition in order to understand its role in neuropsychological function, both normal and disordered. To that end, we have assessed the contribution of exonic variation of *NPAS3* to cognitive function in a cross-sectional population of Western Canadian young adults. The minor alleles of a three SNP haplotype within exon 12 of *NPAS3* were found to be associated with worse performance on a test of verbal working memory, and a trend was observed with generally reduced cognitive function, as assessed by the Screen for Cognitive Impairment in Psychiatry. Taken together, these data expand our understanding of the role of *NPAS3* in normal cognitive functioning and neuropsychiatric dysfunction. NPAS3 has been found to directly regulate two genes with roles in neuronal survival, stress/injury response and metabolic balance. These data support the role of NPAS3 in the survival of proliferating neuroprogenitors during neurodevelopment and neurogenic processes, which can be predicted to contribute to the intellectual disability observed when *NPAS3* expression is reduced or lost.

# Preface

This thesis is an original work by Leiah M. Luoma. The research project involving cognitive assessment of a cross-sectional population of youth received approval from the University of Alberta Research Ethics Board, Project Title “Pathways to Care in First Episode Psychosis”, No. MS4_Pro00003634, Date: 2009-2012.

Some of the research presented here was performed as a part of a collaboration led by Dr Scot Purdon, University of Alberta Department of Psychiatry. Initial population acquisition was led by Dr T Cameron Wild, University of Alberta School of Public Health. The genotyping work was developed in the laboratory of Dr. Diane W. Cox, performed by myself. The neuropsychological testing was performed by members of Dr. Purdon’s group, and the data analysis in Chapter 3 was performed by Dr. Purdon and myself.

*Dedicated to my Dad, Paavo Luoma,*

*he taught me everything I know.*

## Acknowledgements

First and foremost, I must acknowledge the funding agencies who have allowed me to complete my PhD studies. I started my studies funded by an NSERC Master's studentship, followed by a Doctoral studentship from Alberta Innovates (formerly AHFMR) and a Doctoral award from CIHR. I am sincerely grateful for the continued support as I had my family over the course of this degree. I would also like to acknowledge the various sources of operating funding involved in this study, including an Emerging Team Grant from the Faculty of Medicine, as well as a CIHR operating grant for the molecular studies of NPAS3.

I'd like to acknowledge my supervisor, Dr Fred Berry, for being willing to support me in completing my degree, and for having faith in me. I am sincerely grateful that he was willing to take me on, and to let me focus on the project that I had already started, despite all the roadblocks along the way. I would also like to acknowledge my supervisory committee members, Dr Sarah Hughes and Dr Scot Purdon. Both of your insights into very disparate areas of study which I have attempted to correlate have been instrumental in the completion of this project and the construction of this thesis. Sincere thanks to my examiners, Dr Marc Ekker and Dr Roseline Godbout who have examined my thesis and improved it through careful scrutiny.

Thanks to the members of the Berry lab for the companionship and support over the course of the years. To Freda Mirzayans, for technical assisstance, emotional support, and always having a smile to share. To Rotem Lavy, for helping me have a "day away from the lab" here and there, and for being stuck in the same boat and somehow still managing to find the energy to listen to my venting.

I would like to acknowledge the members of Dr Purdon's group at Alberta Hospital Edmonton for their work to coordinate data collection and analysis. I'd like to thank Dan LaFreniere for training me on the cognitive testing paradigms so that I understood how they are administered, and how rigorous they are. Further, I thank Dr Phil Tibbo, Dr Ian Colman and Dr T Cameron Wild for their input, and the work they put in to coordinate the team study lead by Dr Purdon. I am grateful to the study participants for volunteering their time and samples.

The members of the former Cox lab, where I started this program, also should be acknowledged. Thanks to Dr Diane Cox for agreeing to take me on, and for initiating the project. Thanks especially to Dr Georgina Macintyre, who was instrumental in my early training and in initial project design. Thanks to Anna Zeisler for helping to generate some of the *NPAS3* clones and constructs used in this thesis, for reading and revising my thesis, and for being such a good friend. Thanks to Darren Bugbee and to Aleks Dimitrijevic for assistance with genotyping. Thanks to the Department of Medical Genetics for facilitating me transfering to, and completing a PhD.

Thanks to my friend Dr Julia Wong, for being there to chat about science and everything else. Your optimism, poise and grit were major motivation for me to finish.

Now, I would like to thank my family. First, I must thank my partner, Steve, who was steadfast and a stable influence even when I was at my least stable. Without you I would not be where I am, I love you. Thanks to my mom, Hilkka, who revised my thesis, counseled me well, and babysat for countless hours, through sickness and health. Finally, thanks to my children Ewan and Ilona, who constantly remind me that endless curiosity and relentless iterative repetition are the basic tenets of scientific exploration of life.

# Table of Contents

# List of Tables

# List of Figures

# List of Abbreviations

α – alpha, anti
aa – amino acid(s)
AhR – aryl hydrocarbon receptor, also known as the dioxin receptor
ANKRD37 – ankyrin domain-containing 37
ARNT – aryl hydrocarbon receptor nuclear translocator
bFGF – basic fibroblast growth factor
bHLH – basic helix loop helix domain
BSA – bovine serum albumin
CGE – caudal ganglionic eminence
ChIP – chromatin immunoprecpitation
CNS – central nervous system
CNV – copy number variation
COMT – catechol-O-methyltransferase
CPT – Connors Continuous Performance Task
CR – calretinin
Δ – delta, deletion
DCX – doublecourtin
dpc – days post coitum
DR – dioxin receptor, also known as the aryl hydrocarbon receptor
E# – embryonic day #
EV – empty vector control
FGF – fibroblast growth factor
FGFR1 – fibroblast growth factor receptor 1
GABA – gamma aminobutyric acid
GWAS – genome-wide association study
HA – hemagglutinin epitope tag
HIF1A – hypoxia inducible factor 1 alpha
IP – immunoprecipitation
LB – lysogeny broth
LGE – lateral ganglionic eminence
LIF – leukemia inhibitory factor
MGE – medial ganglionic eminence
miRNA – micro RNA
NFY – nuclear factor Y
NPAS3 – neuronal PAS domain containing 3
NPY – neuropeptide Y
P# – postnatal day #
PAS – Period, ARNT, Single-minded domain
PBS – phosphate buffered saline
PV – parvalbumin
SCIP – Screen for Cognitive Impairment in Psychiatry
shRNA – short hairpin RNA
SNP – single nucleotide polymorphism
SOC – super optimal broth with catabolite repression

SST - somatostatin
TAD – transactivation domain
TAE – tris acetate EDTA buffer
TBE – tris borate EDTA buffer
TBS – tris buffered saline
TE – 10 mM Tris pH 8.0 1 mM EDTA buffer
TSS – transcription start site
TXNIP – thioredoxin interacting protein
V5 – V5 epitope tag
VGF – VGF, nerve growth factor inducible (nonacronymic)
VFT – verbal fluency test
VIP – vasointenstinal peptide
VLT – verbal learning test
VMT – visuomotor tracking test
VPT – Visual Patterns Test
WMT – Working Memory Task
WRAT4_R – Wide Range Achievement Test, reading subtest
ZT – zeitgerber time

# Chapter 1: Introduction

## 1.1 Genetics of neuropsychiatric disorders

Schizophrenia is a neuropsychatric disorder with a prevalence of approximately 1% with onset typically in youth and young adulthood. The symptoms of schizophrenia include hallucinations and delusions, anhedonia, avolition, as well as markedly reduced cognitive function (Tandon, Keshavan, Nasrallah 2008). Due to the severity of the disorder and the early age of onset, there is significant societal cost and severe morbidity and mortality to individuals affected (Public Health Agency of Canada 2012). Due to the early onset resulting in severe personal and familial suffering, relatively poor outcomes, and high social costs, significant research has gone into understanding the etiology and biology of the disease in an effort to improve patient outcomes.

Heritability estimates of schizophrenia range up to 80%, suggesting that there is a significant genetic contribution to its pathology (Sullivan, Kendler, Neale 2003). Initial attempts to understand the genetic underpinnings of schizophrenia were undertaken using classical genetic methods, including linkage studies, where disease traits were mapped to linked genetic loci in families. The candidate gene approach was also used, where the genetic cause of heritable human diseases is screened for based on assessing variation in genes predicted to cause disease in affected individuals relative to unaffected individuals (Alaerts and Del-Favero 2009). By identifying genes affected in individuals with disease and then characterizing their normal functions, researchers aim to understand both the normal biological function of affected gene products and pathways, as well as the pathogenesis of disease. Identification of a ‘schizophrenia gene’ would expand our

understanding of the molecular etiology of schizophrenia, as well as of normal neuropsychological functioning. Through the linkage and candidate gene approaches, many genes have been associated with the pathology of schizophrenia, including *DISC1*, *BDNF*, *COMT* and *NRG1* (Farrell et al. 2015). These genes have largely not been validated as contributing to schizophrenia in the population (Farrell et al. 2015; Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014; Sullivan, Daly, O'Donovan 2012). These data indicate that the genetic etiology of schizophrenia is complex and multigenic.

Genome-wide studies have implicated multiple genes in neuropsychiatric dysfunction and further underline the complexity of the genetic architecture of psychiatric disorders (Giegling et al. 2017; Kato 2015). Attempts to understand the genetic underpinnings of schizophrenia through genome-wide association studies (GWAS) have resulted in the identification of common SNPs linked to over 100 genes affecting multiple pathways enriched in processes relevant for brain and immune function (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). Whole exome sequencing (sequencing of all coding DNA in the genome), which allows for identification of *de novo* and rare variants that may contribute to neuropsychiatric disorders, has similarly identified multiple loci as implicated in neuropsychiatric dysfunction (Fromer et al. 2014; Giegling et al. 2017; Kato 2015; Purcell et al. 2014). The identification of many genes as implicated in pathology is unsurprising given the complex pathogenesis and etiology of neuropsychiatric disorders including schizophrenia. The overlapping loci identified through genome-wide studies of multiple neuropsychiatric disorders indicates that many disorders exist on a continuum with

overlapping molecular/genetic architecture, with contributions by both common and rare variants (Burmeister, McInnis, Zollner 2008; Fromer et al. 2014; Sullivan, Daly, O'Donovan 2012).

Diagnostic criteria and clinical heterogeneity confound interpretation and significance of variants identified as associated with schizophrenia and other neuropsychiatric disorders (Liang and Greenwood 2015; Wray, Lee, Kendler 2012). Care must be taken in interpreting GWAS data, as the presence of secondary disorders in patient populations have been shown to affect linkage/non-linkage of variants to specific disorders (Liang and Greenwood 2015; Vieland et al. 2014). Furthermore, the presence of mental illness in control populations, even if considered unrelated to the disorder under investigation, has been found to affect the significance of findings (Liang and Greenwood 2015). The genetic data and effects of secondary diagnoses on genetic associations support the model that mental illness exists on a continuum, notably for this study, the continuum of psychotic and affective illness as these disorders share common heritable markers (Cross-Disorder Group of the Psychiatric Genomics Consortium et al. 2013; Keshavan et al. 2011; Liang and Greenwood 2015). There are several methods to minimize the confounding effects caused by clinical heterogeneity: increased sample sizes to improve power to detect small effects, rigorously selected populations, or assessment of a less heterogeneous/qualitative phenotype that contributes to the larger disorder(s) of interest.

The genetic data support the hypothesis that variants with more severe effects (eg copy number variants [CNVs] causing gene deletion) result in more severe deficits in neuropsychological function than common polymorphic single nucleotide variants, which

are expected to be associated with variants of much smaller effect sizes (Burmeister, McInnis, Zollner 2008; Marshall et al. 2017; Purcell et al. 2014). As such, the same gene may be associated with severe (deletion) and mild (polymorphism) neuropsychiatric dysfunction in overlapping but distinct disorders (Fromer et al. 2014). To simplify the assessment of the genetic contribution and cellular biology of complex traits, it is helpful to look at smaller causative components of these disorders. Endophenotypes are subclinical, reliably quantifiable, state independent, heritable traits that are associated with larger complex traits (Gottesman and Gould 2003; Gur et al. 2007). That endophenotypes are quantifiable and heritable greatly simplifies the association between genotype and phenotype.

## 1.2 Cognitive functions as endophenotypes of neuropsychiatric disorders

Deficits to cognitive functioning can severely affect daily functioning and have been found to be core features of a broad range of neuropsychatric disorders (Greenwood et al. 2016; Sahakian et al. 2015). Cognitive functions represent a powerful potential endophenotype of neuropsychiatric disorders, as they meet many criteria of endophenotypes: (1) variation in cognitive function is present in normal and psychiatrically ill populations, (2) variability of multiple cognitive functions is known to be heritable, (3) in psychiatrically ill populations, the cognitive deficit is commonly state independent (i.e. cognitive deficits are observed both during acute episodes of psychiatric illness, as well as during prodromal periods), and (4) deficits can be assessed by testing paradigms that can be quantitatively scored, facilitating objective assessment of cognitive

functioning that is consistent within and between studies (Greenwood et al. 2013; Greenwood et al. 2016; Gur et al. 2007; Zai et al. 2017).

Several cognitive functions have been assessed as endophenotypes of neuropsychiatric disorders by GWAS and preliminary evidence demonstrates significant heritability of many cognitive functions in both psychiatrically ill individuals and their family members, relative to unrelated controls (Greenwood et al. 2013; Greenwood et al. 2016). Genetic data to date suggest that cognitive testing batteries can be coordinated over multiple sites in order to detect genetic linkage (Greenwood et al. 2013; Greenwood et al. 2016). Although larger sample sizes are required for genome-wide assessment of specific variants and genes contributing to variation in cognition, cognitive function has been found to be an endophenotype of schizophrenia with shared genetic etiology (Lencz et al. 2014). Furthermore, cognitive function and intelligence are heritable, with estimates approximating 30% (Benyamin et al. 2014; Davies et al. 2011; Lencz et al. 2014). As such, cognitive functions can be seen to be a promising target to understand the genetic contribution to the cognitive underpinning of neuropsychological variability in function and dysfunction.

Verbal working memory is of particular interest in this study. Working memory is a short term memory process involved in retaining information that is being actively manipulated (Cowan 2008). Adequately functioning working memory is a critical cognitive function that underpins many other cognitive functions, notably those related to learning and memory (Baddeley 2003; Cuesta et al. 2011; Park and Gooding 2014; Schwarz, Tost, Meyer-Lindenberg 2015). Severe deficits in working memory have been associated with psychiatric pathology, such as psychosis and affective disorders

(Chapman and Chapman 1987; Cuesta et al. 2011). Deficits in verbal working memory have been identified in schizophrenia patients and their relatives, as compared to unrelated controls, demonstrating heritability (Horan et al. 2008). Both visual and verbal working memory function have been found to be highly heritable, although the heritability was found to be complex, likely owing to the many factors that contribute to working memory function (Ando, Ono, Wright 2001; Greenwood et al. 2007). Taken together, these data support assessment of the genetic contribution to working memory function in our understanding of how variants present in the normal population can affect normal (and abnormal) variation in neurocognition, as well as in the larger context of neuropsychological function and neuropsychiatric disorders. To this end, I will assess the contribution of variation in *NPAS3*, a gene that has been associated with neuropsychiatric disorders, to the common feature of the disorders to which it is associated: intellectual disability.

## 1.3 *NPAS3* is associated with disorders featuring intellectual disability

*NPAS3* (neuronal period-ARNT-single minded [PAS] domain containing 3) was originally identified as a gene of interest in human neuropsychiatric function when it was found to be the only gene broken by the reciprocal translocation t(9;14)(q34;q13) in a mother-proband family with significant intellectual impairment, as well as features of psychosis (Kamnasaran et al. 2003; Pickard et al. 2005). *NPAS3* encodes a transcription factor of the bHLH-PAS family, with conserved bHLH (basic helix-loop-helix) DNA-binding domain, and tandem PAS(A)-PAS(B) protein interaction domains (Brunskill et al. 1999). The breakpoint on chromosome 14 mapped to an interval within intron 4

(hg38 Chr14: 33,552,123-33,560,111, intron numbering based on transcript variant 1 NM_001164749.1) of the *NPAS3* gene and was predicted to cause expression of a truncated protein product containing only the DNA-binding domain separated from all other domains. The translocation has been hypothesized to cause the observed neuropsychiatric phenotypes due to either haploinsufficiency, or dominant-negative effects (Kamnasaran et al. 2003; Pickard et al. 2005). The breakpoint on chromosome 9 was mapped to an interval (hg38 Chr9:129,188,768-129,288,806) that had no known genes, and is currently annotated to contain two predicted long non-coding RNA (lncRNA) genes and one validated long non-coding RNA (lncRNA) with no functional annotation, nor any enrichment in neuronal expression (Kamnasaran et al. 2003; O'Leary et al. 2016). As such, the disruption to *NPAS3* was concluded to be causative of the observed neuropsychiatric disorder in the mother-daughter pair. In this family, the father also had affective illness (bipolar disorder), which may modify the observed phenotype in the proband, however, as the disruption of *NPAS3* segregates with psychosis and intellectual impairment in this family, it is likely to result in the observed phenotype in the mother-proband family carriers.

Since its discovery as a schizophrenia candidate gene, *NPAS3* has been repeatedly associated with a wide variety of neuropsychiatric disorders with intellectual disability as a common feature as well as with with severe neurological disorders (Kamnasaran et al. 2005; Phelps et al. 2016; Piccione et al. 2012; Rosenfeld et al. 2010; Visser et al. 2010). *NPAS3* is located within the 2.8Mb region of holoprosencephaly locus 8 (*HPE8*), a severe disorder of the formation of the forebrain (prosencephalon) and craniofacial structures (Kamnasaran et al. 2005). Microdeletions of this locus that involve *NPAS3*

have been associated both with holoprosencephaly, as well as holoprosencephaly microforme characterized by abnormalities in the corpus callosum, microcephaly, and developmental delay (Piccione et al. 2012; Rosenfeld et al. 2010). Furthermore, microdeletions affecting *NPAS3* have been associated with intellectual disability and microophthalmia without holoprosencephaly, suggesting that deletions to the *HPE8* locus may not be sufficient for holoprosencephaly in isolation, however, it is consistently associated with intellectual impairment and neurodevelopmental dysfunction (Fonseca et al. 2012; Piccione et al. 2012).

Deletions physically affecting only *NPAS3* have also been found to cause intellectual disability and neuropsychiatric disorders. A 155 kb deletion of *NPAS3* exon 1 and proximal intronic and promoter regions, physically affecting no other genes, has been identified in Sotos syndrome with intellectual disability as the primary phenotype. Although the deletion was also found in the proband's mother (borderline normal intelligence), in the eight siblings of the proband, learning disability was noted as prevalent (Visser et al. 2010). A 210 kb deletion described as affecting a portion of *NPAS3* (deletion interval not specified) was identified as the cause of the intellectual disability and disorganized schizophrenia in an 11 year old child (Phelps et al. 2016). These data indicate that larger deletions of this region are associated with gross neurological abnormalities, that deletions that include *NPAS3* appear to contribute to larger scale disorders, and that more limited loss of the *NPAS3* locus is associated with intellectual disability and psychiatric illness.

Similar to the large-scale deletions, common variants have linked *NPAS3* to neuropsychiatric disorders with features of intellectual impairment. A study of 70 SNPs

across the *NPAS3* locus identified multiple SNPs in four regions that act as haplotypes that are associated with both risk and protection from bipolar disorder and schizophrenia (Pickard et al. 2008). Although the association was found to be nominally significant, genotypes in the regions associated with protection or risk were found to be predictive of disease vs control status (Pickard et al. 2008). Analysis of *NPAS3* variation in a population of individuals with autism spectrum disorders identified seven non-synonymous variants in *NPAS3* as present only in the autism spectrum population (Stanco et al. 2014). In a cohort of schizophrenia patients, exonic variants in *NPAS3*, including a coding variant c.1654G>C (p.Ala552Pro, rs12434716) and two non-coding variants c.2208C>T (rs10141940) and c.2262C>G (rs10142034) were found to be enriched in the schizophrenia population relative to controls (Macintyre et al. 2010). In the same study, several other exonic variants, both synonymous and non-synonymous, were observed in the schizophrenia patient cohorts and/or in the control individuals at low frequencies. Similarly, a rare coding variant, c.910G>A (p.Val304Ile, rs146677388), was identified as co-segregating with schizophrenia in a small family with significant loading of neuropsychiatric illness (Yu et al. 2014). Thus at the level of exonic single nucleotide variants of uncertain significance, *NPAS3* has also been robustly linked to multiple neuropsychiatric disorders, strongly suggestive of its role in neuropsychological function.

GWAS are a method of querying variants across the genome to identify loci (either genotyped or linked regions) that are associated with disorders. GWAS have identified hundreds of loci associated with psychiatric disorders, however not all findings have been replicated. The variability in GWAS findings is contributed to by the interplay

of many factors, including diagnostic considerations, common genetic factors between psychiatric disorders, allelic heterogeneity and the ability to detect variants with small effect sizes being limited without large populations (Geschwind and Flint 2015; Psychiatric GWAS Consortium Steering Committee 2009). As psychotic and mood disorders exist on a continuum confounded by diagnotic variability, GWAS looking at SNPs associated with neuropsychiatric disorders across diagnostic boundaries have been undertaken.

Neuropsychiatric disorder associations for *NPAS3* found in GWAS are preliminary, but reported here as supportive evidence of the assocation of *NPAS3* with cross-disorder features of psychiatric disorders. *NPAS3* has been linked to bipolar disorder by GWAS, which was nominally significant in a replication association study of selected bipolar disorder-associated genes (Ferreira et al. 2008; Weber et al. 2011). Further support was offered in a meta-analysis of four GWAS of bipolar disorder found a significant association of *NPAS3* with bipolar disorder (Nurnberger et al. 2014). Furthermore, the same study included analysis of microarrays of post-mortem cortical samples that identified altered expression of *NPAS3* in the dorsolateral prefrontal cortex, a region associated with psychiatric illness including bipolar disorder and schizophrenia, in bipolar disorder patients relative to normal controls. In an independent meta-analysis, twelve SNPs within intron 1 of *NPAS3* were found to be associated with bipolar disorder, schizophrenia and major depression, albeit not at a genome-wide significant level (Huang et al. 2010). A SNP within the second intron of *NPAS3* was found to be part of a set of SNPs discriminative between schizophrenia patients and controls, and furthermore, when compared with MRI-based diffusion tensor imaging, was found to be predictive for

reduced white matter connectivity in schizophrenia patients relative to controls (Gupta et al. 2015). These data suggest that the linkage of SNP variation linked to *NPAS3* with neuropsychiatric disorders identified in GWAS is supported by correlations with variation at the mRNA expression and neurocircuit level.

In a GWAS of patients undergoing a trial treatment with a novel antipsychotic, iloperidone, a serotonin 5HT2A receptor and D2 dopamine receptor antagonist, *NPAS3* genotype was associated with optimal outcomes. Individuals with the minor allele of an intronic SNP in *NPAS3* had significantly higher improvement of symptoms on all scales assessed (Lavedan et al. 2009). Therefore, *NPAS3* genotype appears to be informative for treatment selection in neuropsychiatric disease.

Meta-analysis of 52 GWAS of schizophrenia did not replicate a link between the region proximal to *NPAS3* and schizophrenia, with the closest linked region mapping over 2Mb away (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). SNPs associated with schizophrenia were enriched for genes affected by rare variants, such as CNVs, e.g. small deletions and duplications, supporting the role for rare variants with larger effect sizes to non-genome wide significant genes in the etiology of schizophrenia in individuals that carry them (Schizophrenia Working Group of the Psychiatric Genomics Consortium 2014). Taken together, these data are supportive for a contribution of *NPAS3* variation in the etiology of neuropsychiatric disorders. The variation to *NPAS3* can also be seen to be relevant to treatment and patient outcomes. However, the means by which variants affecting *NPAS3* contribute to normal and abnormal function is not always easy to interpret in the scope of individual psychiatric disorders. When combined with assessment of rare variants of *NPAS3,* which are likely

to have larger effect sizes and cause the neurodevelopmental and psychiatric disorders in the individuals who carry them, these data support that variation affecting *NPAS3* function does contribute to the relative neuropsychological functioning of the individuals, notably affecting intellectual function. Further characterization of how these variants contribute to the observed phenotypes is required, including assessment of contribution to endophenotypes of psychiatric disorders, as discussed above. Furthermore, in order to understand how these variants contribute to the observed phenotypes, the cellular and molecular function of *NPAS3* must be assessed.

## 1.4 *NPAS3* in evolution and mosaicism

*NPAS3* (NCBI Gene ID 64067) is located on chromosome 14q13.1 and is 864kb in length (Kamnasaran et al. 2003; O'Leary et al. 2016). The longest isoform (NM_001164749.1, transcript variant 1) is encoded from 12 exons across its 864 kb length. Alternative exons are encoded and currently there are 18 predicted alternate transcripts from up to 18 predicted exons. Although its coding sequence has been under purifying selection over the human evolutionary lineage from chimpanzee, *NPAS3* has been shown to hold the highest number of the human-accelerated regions (HARs) (Kamm et al. 2013a). HARs are non-coding predicted regulatory sequences which are highly conserved, but found to have enhanced rates of variation over the development of the human lineage (Pollard et al. 2006a; Pollard et al. 2006b). Even though most genes (97.5%) do not contain HARs, and of the remaining 2.5%, 84% contain 1-3 HARs, *NPAS3* was found to contain 14 HARs, which was still found to be significantly enriched when corrected for gene size (Kamm et al. 2013a). Most (11) of these sequences were

shown to act as enhancers directing expression of *NPAS3* to specific locations in the developing zebrafish (*Danio rerio*) central nervous system and all human-specific variants were predicted to affect transcription factor binding sites, hinting at its importance in proper development and function of the human brain (Kamm et al. 2013a). Of note, one of the HARs (HAR-202), which was not found to drive expression in zebrafish, when mutated from human sequence to the chimpanzee sequence, was found to drive specific expression with some neuronal/CNS expression. Contralaterally, another HAR (HAR-142) was found to drive expanded expression into the developing cortex, with the human sequence relative to the mouse or chimpanzee variants (Kamm et al. 2013a; Kamm et al. 2013b). These data suggest that modification of *NPAS3* expression may be involved in the human specific evolution of neurodevelopmental processes and thus contribute to neuropsychological function. Assessments of SNPs affecting predicted brain-specific regulatory elements in the HARs within *NPAS3* within a schizophrenia patient population did not find any significant association with schizophrenia, although variation was observed in patient and control populations (Gonzalez-Penas et al. 2015).

Finally, *NPAS3* has been shown to be a hot spot for double-stranded DNA breaks in neural stem cells, a phenomenon that may contribute to the observed CNV mosaicism among neurons in post-mortem samples from human brains, as well as in neurons differentiated in culture (McConnell et al. 2013; Wei et al. 2016). These CNVs and double-stranded breaks can be seen to relate to variation in neurological function, both in normal individuals, as well as contributing to the pathology of neuropsychiatric disorders, which would not be detected by GWAS and sequencing of DNA derived from non-neuronal cell types. In summary, variation affecting expression of *NPAS3* can be seen to

contributes to variation in neurological function. However, little is known of the normal molecular function of *NPAS3*, which is discussed below.

## 1.5 *NPAS3* variants of uncertain significance

Multiple variants predicted to affect *NPAS3* expression and function have been identified in cross-sectional populations, as well as neuropsychological and psychiatric disorder populations. Figure 1-1 depicts the exonic structure of *NPAS3* and the *NPAS3* variants of interest for further functional analysis. Single nucleotide variants of *NPAS3* are numbered based on transcript variant 1 of *NPAS3*, NM_001164749.1, where +1 is the first nucleotide of the transcript and the encoded protein is isoform 1 NP_001158221.1, 933 amino acids (aa).

The variants that have been studied in this thesis project are as follows: the common coding variant of exon 12 (rs12434716, c.1654G>C, p.Ala552Pro), which has been identified as enriched in a schizophrenia patient population, alongside two proximal non-coding SNPs (rs10141940, c.2208C>T and rs10142034, c.2262C>G), which were found to be inherited as a haplotype (Macintyre et al. 2010). The first variant was predicted by bioinformatic analysis to be benign due to its relative non-conservation. Of note, its ancestral state is a proline residue, rather than an alanine. The c.1654G>C variant was found at the same frequency in a cohort of over 900 autism spectrum disorder cases and neurologically normal controls (Stanco et al. 2014). As the c.1654G>C variant has been inconsistently associated with neuropsychiatric disorders and is found in the normal population, characterization of its functional consequences are of interest for assessment

of its pathogenicity and/or potential role in normal variation in neuropsychological functioning.

In the same study that identified c.1654G>C as associated with schizophrenia, another coding variant (rs141427321, c.2089G>A, p.Gly697Ser) was identified at a low frequency in the control sample and has since been identified at a frequency of 0.0003-0.002 in various populations (Lek et al. 2016; National Center for Biotechnology Information, National Library of Medicine 2017; Sherry et al. 2001). This variant is localized to a poly-glycine repeat, which is in the predicted transactivation domain of NPAS3 and is of interest due to its potential effects on gene regulatory function of NPAS3.

A separate family-based study identified a coding variant of *NPAS3* (rs146677388, c.910G>A, encoding p.Val304Ile) as co-segregating with schizophrenia. This variant is localized between the PAS domains of NPAS3, domains which are known to be involved in interaction between bHLH-PAS proteins and may affect the ability of NPAS3 to interact with other bHLH-PAS proteins (Kikuchi et al. 2003; Lindebro, Poellinger, Whitelaw 1995; Wu et al. 2013a; Wu et al. 2015). These variants can all be thought to affect NPAS3 function in some manner, including predicted deficits in splicing for the non-coding variants and affecting relative protein function. As such, functional assessment will expand our understanding of the significance of these variants in the individuals that carry them, both in normal and abnormal functioning.

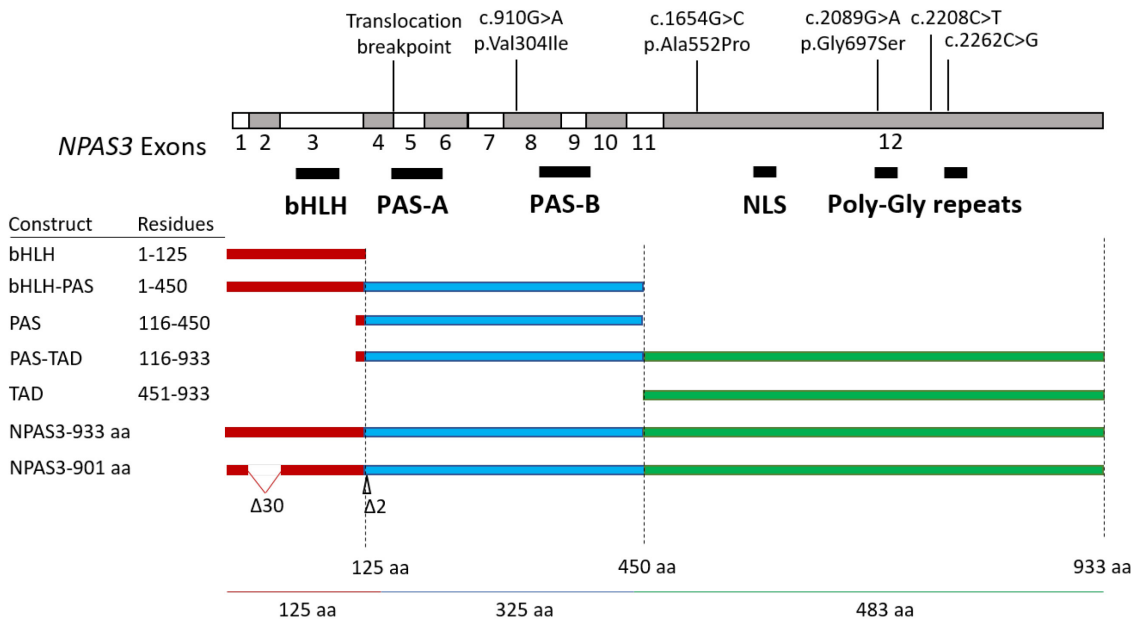



**Figure 1-1: Ideogram of *NPAS3* exonic and domain structure with variants and constructs assessed in this thesis.**
Structure of the NPAS3 protein is based on NP_001158221.1, 933 aa isoform 1 encoded by NM_001164749.1 transcript variant 1. Protein domains and motifs were predicted by ScanProsite tool and InterPro (de Castro et al. 2006; Finn et al. 2017). Relative deletions of the 901 aa isoform 2 (NM_022123.1, NP_071406.1) flanking, but not including, the bHLH DNA-binding domain are indicated by triangles. Constructs used for domain analysis are indicated with colour coding indicating the domain(s) included and the size of each domain. Exons are numbered based on transcript variant 1 as annotated in human genome build hg19.

## 1.6 NPAS3 in neurodevelopment

The nervous system is largely derived from neuroectodermal cells of the neural plate. This plate invaginates into neural folds which circularize to create the neural tube (Mueller, Hassel, Grealy 2015 p.455). The neural tube further forms bulbous structures called vesicles, which delineate the anatomical division of the brain into fore- (prosencephalon), mid- (mesencephalon) and hindbrain (rhombencephalon) (Mueller, Hassel, Grealy 2015 p.455-7). The prosencephalon is further subdivided into the

telencephalon, the developmental structure that will form the neocortex, the basal ganglia including the limbic system and hippocampus, and the diencephalon which forms the thalamus, hypothalamus and the eye vesicles (Mueller, Hassel, Grealy 2015 p.457-8). The midbrain is formed from the mesencephalon. The rhombencephalon subdivides into the metencephalon and myelencephalon, which form the hindbrain structures: the pons, cerebellum and medulla (Mueller, Hassel, Grealy 2015 p.457-9). Neurons throughout the developing brain are generated from proliferative neural stem cells which are produced by symmetric division to expand the neuroprogenitor population, followed by asymmetric division and differentiation into various cell types of the nervous system, including neurons, glia and astrocytes (Brown et al. 2011; Hansen et al. 2010; Hansen et al. 2013; Kornack and Rakic 1995; Noctor et al. 2001). Populations of neuronal stem cells are maintained in discrete regions of the central nervous system that remain proliferative into adulthood in mammals (Eriksson et al. 1998; Spalding et al. 2013).

*Npas3* is expressed throughout the developing mouse (*Mus musculus*) central nervous system, starting caudally in the neural tube as early as E9.5 and proceeding rostrally to include the rhombencephalon, mesencephalon and telencephalon with apparently maximal expression at E13.5 (Brunskill et al. 1999). Expression of *Npas3* is restricted to specific cells in the developing CNS, including cells localized to the ventricular zone, the mantle zone of the subpallial ganglionic eminences, the basal ganglia and regions of the developing midbrain and hindbrain (Stanco et al. 2014). *In situ* hybridization of developing mouse embryos demonstrated *Npas3* expression at E10.5 mostly localizing to neural precursors expressing *Sox2*, but not differentiating neurons expressing *Dcx*, *Tuj1,* or *β(III)-tubulin* (Kamm et al. 2013b). In adult mice, *Npas3*

expression was observed in cells throughout the cortex, in the basal ganglia, midbrain and granular layer of the cerebellum, as well as in the hippocampus, with strongest staining in the subgranular cell layer of the dentate gyrus (Brunskill et al. 2005; Erbel-Sieler et al. 2004; Sha et al. 2012; Stanco et al. 2014). This staining is consistent with expression of *Npas3* in proliferating neuroprogenitors, with more restricted *Npas3* expression in differentiating and differentiated cells. In zebrafish, a similar trend was observed: at 24 hours post-fertilization, *npas3* mRNA expression was found throughout the developing fish and its expression became specific to the developing CNS, eye, heart, limbs, branchial arches and mouth (Kamm et al. 2013a). In chickens, *Npas3* expression has been shown in the ventricular zone of the developing CNS, in a manner coordinate with those observed in other species (Shin et al. 2010; Shin and Kim 2013). These data demonstrate conservation of *NPAS3* expression among vertebrates.

In order to further characterize the role of *NPAS3* in neurodevelopment, *Npas3* knockout mice were developed (Brunskill et al. 2005; Erbel-Sieler et al. 2004). In mice where *Npas3* was deleted by replacement of the exon encoding the bHLH DNA-binding domain with a neomycin resistance cassette, homozygous knockout mice were found to be smaller than wild-type. The volume of the cortex, basal ganglia and dentate gyrus of the hippocampus of these mice were all found to be reduced, accompanied by enlarged ventricles, reduced commissural fibres crossing the corpus callosum and altered cerebellar foliation (Brunskill et al. 2005; Stanco et al. 2014). These data suggest deficits in neuroproliferation and neuroconnectivity as caused by *Npas3* deletion.

Immunohistochemical studies of developing human brains using an NPAS3 antibody have been performed, and the observed staining largely is concordant with

expression data from developing mice (Brunskill et al. 1999; Erbel-Sieler et al. 2004; Gould and Kamnasaran 2011). Expression was observed in the ventricular zone of the cortex, as well as in the dentate gyrus of the hippocampus and in the developing cerebellum. NPAS3 positive cells were also distributed throughout the cortex in a pattern reminiscent of those observed in mice (Gould and Kamnasaran 2011). Taken together, these data support a role of *NPAS3* in the development of the mammalian CNS and its linkage to the etiology of neuropsychiatric disorders.

*Npas3* has been shown to be involved in the generation of interneurons during mouse neurodevelopment (Stanco et al. 2014). Here, I refer only to GABAergic interneurons in the discussion of interneurons. In adult mice, *Npas3* expression was observed throughout the brain: distributed throughout the cortex, as well as in the dentate gyrus of the hippocampus, in the basal ganglia, midbrain, hindbrain and the granular layer of the cerebellum. Closer assessment of cortical *Npas3* staining demonstrated varying overlap with markers of interneurons (GABA and GAD67), suggesting its expression was primarily in inhibitory interneurons (Erbel-Sieler et al. 2004; Stanco et al. 2014). Preliminary evidence gained from *Npas3* knockout mice found that even with loss of *Npas3*, interneurons were still produced normally, demonstrated by GAD67 expression. However, another cortical interneuron marker that marks a subset of interneurons, *Reln,* was found to be reduced (Erbel-Sieler et al. 2004; Pesold et al. 1998). These data suggest that *Npas3* may be involved in the development of a subpopulation of cortical interneurons during embryogenesis.

Interneurons are a class of neuron that are responsible for the inhibitory signalling required for maintenance of excitation/inhibitory balance in the cortex (Rudy et al. 2011).

The balance between excitation and inhibition has been associated with multiple neuropsychiatric disorders, such as schizophrenia, with which *NPAS3* has been associated (Marin 2012; Nakazawa et al. 2012). During embryonic neurodevelopment, interneurons are born in subcortical progenitor domains, primarily in the medial and caudal ganglionic eminences (MGE and CGE, respectively), and the majority migrate tangentially into the cortex, while some migrate ventrally into the striatum and pallidum, where, in both cases, they integrate into local circuits and act as gating factors that inhibit excitation of target neurons (Figure 1-2) (Anderson et al. 1997; Hansen et al. 2013; Marin, Anderson, Rubenstein 2000; Miyoshi et al. 2007; Miyoshi et al. 2010). Newly born neuroprogenitors from the MGE and CGE that are destined to terminally differentiate into cortical interneurons express early markers of interneurons, including homeobox transcription factors of the Dlx family, including *Dlx1*, *Dlx2*, *Dlx5* and *Dlx6,* as well as other lineage specific markers (Batista-Brito et al. 2008; Eisenstat et al. 1999; Hansen et al. 2013; Kanatani et al. 2008; Long et al. 2009). These markers allow for tracking of interneurons born from subpallial structures throughout their migration and terminal differentiation into mature interneurons.

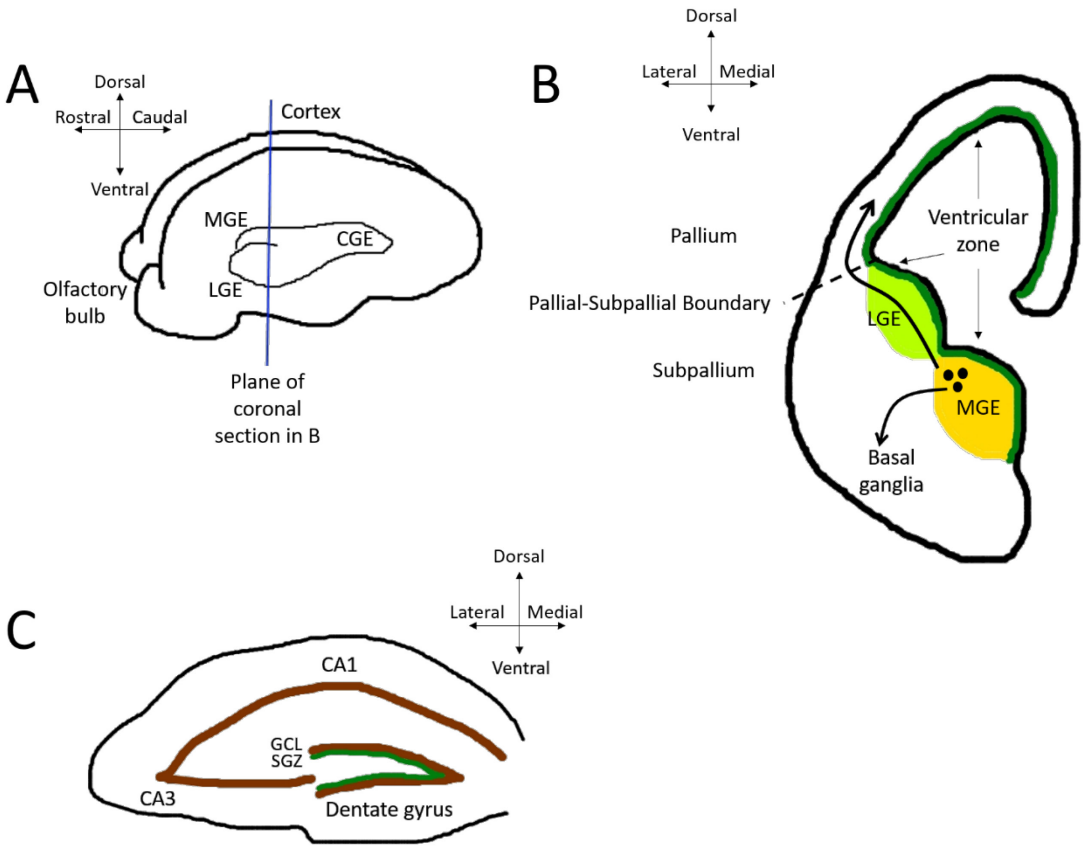



**Figure 1-2: Neuroanatomical structures where *Npas3* is expressed during neurodevelopment and neurogenesis.**
(A) Side view of a the developing mouse brain indicating the plane of the coronal hemispheric section in (B). MGE: medial ganglionic eminence. LGE: lateral ganglionic eminence. CGE: caudal ganglionic eminence. (B) Ideogram of the E13.5 mouse right cortical hemisphere. Dark green: ventricular zone, where *Npas3* expression is observed, light green: LGE, yellow: MGE. Curved arrows indicate migratory paths of interneurons into the pallium (neocortex) and basal ganglia of the subpallium. Black circles: interneurons born in the MGE. (C) Ideogram of the coronal section of the right hippocampal formation of the adult mouse. Brown areas indicate the granular cell layer (GCL) of the dentate gyrus and the pyramidal cell layer of CA1-CA3. Dark green areas indicate regions of *Npas3* expression, concordant with the subgranular zone (SGZ). Drawn by L.M.Luoma.

Lineage tracing involves targetted expression of fluorescent markers driven by lineage specific regulatory sequences, such as enhancer elements that drive *Dlx* gene expression in interneurons, to follow cells through migration and differentiation (Batista-Brito et al. 2008; Potter et al. 2009; Solek et al. 2017). Using lineage tracing, interneurons born in the MGE and CGE have been tracked as they migrate into the pallium (developing neocortex), as well as the basal ganglia and striatum in the subpallium (Miyoshi et al. 2007; Miyoshi et al. 2010). In support of its role in the generation of the interneuron population during neurogenesis, *Npas3* expression was found to be enriched in a population of *Dlx5/6* expressing cortical interneuron progenitor cells isolated from the mouse pallium on E13.5 and E15.5, when interneuron-fated neuroprogenitor cells are migrating tangentially into the cortex (Batista-Brito et al. 2008). Further, *Npas3* mRNA expression in the developing mouse CNS shows concordance in staining with markers of interneurons (Erbel-Sieler et al. 2004; Stanco et al. 2014). Expression of *Npas3* has been detected in the ventricular zone of both the pallium and subpallium, including the ventricular zone of the MGE and CGE at E13.5 (Figure 1-2), as well as strong staining in the basal ganglia, including the globus pallidus starting at E15.5 (Stanco et al. 2014). Postnatally, *Npas3* expression was observed in interneurons throughout the mouse cortex at P30, including interneurons expressing markers of multiple interneuron subtypes that represent the majority of interneurons in the neocotex, somatostatin (SST), parvalbumin (PV), Reelin (RELN), calretinin (CR) and neuropeptide Y (NPY) (Rudy et al. 2011; Stanco et al. 2014). Although *Npas3* expression was identified in the majority of interneurons of a variety of subtypes, *Npas3* was not found to be exclusively expressed in interneurons, with 67% of *Npas3* expressing cells at P5 being GAD67 positive but only

21% remain GAD67+ at P30 (Stanco et al. 2014). Further, studies of selected cortically migrating interneuron populations based on *Dlx5/6* expression showed 3- to 5-fold enrichment of *NPAS3* relative to a non-selected population of cortical neuroprogenitors from the same timepoint, rather than exclusive expression, suggesting that *Npas3* plays a role in the generation of multiple types of neurons, including interneurons (Batista-Brito et al. 2008). This finding is consistent with the observed expression of *Npas3* in the developing pallial ventricular zone, which produces cortical excitatory neurons, among other cell types, but not interneurons (Anderson et al. 2002; Gorski et al. 2002; Kon, Cossard, Jossin 2017; Stanco et al. 2014).

Deletion of *Npas3* has been shown to reduce the cortical interneuron population, as well as cortical size. Assessment of interneuron number and cortical volume in mice at P30 demonstrated that deletion of *Npas3* results in a decrease in cortical size by 21%, decreased proliferation in the MGE but not the CGE at E13.5, and a coordinate decrease in SST+, VIP+, NPY+ and RELN+, but not PV+, interneuron numbers throughout the cortex (Stanco et al. 2014). The subtypes of interneurons lost have been shown to be derived from both the MGE (PV+, SST+) and CGE (RELN+, CR+, VIP+), suggesting that Npas3 function is important in both of these embryonic sources of interneurons, despite the lack of observed neuroproliferative deficit in the CGE of *Npas3-/-* mice at the timepoints assessed (Miyoshi et al. 2007; Miyoshi et al. 2010; Rudy et al. 2011; Xu et al. 2004). Assessment of ERK phosphorylation, used as a marker for responsiveness to fibroblast growth factor (FGF) signalling, demonstrated that at both the MGE and CGE, loss of *Npas3* resulted in decreased ERK phosphorylation relative to basal rates (Stanco et al. 2014). This may be suggestive of a deficit in normal function, despite the detection

of reduced proliferation only in the MGE. Although rates of apoptosis were not assessed during development, the observed loss of interneurons postnatally, and reduction in cortical volume, was not correlated with an increase in rates of apoptosis (Stanco et al. 2014). This finding suggests the loss of interneurons occurs during development either by reduced proliferation or increased apoptosis. These data directly demonstrate a role for *NPAS3* in the generation of cortical interneurons, as well as demonstrating that NPAS3 is involved in the proliferation of neuroprogenitor cells during development, contributing to the generation of excitatory neuronal subtypes in cortical, subcortical forebrain structures as well as mid- and hindbrain structures during embryogenesis.

## 1.7 Behavioural changes associated with *Npas3* deletion in mice

*Npas3* knockout mice have been generated by deletion of exon 3, encoding the DNA-binding domain, which is predicted to result in a null allele. Homozygous knockout mice displayed altered reflexes to tail suspension, as well as altered nesting behaviour and poor mothering, resulting in pup mortality (Brunskill et al. 2005; Erbel-Sieler et al. 2004). In behavioural testing paradigms, *Npas3* knockout mice have been found to have reduced prepulse inhibition of acoustic startle response, despite normal baseline startle responses. These data suggest altered sensorimotor gating and emotional tone in mice deficient in *Npas3* function (Brunskill et al. 2005; Erbel-Sieler et al. 2004).

Due to the observed reduction in hippocampal volume in *Npas3* knockout mice, recognition memory, for which the hippocampus plays a role in encoding, was assessed (Brunskill et al. 2005; Kim et al. 2014). When presented with a novel object and a known object, mice are expected to explore the novel object more. However, in *Npas3* knockout

mice, recognition memory appeared to be reduced, resulting in decreased time spent with the novel object relative to the non-novel object (Brunskill et al. 2005). Similarly, a deficit in social recognition has been observed in *Npas3* knockout mice, which may be contributed to by decreased social interaction with novel animals upon introduction, or due to decreased social recognition memory (Erbel-Sieler et al. 2004). These data are supportive of deficits in hippocampal-dependent learning and memory processes caused by the loss of *Npas3* expression.

*Npas3* knockout mice were observed to have altered emotional state. Enhanced open-field activity and increased time spent in open sections of an elevated zero maze was observed, suggesting altered anxiety (Brunskill et al. 2005). The mice were also found to be hyperactive in certain environments, however hyperactivity was not consistently observed by different testing methods; notably hyperactivity was associated with lack of habituation to the environment, which may be associated with memory impairment (Brunskill et al. 2005; Erbel-Sieler et al. 2004). Finally, the mice were observed to have gait defects and difficulty crossing narrow beams, which may be associated with the observed reduction in cerebellar foliation in *Npas3*$^{-/-}$ mice (Brunskill et al. 2005). The observed deficits in learning and memory associated with altered behavioral outcome in *Npas3* knockout mice are understandable in the context of NPAS3 function being involved in cognition and neuropsychiatric illness in humans.

## 1.8 NPAS3 in adult neurogenesis

The identification of neurogenesis in the dentate gyrus of the adult human brain revolutionized understanding of how the brain works, as the brain was previously widely

thought to be a terminally differentiated organ (Eriksson et al. 1998). Since its discovery, adult neurogenesis has been found to occur at rates relevant to human behaviour, and to persist throughout the lifespan, up to 80 years of age (Knoth et al. 2010; Spalding et al. 2013). Hippocampal-dependent learning and memory have been associated with adult neurogenesis (Gould et al. 1999; Leuner et al. 2004). Reduction in rates of adult neurogenesis in rodent models has been shown to be associated with reduced learning and altered cognitive flexibility (Dupret et al. 2008; Winocur et al. 2006). These data suggest that reductions in the rate of adult neurogenesis causes reduced learning and memory, resulting in behavioural outcomes like those observed in *Npas3* knockout mice: without adequate rates of neurogenesis, they cannot encode, or store new memories as efficiently as wild-type mice. Furthermore, adult neurogenesis has been shown to be reduced in neuropsychiatric illness and responsive to treatment with pharmacological agents (Kodama, Fujioka, Duman 2004; Reif et al. 2006; Walker et al. 2015). *Npas3* was found to be expressed in neuronal cells throughout the adult mouse brain including the dentate gyrus of the hippocampus (Erbel-Sieler et al. 2004; Stanco et al. 2014). Given the observed deficits in learning and memory in mice deficient in *Npas3,* combined with the observed expression in the subgranular layer of the dentate gyrus of the hippocampus, a known neurogenic zone of the adult mammalian brain, the role of Npas3 in adult neurogenesis has been examined (Goncalves, Schafer, Gage 2016; Knoth et al. 2010).

Loss of *Npas3* results in almost complete loss of neurogenesis in the dentate gyrus of the hippocampus, and a corresponding reduction in the thickness of the granular layer of the dentate gyrus (Pieper et al. 2005). The deficit in neurogenesis of *Npas3*$^{-/-}$ mice was found not to be in the basal rate of neuroproliferation, but in decreased survival of the

newly born neurons, likely due to increased rates of apoptosis, given the observed increase in cleaved caspase 3, a marker of apoptosis (Pieper et al. 2010). This deficit was found along with a decrease in *Fgfr1* expression and unresponsive to treatment with bFGF (Pieper et al. 2005). However, the deficit in neurogenesis was partially responsive to electro-convulsive seizure, known to increase rates of neurogenesis, suggesting that Npas3 activity is important for neurogenesis in the dentate gyrus of adult animals and especially with FGF-mediated adult neurogenesis (Pieper et al. 2005). Further analysis of the neuroconnectivity of the hippocampus in *Npas3* knockout mice showed that downstream of the initial deficit in survival rates of newly born neurons there is significantly reduced dendritic branching and neuronal connectivity, as measured by the number of synaptic spines and staining for synaptic vesicles (Pieper et al. 2010). This likely contributes to the observed hyperexcitability of synaptic transmission in the hippocampus. Treatment with an anti-apoptotic, pro-neurogenic compound (P7C3) was found to rescue the loss of neurogenesis in *Npas3*$^{-/-}$ mice and was found to return the thickness of the granular layer of the dentate gyrus to wild-type levels (Pieper et al. 2010). These data indicate that Npas3 function is critical for the survival of newly born neuroprogenitor cells in which it is expressed and in the dentate gyrus, this may contribute to secondary effects on surviving cells that have decreased neuronal connectivity, possibly due to the loss of *Npas3* expression within the surviving cells, or due to the loss of proximal *Npas3* expressing cells.

Single-cell RNA sequencing of clonal populations of neural stem cells isolated from the subgranular zone of the dentate gyrus identified *Npas3* as expressed in quiescent and activated neural stem cells and was down-regulated in intermediate progenitor cells

although its expression is retained in a subset of these cells (Shin et al. 2015). The down-regulation of *Npas3* was correlated with decreased expression of multiple markers of neuronal stem cells, including *Sox2* and members of the Notch signalling pathway known to be important for maintenance of neural stem cells, including *Hes1* and *Notch2* (Axell, Zlateva, Curtis 2009; Kageyama et al. 2005; Shin et al. 2015; Wilhelmsson et al. 2012). In adult mouse hippocampi, Npas3 protein has been shown to be expressed largely in cells positive for *Dcx*, which is expressed in late stage neuroprogenitors and immature neurons for approximately 2 weeks (Erbel-Sieler et al. 2004; Goncalves, Schafer, Gage 2016; Sha et al. 2012). Final maturation of hippocampal neurons, where *Npas3* is expressed in a subset of cells, extends to six weeks (Erbel-Sieler et al. 2004; Goncalves, Schafer, Gage 2016). As such, there is a long temporal period where Npas3 function can be seen to affect neurogenesis in the dentate gyrus.

## 1.9 Neuroconnectivity deficits associated with NPAS3 dysfunction

As *Npas3* expression partially co-localizes with interneuronal markers throughout the cortex and its deletion results in an alteration in interneuron population, widespread alterations in neurotransmission and neurocircuit tone can be hypothesized (Erbel-Sieler et al. 2004; Stanco et al. 2014). The hyperactivity that has been observed in *Npas3* knockout mice in an open field test was used to assess neurotransmission in these mice, assessing the response of this hyperactivity to pharmaceuticals affecting a range of neurotransmitters (Brunskill et al. 2005). In this study, hyperactivity in *Npas3* knockout mice was found to be non-responsive to methamphetamine, a dopamine agonist. Reciprocally, muted response was observed with the dopamine D2 receptor antagonist

haloperidol, suggesting perturbations in dopamine receptor signalling. Tests of dopamine D1 receptor antagonists were found to be conflicting, with opposite and muted response by *Npas3* knockout mice relative to wild-type littermates. Altered response to an antagonist of the NMDA receptor, involved in glutamate signalling, was also noted, where the activity of *Npas3* knockout mice was suppressed instead of activated as observed in the wild-type littermates. In both the assessment of dopamine and NMDA receptors, the amount of each targeted receptor was not found to be altered, suggesting that the mechanism of altered signalling response does not lie with the expression of the receptors themselves, but in the response to the signals received via the receptors (Brunskill et al. 2005). These data are reminiscent of the decreased neuroconnectivity observed in the hippocampus of *Npas3* knockout mice, resulting in increased post-synaptic potential, that can be linked to hyperactivity in signalling (Pieper et al. 2010). As dopamine and glutamate signalling have been linked to human neurocognitive function, and shown to be dysregulated in neuropsychiatric dysfunction, alterations in *NPAS3* expression and function may result in significant effects on wide-ranging neuronal circuits (Zai et al. 2017). For example, dopamine signalling has been shown to exist on a functional curve where moderate levels are associated with maximal cognitive function (Callicott et al. 2003; Drapier et al. 2008; Glahn et al. 2005; Seidman et al. 2006).

Assessment of *NPAS3* expression in the dorsolateral prefrontal cortex (DLPFC) of human neonates through adults demonstrated that *NPAS3* is expressed and its translation is regulated by microRNAs, small RNAs involved in regulation of mRNA expression and stability (Vreugdenhil and Berezikov 2010; Wong et al. 2012). *NPAS3* mRNA expression was detected at high levels in neonates, which decreased over the first decade of life

before becoming stable, whereas NPAS3 protein expression was found to increase in the first decade of life and decrease thereafter (Wong et al. 2012). microRNA miR-17 was found to be regulated over the same time-span in a reciprocal manner to the observed protein expression and was found to be able to reduce expression of a reporter fused to the *NPAS3* 3'UTR (untranslated region), which contains a predicted miR-17 target sequence. *In situ* hybridization of a cohort of schizophrenia patient post mortem DLPFC slices found no change in *NPAS3* mRNA expression level corresponding to disorder status. In female patients, NPAS3 protein was found to be reduced, which was correlated with an observed increase in miR-17 in a separate set of DLPFC samples from schizophrenia patients, however sex-specific effects were not observed for miR-17 levels. These data suggest that NPAS3 regulation is associated with neuropsychiatric disorders, and that *NPAS3* expression is regulated at multiple levels, including by miRNAs, which should be taken into account when interpreting expression data. Further, sex specific effects appear to be relevant to the contribution of *NPAS3,* and variation thereof, to neuropsychological function (Wong et al. 2012). The detection of expression of *NPAS3* in the human cortex combined with the observation that its expression is actively regulated, support its role in human neuropsychological function. A reduction in *NPAS3* in cells throughout the prefrontal cortex may indicate altered neurotransmission and neuronal circuit tone, which may contribute to neuropsychiatric dysfunction.

## 1.10 NPAS3 and metabolism

A recent microarray, assessing genes differentially regulated by expression of *NPAS3,* identified enrichment for genes involved in the regulation of energy metabolism.

Notably, differentially regulated genes were found to be involved in glycolysis and oxidative balance, with a concordant alteration in metabolomic markers observed (Sha et al. 2012). The finding that NPAS3 is associated with metabolic balance was supported by a SNP association study of obese individuals compared to never-obese sibling(s) and parent(s) which identified an association between *NPAS3* and phenotypes of metabolic syndrome (Zhang et al. 2013b). Metabolic syndrome is a disorder of metabolism resulting in insulin resistance, type 2 diabetes, glucose intolerance, dyslipidemia, increased plasma cytokines, hypertension, increased visceral fat and increased risk for multiple cancers. Specifically, *NPAS3* SNPs were linked to increased plasma triglycerides and cholesterol, as well as increased plasma cytokines IL-1β (interleukin 1β) and IL-6 (interleukin 6), which are markers of inflammation. These data suggest that NPAS3 function plays a role in energy metabolism (Zhang et al. 2013b).

## 1.11 NPAS3 in cancer

Variants affecting *NPAS3* have been associated with multiple forms of cancer, with different effects (Bao et al. 2016; Bavarva et al. 2013; Krishnan et al. 2016; Moreira et al. 2011). Assessment of DNA methylation in tongue squamous cell carcinoma demonstrated reduced methylation of NPAS3 over tumour stage progression, suggestive of a role of *NPAS3* expression in tumour progression. However, expression of *NPAS3* was not directly assessed to confirm the direction of the regulation by hypomethylation (Krishnan et al. 2016). In prostate cancer, an intronic SNP was found to be associated with decreased recurrence rates in patients treated by radical prostatectomy (Bao et al. 2016). This SNP was localized in an ultra-conserved non-coding region predicted to be a

regulatory sequence (rs8004379). Bioinformatic analysis of previously generated data found that this variant was correlated with higher expression levels of *NPAS3* and improved survival in two independent study populations assessing prostate cancer survival and genotype (Bao et al. 2016; Gulzar, McKenney, Brooks 2013; Taylor et al. 2010).

*NPAS3* expression has been shown to be reduced in cancer, indicating that NPAS3 has tumor suppressor-like properties (Bao et al. 2016; Moreira et al. 2011). A study looking at nicotine stress effects on gene regulation identified *NPAS3* as down-regulated by nicotine treatment (Bavarva et al. 2013). As nicotine is associated with increased growth of solid tumors and carcinogenesis, the demonstrated nicotine induced down-regulation of *NPAS3* is in agreement with its role as a tumor-suppressor. In astrocytomas, loss of *NPAS3* expression through deletion, epigenetic silencing and somatic point mutations was found to be associated with increased grade and proliferation. Knockdown of *NPAS3* was found to recapitulate the increased proliferation and aggressiveness, while reintroduction of NPAS3 was found to reduce the proliferation of immortalized fetal astrocytes, and alter the expression of genes involved in cell cycle progression, apoptosis and migration/invasion (Moreira et al. 2011). These data are somewhat in conflict with the role of NPAS3 in survival of newborn neurons, whereby loss of NPAS3 is pro-apoptotic, whereas in tumors loss of NPAS3 appears to be oncogenic. Understanding the molecular function of NPAS3 will enhance our understanding of how NPAS3 regulation is reminiscent of a tumor suppressor during oncogenesis while its expression in newly born neuroprogenitors is anti-apoptotic, which would be expected to be oncogenic in the context of a tumour.

## 1.12 NPAS3 in respiratory development

*NPAS3* is highly homologous to *trachealess*, a bHLH-PAS transcription factor involved in tubulogenesis and FGFR1 (*breathless*) mediated branching morphogenesis in the *Drosophila melanogaster* tracheal system (Isaac and Andrew 1996; Klambt, Glazer, Shilo 1992; Reichman-Fried et al. 1994; Wilk, Weizman, Shilo 1996). Trachealess directly regulates the expression of *breathless*, and its loss results in complete loss of tracheal development by inhibiting tubulogenesis (Isaac and Andrew 1996; Ohshiro and Saigo 1997; Wilk, Weizman, Shilo 1996). Inhibition of *trachealess* expression later in tracheal development, after the tracheal tubes have formed, inhibits the tracheal migration and branching through *breathless*-mediated FGF signalling (Ohshiro and Saigo 1997; Wilk, Weizman, Shilo 1996). Similar roles for trachealess have been observed in migration of cells in the salivary glands indicating a conserved role of *trachealess* in branching morphogenesis (Isaac and Andrew 1996; Klambt, Glazer, Shilo 1992; Kuo et al. 1996).

In humans, genome wide-association and linkage studies have identified *NPAS3,* or the region proximal, as associated with asthma (Blumenthal et al. 2004; Collaborative Study on the Genetics of Asthma (CSGA) 1997). *Npas3* was found to be expressed in the developing mouse lung, with highest expression early in development (E10.5), decreasing expression until birth and low-level maintenance during the first ten months of life. *Npas3* was found to be expressed in epithelial and mesenchymal cells of the developing lung (Zhou et al. 2009).

An *Npas3* knockout mouse was generated by inserting a premature stop codon in exon 5, resulting in a transcript encoding only the bHLH domain and a portion of the

PAS(A) domain (compared to the bHLH-domain/exon 3 knockout characterized for behavioral and neuroanatomical deficits) (Zhou et al. 2009). This *Npas3*$^{exon4-/-}$ mouse model demonstrated decreased lung branching morphogenesis, decreased vascularization of the lung and lethality within 48 hours of birth. Comparatively, the exon 3 deletion *Npas3*$^{-/-}$ mice are not known to have altered lifespan (Erbel-Sieler et al. 2004; Pieper et al. 2005). In the lungs of *Npas3*$^{exon4-/-}$ mice expression of members of the *Shh* (*sonic hedgehog*) and *FGF* signalling pathways were found to be altered. In a study of rat lung development using nitrofen to disrupt a later stage of lung development resulting in diaphragmatic hernia, Npas3 expression was found to be reduced, a further association of Npas3 function in lung development (Pereira-Terra et al. 2015). The identification of altered branching in the lungs of *Npas3*$^{exon4-/-}$ mice and the concomitant dysregulation of the Fgf/Fgfr1 signalling pathway suggest that *Npas3* function shows conservation with *trachealess*, its closest orthologue, including a role in respiratory development and branching morphogenesis.

## 1.13 NPAS1: a paralogue of NPAS3 involved in interneuron development

*NPAS1*, a bHLH-PAS transcription factor paralogue of *NPAS3*, with high conservation in the bHLH and PAS domains but not the C-terminus, is often characterized in parallel with NPAS3 (Brunskill et al. 1999; Pieper et al. 2005; Zhou et al. 1997). *Npas1* expression is highly enriched in the CNS and has also been shown to be associated with interneuron development and specification (Stanco et al. 2014; Zhou et al. 1997). Variation of *NPAS1* has been associated with autism and deletions including *NPAS1* are linked to a microdeletion syndrome featuring developmental delay (Stanco et

al. 2014; Travan et al. 2017). *Npas1* expression has been found to be much more tightly associated with interneurons with almost no expression observed outside of interneurons in the cortex (Stanco et al. 2014). During development, *Npas1* is highly expressed in the subventricular zone of the MGE and CGE, and cells expressing *Npas1*, along with other markers of interneurons, have been shown to migrate into the cortex, where *Npas1* expression increases briefly postnatally as interneurons are radially migrating and maturing (Cobos et al. 2006; Stanco et al. 2014). Its expression is retained in the adult mouse cortex in a subset of interneurons, and *Reln* expression is altered in *Npas1-/-* mouse cortices (Cobos et al. 2006; Erbel-Sieler et al. 2004). *Npas1* expression was also found to be enriched in the hippocampus in a similar pattern to *Npas3*, however no effects on neurogenesis or responsivity to bFGF were observed in *Npas1* knockout animals (Erbel-Sieler et al. 2004; Pieper et al. 2005; Stanco et al. 2014). Knockout mice for *Npas1* have been assessed for behavioural deficits in parallel with *Npas3* knockout and compound *Npas1/3* knockout mice. Despite the similar expression patterns of *Npas1* and *Npas3* in the adult mouse brain, no significant behavioural deficits were observed in *Npas1*$^{-/-}$ animals without concomitant deletion of *Npas3* (Erbel-Sieler et al. 2004). Loss of *Npas3* was associated with a stereotypical darting behaviour, with all mice demonstrating this behaviour being homozygous for deletion of *Npas3* and either heterozygous or homozygous for loss of *Npas1* (Erbel-Sieler et al. 2004).

Closer assessment of the developmental contribution of *Npas1* determined that deletion of *Npas1* results in the opposite effect of *Npas3,* whereby interneuron number, cortical volume and basal ganglia volume are increased (Stanco et al. 2014). A similar effect was observed in adult neurogenesis in the dentate gyrus where *Npas1*$^{-/-}$ mice had

enhanced neurogenesis (Stanco et al. 2014). However, another study observed no difference in neurogenesis in *Npas1*$^{-/-}$ mice (Pieper et al. 2005). *Npas1* appears to be involved in the generation of cortical and globus pallidus neurons derived from multiple subpopulations of cells in the subpallium (Glajch et al. 2016; Nobrega-Pereira et al. 2010; Stanco et al. 2014). Study of double knockouts of *Npas1* and *Npas3* suggest that *Npas3* function in interneuron neurodevelopment is epistatic to *Npas1*, where the double knockout mice show largely the same interneuronal and behavioural phenotypes of the *Npas3* knockout (Erbel-Sieler et al. 2004; Stanco et al. 2014). NPAS1 has been shown to interact with ARNT (aryl hydrocarbon receptor nuclear translocator) *in vivo,* and to directly regulate target genes, such as repressing expression of *Epo*, *Arx* and *Tyrosine hydroxylase* (*TH*), potentially through binding of hypoxia response elements (Ohsawa et al. 2005; Stanco et al. 2014; Teh et al. 2006). From these data, *Npas1* and *Npas3* can be seen to have distinct functions in neurodevelopment with overlapping mechanistic features, that can be hypothesized to be due to conservation of their functional domains.

## 1.14 bHLH-PAS proteins

*Npas3* was first identified as a highly expressed gene in the mouse developing CNS and was characterized as a bHLH-PAS transcription factor based on homology to the *Drosophila* bHLH-PAS protein *trachealess* and mouse *Npas1* (Brunskill et al. 1999). Analysis of the protein sequence identified a predicted bHLH DNA-binding domain and two degenerate PAS repeats, PAS(A) and PAS(B), followed by a C-terminal domain without any conserved domain structure predicted to be a transactivation domain (Brunskill et al. 1999; Pickard et al. 2006). bHLH-PAS proteins function as homo- or

heterodimers that interact through their bHLH and PAS domains. Only some bHLH-PAS proteins are capable of homodimerization, these have been categorized as Class 2 and are considered generic heterodimeric bHLH-PAS partners, also capable of heterodimerization with tissue- or stimulus-specific bHLH-PAS factors that are categorized as Class 1 (Kewley, Whitelaw, Chapman-Smith 2004). ARNT is considered the general heterodimeric partner of bHLH-PAS proteins and is capable of interacting with bHLH-PAS proteins involved in signalling multiple environmental stimuli, including oxygen state, redox balance and the presence of environmental toxins (Hoffman et al. 1991; Jiang et al. 1996; Rutter et al. 2001). bHLH-PAS proteins contain two common domains: the bHLH domain, which is involved in DNA binding and interaction with other bHLH proteins through the basic and helix-loop-helix regions respectively; the PAS domains, which are degenerate 70 aa repeats over approximately 275 residues encoding PAS(A) and PAS(B) folds, responsible for interaction with other PAS-domain containing proteins; and a C-terminus may be present, commonly containing transactivation or repressive domains which regulate expression of target genes (Ellenberger et al. 1994; Jiang et al. 1996; Kikuchi et al. 2003; Lindebro, Poellinger, Whitelaw 1995; Ma et al. 1994; Moffett, Reece, Pelletier 1997; Pongratz et al. 1998; Sogawa et al. 1995; Whitelaw, Gustafsson, Poellinger 1994; Xiong et al. 2016).

ARNT is a class 2 bHLH-PAS protein considered to be the generic heterodimeric partner of bHLH-PAS proteins (Kewley, Whitelaw, Chapman-Smith 2004). Canonically, when a ligand-bound AhR (aryl hydrocarbon receptor) translocates to the nucleus it interacts with ARNT and alters gene regulation (Hoffman et al. 1991; McGuire et al. 1994; Reyes, Reisz-Porszasz, Hankinson 1992). AhR localized to the nucleus interacts

with ARNT to promote binding to xenobiotic response element sequences, 5'-TNG GTG-3', as compared to ARNT homodimers which bind to the E-box sequence 5'-CAC GTG-3' (Swanson, Chan, Bradfield 1995). By similar mechanisms, ARNT can interact with multiple bHLH-PAS proteins, with and without cofactors, to change its DNA-binding specificity and the regulation of target genes in response to extracellular stimuli (Kikuchi et al. 2003; Kinoshita et al. 2004; Swanson, Chan, Bradfield 1995; Wang et al. 1995; Wharton et al. 1994). Furthermore, this effect is seen among other bHLH-PAS heterodimers, such as the circadian complex NPAS2::BMAL1, where the oxidation state of a bound cofactor (which signals the cellular environment) affects the affinity of the complexes for DNA and thus can affect regulation of target genes (Dioum et al. 2002; Rutter et al. 2001). The PAS domain can also contribute to DNA binding specificity of bHLH-PAS heterodimers (Zelzer, Wappner, Shilo 1997). In the crystal structure of the HIF1α::ARNT and AhR::ARNT heterodimers, regions of the PAS domain were found to make contact with co-crystallized DNA, suggesting that although the PAS domain is not the primary site for DNA sequence recognition and binding, it may contribute to the DNA sequence selection in some cases (Seok et al. 2017; Wu et al. 2015).

The C-terminus of ARNT contains a potent transactivation domain, notably the final 34 aa of the protein and a proximal polyglutamine repeat that has been shown to be critical for activation of target genes (Li, Dong, Whitlock 1994; Sogawa et al. 1995; Whitelaw, Gustafsson, Poellinger 1994). When ARNT is part of a heterodimeric complex, the interacting partner provides conditional regulation of transactivation; for example, differential regulation driven by the hypoxia response element by selection of the activator HIF1A (hypoxia inducible factor 1 alpha) or the repressor NPAS1,

suggesting that the transactivation domain of ARNT is not the sole determinant of the regulatory effect of bHLH-PAS heterodimers (Jiang et al. 1996; Teh et al. 2006; Wu et al. 2013b). Similarly, heterodimers of Arnt and Sim1 or Sim2, two class I bHLH-PAS proteins involved in central nervous system development, result in different transcriptional output, which is associated with differential transactivation domain function between Sim1 (transactivation) and Sim2 (repression) (Moffett and Pelletier 2000; Woods and Whitelaw 2002). Thus, although ARNT functions as an activator when homodimeric, its regulatory activity, as well as its DNA binding specificity, is specified by the tissue- or stimulus-specific heterodimeric partner.

NPAS3 is predicted to be a tissue-specific partner that specifies ARNT binding to divergent E-box-like elements, due to its predicted bHLH and PAS domains. The closest orthologue to NPAS3, the *Drosophila* bHLH-PAS protein trachealess heterodimerizes with tango, the *Drosophila* ARNT orthologue, in order to regulate target genes by CNS midline enhancer elements, E-box like sequence (Morozova et al. 2010; Sonnenfeld et al. 1997). NPAS3 and ARNT have been shown to act cooperatively to regulate genes involved in Shh and FGF signalling, however, the nature of the regulatory interaction has not been characterized (Zhou et al. 2009). The interaction of NPAS3 and ARNT *in vivo* requires experimental validation. Furthermore, direct targets of the NPAS3::ARNT heterodimer, including the specific sequences to which the proposed heterodimer bind, have not yet been identified. These data will fill in gaps in our knowledge of the contribution of NPAS3 to the regulation of target genes in cells expressing *NPAS3* and facilitate understanding of how variation affecting *NPAS3* results in the observed phenotypes described above.

## 1.15 Potential targets regulated by NPAS3

At the start of this study no direct targets of NPAS3 were known. However, regulatory effects of the loss of *NPAS3* as well as overexpression have been assessed in various systems. In the original study identifying *Npas3* as critical for FGF-mediated neurogenesis in the adult mouse hippocampus, co-expression of NPAS3 and ARNT in neuroblastomas followed by qPCR assessing members of the FGF signalling pathway had been performed (Pieper et al. 2005). Similar results were observed in lung development (Zhou et al. 2009). However, whether the observed synergystic effects were direct or resulted in differential regulation of the target genes *in situ* was not pursued.

Transcriptomic analysis of the effects of *NPAS3* expression has been performed using microarrays to assess differentially regulated transcripts with expression of *NPAS3* in HEK 293T cells (Sha et al. 2012). This study identified 282 up-regulated and 359 down-regulated genes. Up-regulated genes were found to be enriched for transcription factors, while down-regulated genes were found to be enriched in metabolic functions, including glycolysis and redox balance. Although this analysis was performed using expressed *NPAS3* in HEK 293 cells and not validated *in vivo*, the metabolic effects observed were confirmed in brains of *NPAS3*$^{-/-}$ mice. Metabolomic analysis of the brains of the *NPAS3*$^{-/-}$ and wild-type littermates demonstrated multiple differences, notably affecting metabolites of the TCA cycle, the urea cycle and glycolysis, supporting the finding of genes involved in energy metabolism as regulated by NPAS3. Finally, the regulatory function of NPAS3 was found to vary over the circadian cycle, where it acted primarily as a repressor at ZT12 (Zeitgerber time +12h), and as both an activator and repressor at ZT24, the opposing ‘end’ of the circadian cycle (Sha et al. 2012). These data

suggest that NPAS3 may interact with circadian bHLH-PAS transcription factors, such as BMAL1, to differentially regulate gene expression.

As these studies assessed differences in mRNA and protein expression in response to NPAS3 and notably as several transcription factors were found to be differentially regulated by expression of *NPAS3* in the microarray study, direct targets of NPAS3 cannot be predicted by these data alone. ChIP data would be required to determine which targets are direct (regulated) and which are secondary effects (genetically regulated).

## 1.16 Project outline

Although NPAS3 function has been robustly linked to altered neuropsychiatric function in a manner consistent with its function, as characterized in the brains of mice, how NPAS3 function affects the observed neuropsychological functioning of both mentally ill and healthy individuals who carry variants in *NPAS3* is not known. To this end, we undertook a pilot study to assess the correlation of *NPAS3* genotype to cognitive function. We assembled a population of western Canadian students before the typical onset age of neuropsychiatric disorders, to assess *NPAS3* genotype relative to neuropsychological function in the population as a whole. These data will enhance our understanding of how the observed reduction in adult neurogenesis, as well as the observed behavioural and neural outcomes caused by loss of *Npas3* in mice may relate to the neurological functioning of individuals carrying variants affecting *NPAS3* which has been linked to neuropsychiatric dysfunction, but is also present in the normal population.

Second, I set out to understand the molecular underpinnings of NPAS3 function. The molecular function of NPAS3 is largely uncharacterized and NPAS3 function is presumed based on bioinformatic data. *NPAS3* and *ARNT* have been shown to affect the regulation of common genes, however, whether the observed gene regulation is due to direct effects has not been shown. Furthermore, whether ARNT and NPAS3 act as a heterodimer had not been experimentally demonstrated. For all genes found to be differentially regulated by NPAS3 (in the presence or absence of ARNT), none of the effects have been demonstrated to be direct. In order to assess the molecular function of NPAS3, I performed assays to determine whether NPAS3 and ARNT do interact, whether they can regulate common targets and whether this is direct. Further, I experimentally determined the function of the predicted domains of NPAS3. Three coding variants of *NPAS3* identified in humans were also assessed in relevant functional assays developed for the characterization of wild-type NPAS3, in order to assess their functional significance to the individuals that carry them (Figure 1-1).

Finally, I set out to connect the molecular data to the cellular effects observed by others. The expression and regulation of *NPAS3* were assessed and multiple factors known to affect regulation by bHLH-PAS genes were screened to determine whether they affect regulation driven by NPAS3. To that end, inducible *NPAS3* cell lines and shRNA knockdown cell lines were generated and characterized. Human neuroprogentior cells were also assessed for expression of *NPAS3,* to develop a model to assess potential NPAS3 functions in human neural stem cells.

Taken together, the data presented in this thesis will fill gaps in the existing knowledge presented above: characterizing the molecular function of NPAS3 and

attempting to connect it to cellular responses in *Npas3* deficient mice, as well as connecting the effect of coding SNPs likely to affect NPAS3 function, to endophenotypes of disorders associated with NPAS3 genotype and to their molecular consequences. These data will enhance the our understanding of the function of NPAS3 in the context of the cell and the individual, as well as the contribution of variants affecting NPAS3 to the normal and abnormal variation in human neuropsychiatry.

# Chapter 2: Materials and Methods

## 2.1 Identification of neuropsychological endophenotypes associated with *NPAS3* variation

### 2.1.1. Consent and cohort acquisition

The study population of western Canadian high school students (*n*=78) was recruited from a related study into the attitudes towards mental health issues, which involved outreach presentations to students addressing stigma towards mental health concerns. Undergraduate psychology students from the University of Alberta were also recruited (*n*=9). The 87 participants who were recruited indicated an interest in participating in related research at the end of a knowledge and attitude survey, and were re-contacted by the research team. Participants agreed to attend the University of Alberta or Alberta Hospital Edmonton (AHE) to complete a neuropsychological evaluation and to provide a saliva sample for DNA analysis. The study protocol was approved by the Human Research Ethics Board of the University of Alberta. Informed consent was obtained for all participants included in the study.

### 2.1.2. Neuropsychological testing

Once participants arrived at the testing location they were administered tests of neuropsychological function including: the Reading Subtest from the Wide Range Achievement Test (WRAT4_R) (Wilkinson and Robertson 2006), the Conners Continuous Performance Test (CPT-0X) (Conners 1994), the Visual Patterns Test (VPT) (Della Sala, S., Gray, C., Baddeley, A., Wilson, L. 1997), and the Screen for Cognitive

Impairment in Psychiatry (SCIP) (Purdon 2005). The WRAT4 is a list of written vocabulary words whose pronunciation is used to provide an estimate of premorbid intellect (Wilkinson and Robertson 2006). The CPT-0X is a computer-based test where the user is shown a series of letters and are told to press a key for every letter except X. Performance provides an estimate of sustained attention (Rapisarda et al. 2014). The VPT is also computer-based, assessing non-verbal working memory by means of pattern recall (Della Sala, S., Gray, C., Baddeley, A., Wilson, L. 1997). The SCIP is a brief pencil and paper based test for the estimation of relative functioning in a broad range of cognitive functions associated with psychotic and affective illness (Cuesta et al. 2011; Purdon 2005). It includes brief tests of verbal list learning with immediate and delayed free recall (VLT-I and VLT-D), verbal working memory (WMT), verbal fluency (VFT), and visuomotor tracking (VMT). These tests were given in one sitting, after which the saliva sample was taken for DNA analysis.

### 2.1.3. Saliva sample collection and storage

Saliva was collected using Oragene-DNA OG-250 (DNAGenotek, OG-250) collection kits as per the user manual. Participants were asked not to eat or drink anything for at least 30 minutes prior to sample collection, to this end samples were collected after cognitive testing was completed. Before sample collection participants were asked to lick the inside of their cheeks to stimulate salivation, to fill the sample collection disk to the fill line and to reseal it, ensuring that the lid was on firmly to release the saliva stabilization solution. Samples were then mixed by the collecting technician, labelled

with date of collection, sample identification numbers and stored in a locked, sealed container at room temperature.

### 2.1.4. Genomic DNA isolation from saliva

Genomic DNA was isolated from 1 mL of the stabilized saliva sample collected in the Oragene DNA kit, per the manufacturer's protocol. Briefly, the saliva samples were mixed by inversion and incubated at 50°C for 2 hours in an air incubator (hybridization oven). An aliquot of 1 mL of the mixed sample was transferred to a microcentrifuge tube and 40 µL of Oragene-DNA purifier (OG-L2P) was added followed by vortexing and incubation on ice for 10 minutes. Samples were centrifuged for 15 minutes at 15 000×g at room temperature and the supernatant was retained in a new tube. To pellet DNA, 1 mL of 95% ethanol was added to the sample and the tube was mixed by inversion 10 times, followed by incubation at room temperature for 10 minutes for DNA precipitation. The DNA was pelleted at 15 000×g for 2 minutes, the supernatant was discarded. The DNA pellet was washed with 500 µL of 70% ethanol for 1 minute at room temperature and the ethanol was completely removed. The DNA was rehydrated in 100 µL of 10 mM Tris-HCl pH 8.0, and incubated overnight at room temperature followed by 1 hour at 50°C with intermittent vortexing to ensure complete rehydration. The DNA was stored at 4°C and working stocks were made at 50 ng/µL for genotyping.

### 2.1.5. *COMT* genotyping

*COMT* rs4680 (NM_000754.3: c.472G>A, NP_000745.1: p.V158M) was genotyped using an RFLP method modified from Malhotra et al. 2002. A 109 bp

fragment of *COMT* was amplified from genomic DNA by PCR using primers *COMT*genF: 5'-CTCATCACCATCGAGATCAA-3' and *COMT*genR 5'-CCAGGTCTGACAACGGGTCA-3'. Reaction conditions were standard for NEBTaq (NEB M0273), with 1X ThermoPol reaction buffer (10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, pH 8.3), 20 µM dNTPs, and 100 ng primers in a 20 µL reaction. Cycling conditions include an initial denaturation for 2 minutes at 95°C, followed by 32 cycles of 1 minute denaturation at 95°C, 1 minute annealing at 61.9°C, and 15 second extension at 72°C. The reaction was completed by a 5 minute final extension at 72°C. PCR products were visualized using the QIAxcel capillary electrophoresis apparatus (QIAGEN), with OL700 run parameters with 15 + 500 bp alignment markers on a high resolution cartridge. The PCR product was directly digested with a FastDigest *Nla*III (Fermentas FD1834) as follows: 10-17 µL of PCR product with 0-7 µL of ddH2O to bring the volume to 17 µL, followed by addition of 2 µL of 10X FastDigest buffer and 1 µL of FastDigest *Nla*III (final volume 20 µL) and incubated for 30 minutes at 37°C followed by heat inactivation at 80°C for 5 minutes. The digestion products were visualized using the QIAxcel capillary electrophoresis apparatus (QIAGEN) per above for undigested PCR products.

The A (Met) allele creates an additional 5'-C**A**TG-3' *Nla*III cut site relative to the G (Val) allele. As such, Met alleles result in the resolvable 68 bp and 18 bp digestion products. Val (G) alleles do not contain this site and result in the resolvable 86 bp and 22 bp products. Heterozygotes produce all four (86, 68, 22 and 18 bp) fragments. Genotypes were called based on the presence of the 86 bp and/or 68 bp bands, which were

consistently resolved and detected. Genotype accuracy was confirmed for a subset of the tested variants from an independent DNA aliquot.

### 2.1.6. *NPAS3* genotyping

Genotyping of *NPAS3* coding SNPs was achieved using PCR amplification of a region of exon 12 by two PCR products, 12A and 12B. Primers for PCR amplification and sequencing are as follows: *NPAS3*-12A-F: 5'- CCCGCTAACCTGGTGTCTTC-3', *NPAS3*-12A-R: 5'- CCGCCCTTTTGCCGTTTC-3', *NPAS3*-12B-F: 5'-GCACTCGGACTTTGAGAACC-3', *NPAS3*-12B-R: 5'-GGCTGTTAACGTCCACGAAG-3'. PCR conditions were as follows: for 12A 0.06 U Taq (QIAGEN PN: 201205) with included 10X reaction buffer diluted to 1X, 20 nM dNTPs, Q-solution used at 1X, 5 ng/µL primers in a 10 µL volume, with 25 ng of template genomic DNA. For 12B the same reaction conditions were used as 12A, but at double volume with addition of 150 ng of template genomic DNA. Thermocyling parameters were as follows for both fragments: 98°C for 10 min; 35 cycles of 98°C for 30 seconds, 57°C annealing for 45 seconds, 72°C extension for 2 minutes; reaction completed at 72°C for 5 minutes. PCR product 12A was purified by oligonucleotide removal (ExoSAP-IT, Affymetrix) and for 12B, PCR reactions were run on a 1% agarose 1X TAE gel for one hour at 110V, bands were extracted on a transilluminator, and DNA was purified by QIAquick Gel Extraction Kit (QIAGEN). Purified samples were quantified and 150 ng template was submitted for Sanger sequencing (The Applied Genomics Center, http://tagc.med.ualberta.ca/) using the PCR primers diluted to 3.2 pmol per reaction as sequencing primers. Sequencing reads were compared to the *NPAS3* exon

12 region using SeqScape sequencing analysis software (Thermo). Sequencing quality and genotype calls were confirmed by a blinded second reader. Genotype accuracy was confirmed for a subset of the tested variants from an independent DNA aliquot.

### 2.1.7. Statistical analysis of neuropsychological measures and genotype

General population assessment of descriptive statistics for IQ, age, ethnicity, sex as well as genotype and allele frequency calculations were performed using Excel (Microsoft). Variation in genotype frequency between our study population and the previous study population was performed using Fisher's Exact test, accessed from http://in-silico.net/statistics/fisher_exact_test/3x2 (In-silico 2016). Comparisons of *NPAS3* genotype with neuropsychological measures were performed in R (R Core Team 2013). Individuals homozygous for the minor allele were excluded due to small sample size ($n$=2). Individuals were grouped by genotype and for each measure compared, Levene's test was performed to test whether the groups had equal or unequal variance, after which a two sample $t$-test was performed for either equal or unequal variance. Association between *NPAS3* genotype and performance on cognitive tests was assessed using regression analysis assuming a linear model. The threshold $p$-value was set as 0.00455 by Bonferroni correction for multiple comparisons (11 cognitive measures by *NPAS3* genotype).

## 2.2 Characterizing the interaction between NPAS3 and ARNT

### 2.2.1. Identification of potential NPAS3 and ARNT co-targets

Microarray data of the top 50 up- and down-regulated genes by expression of *NPAS3* (Sha et al. 2012) was used as the source of genes to screen the ENCODE ARNT ChIP-seq data (experiment ENCSR155KHM, accessed from https://www.encodeproject.org/experiments/ENCSR155KHM/) visualized on the UCSC genome browser on human genome build 19 (Kent et al. 2002; Rosenbloom et al. 2010; Rosenbloom et al. 2015). The region within 1 kb of the transcription start site of the *NPAS3* regulated gene was screened for peaks called in both replicates 1 and 2 of the ChIP-seq experiment, with *p*-values reaching significance in both individual and pooled samples, as well as with optimal and conservative peak calling algorithms (score out of 5).

### 2.2.2. Antibody generation

An antibody against NPAS3 was generated by Pocono Rabbit Farms and Laboratory using the 28 day protocol and including the peptide design and synthesis service. The sequences of isoform 1 of human and mouse NPAS3 (NP_001158221.1, NP_038808.2 respectively) and NPAS1 (NP_002508.2, NP_032744.1) were aligned using ClustalW2 (Larkin et al. 2007) and the regions C-terminal to the final PAS motif (after residue 450) were used to design two potential peptides for antibody production (Figure 2-1). The more N-terminal peptide Nter-GNQSENSEDPEPDRK-Cter was selected due to its relative lack of homology to NPAS1, and its higher antigenic potential according to in-house analysis performed by Pocono Rabbit Farms and Laboratory. The

peptide was generated and 4 mg was KLH conjugated for injection into a guinea pig along side an immune stimulator and Freund's complete adjuvant. On days 7 and 14 after initial injection the guinea pig was boosted with antigen, immune stimulator and Freund's incomplete adjuvant. On day 21 a bleed was collected and serum was used to test the antibody on a western blot of untransfected U2OS and U2OS expressing an HA-tagged clone of NPAS3. After this test, the guinea pig was boosted one final time at day 44 and then exsanguinated on day 56. The NPAS3 antibody was affinity purified by Pocono Rabbit Farms and Laboratory from half of the serum using the unconjugated peptide and ELISA was performed to determine antibody concentration. The antibody was aliquoted and stored at -80°C, along side the unpurified serum. Working aliquots of the antibody were stored at 4°C.

A

```
NP_002508      NPAS1 432 ----------------------PGPEPTEPEPPTEGKQAAPAEN---EAPQTQGK----- 461
NP_032744      Npas1 433 ----------------------PGPEPTEEEPPVDGKQAVPADQDKDKDPQARGK----- 465
NP_001158221   NPAS3 474 ITEDNENSKSDEKGNQSENSEDPEPDRKKSGNACDNDMNCNDDGHSSSNPDSRDSDDSFE 533
NP_038808      Npas3 479 ITEDNENSKSDEKGNQSENSEDPEPDRKKSGSACDNDMNCNDDGHSSSNPDSRDSDDSFE 538
                                       * *: .:   . :..     :    . *:::..

NP_002508      NPAS1 462 --------------------RIKVE-------------------PGPRETKGSEDSGDED 482
NP_032744      Npas1 464 --------------------RIKVE-------------------ASPKEARGSEDSGEEE 486
NP_001158221   NPAS3 534 HSDFENPKAGEDGFGALGAMQIKVERYVESESDLRLQNCESLTSDSAKDSDSAGEAGAQA 593
NP_038808      Npas3 539 HSDFEHPKAAEDGFGALGPMQIKVERYVESEADLRLQPCESLTSDSAKDSDSANEAGAQA 598
                                      :****                    ..::: .: ::* :

NP_002508      NPAS1 483 PSSHP------------ATPRPEFTSVIRAGVLKQDPVRP-------------------- 510
NP_032744      Npas1 487 LSDPP------------APPRPEFTSVIRAGALKHDPVLP-------------------- 514
NP_001158221   NPAS3 594 SSKHQKRKKRRKRQKGGSASRRRLSSASSPGGLDAGLVEPPRLLSSPNSASVLKIKTEIS 653
NP_038808      Npas3 599 SSKHQKRKRRRKRQKGGSASRRRLSSASSPG-LDAGLVEPPRLLSSPHSASVLKIKTEIA 657
                          *.              :..* .::*.  .* *. . * *

NP_002508      NPAS1 511 -----------WG---------------------------------------------- 512
NP_032744      Npas1 515 -----------WG---------------------------------------------- 516
NP_001158221   NPAS3 654 EPINFDNDSSIWNYPPNREISRNESPYSMTKPPSSEHFPSPQGGGGGGGGGGGLHVAIPD 713
NP_038808      Npas3 658 EPINFDNDSSIWNYPPNREISRNESPYSMTKPPTSEHFPSPQGQGGSSGGGGALHVAIPD 717
                                   *.

NP_002508      NPAS1 513 --LAPPGDP------------------------------------------------- 519
NP_032744      Npas1 517 --LTTPGDP------------------------------------------------- 523
NP_001158221   NPAS3 714 SVLTPPGADGAAARKTQFGASATAALAPVASDPLSPPLSASPRDKHPGNGGGGGGGGGA 773
NP_038808      Npas3 718 SVLTPPGADGTAGRKTQFSG-----TAPVPSDPLSPPLSASPRDKHPG-------GGAGS 765
                           *:.**

NP_002508      NPAS1 520 -------PPTLLHAG-------------FLPPVVRG-------------LCTPGTIRYGP 546
NP_032744      Npas1 524 -------SPALLHAG-------------FLPPVVRG-------------LCTPGTIRYGP 550
NP_001158221   NPAS3 774 GGGGPSASNSLLYTGDLEALQRLQAGNVVLPLVHRVTGTLAATSTAAQRVYTTGTIRYAP 833
NP_038808      Npas3 765 GGGGPGASNSLLYTGDLEALQRLQAGNVVLPLVHRVTGTLAATSTAAQRVYTTGTIRYAP 825
                                . :**::*             .** * *             : *.*****.*

NP_002508      NPAS1 547 AELG--------------------------------LVYPHLQRLGPG------------ 562
NP_032744      Npas1 551 AELS--------------------------------LMYPHLHRLGAG------------ 566
NP_001158221   NPAS3 834 AEVTLAMQSNLLPNAHAVNFVDVNSPGFGLDPKTPMEMLYHHVHRLNMSGPFGGAVSAAS 893
NP_038808      Npas3 826 AEVTLAMQGNLLPNAHAVNFVDVNSPGFGLDPKTPMEMLYHHVHRLNMSGPFGGAVSAAS 885
                         **:                                ::* *::**. .

NP_002508      NPAS1 563 -PALP--------EAFYPPLGLPYPG---PAGTRLPRKGD 590
NP_032744      Npas1 567 -PSLP--------EAFYPTLGLPYPG---PTGTRVQRKGD 594
NP_001158221   NPAS3 894 LTQMPAGNVFTTAEGLFSTLPFPVYSNGIHAAQTLERKED 933
NP_038808      Npas3 886 LTQMPGGNVFTTAEGLFSTLPFPVYSNGIHAAQTLERKED 925
                          . :*        *.::..* :*  .    :.  : ** *
```

B

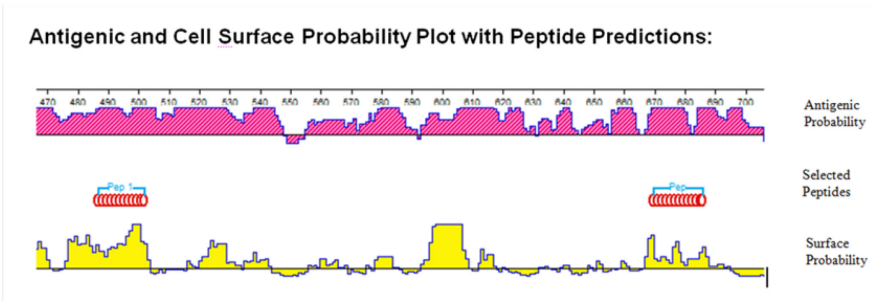



**Figure 2-1: Alignments and peptide analysis used to select NPAS3 peptides for antibody generation in guinea pigs.**
(A) ClustalW2 Alignment of the C-terminal half of NPAS3 isoform 1 with NPAS1 isoform 1, as well as mouse Npas3 isoform 1 and Npas1 isoform 1, from residue 474, which do not contain the bHLH DNA binding or PAS interaction domains, regions of high homology between NPAS3 and NPAS1. Two peptides were selected for further analysis performed by Pocono Rabbit Farms and Laboratory (B). Peptide 1 (N-terminal) was selected for synthesis, KLH conjugation and injection into three guinea pigs for antibody production.

### 2.2.3. cDNA clones

Clones of the complete coding sequence of two *NPAS3* transcript variants were used in this study. *NPAS3* cDNA clones of transcript variant 1, NM_001164749.1 (encoding NP_001158221.1, 933 aa) coding sequence (Origene RG228619), and transcript variant 2, NM_022123.1 (formerly transcript variant 1, encoding NP_071406.1, 901 aa) coding sequence (Origene RG215493) were purchased from Origene. Unless otherwise specified, NPAS3 isoform 1 (933 aa) was used in experiments.

Two *ARNT* clones were purchased. HaloTag-*ARNT* was purchased from Promega (FHC01533, NM_001286036.1, transcript variant 6). Relative to transcript variant 1 it is alternatively spliced, resulting in a difference of two residues, c.2740-2745delAGGAAT (p.R600-N601del) downstream of the bHLH and PAS domains (https://www.ncbi.nlm.nih.gov/gene/405), as well as a synonymous variant c.768G>C (p.Val188=) that is not predicted to change the protein. A clone of *ARNT* transcript variant 1 was also purchased (GenScript OHu21515, NM_001668.3, untagged) which was found to have no variants and has been used for all other applications.

*HIF1A* (DNASU ID HsCD00505780, NM_001530.3, transcript variant 1) and *BMAL1* (*ARNTL*, DNASU ID HsCD00079688, NM_001178.5, transcript variant 1) were purchased from DNASU in Gateway entry vectors (Seiler et al. 2014). The *HIF1A* clone contains a synonymous variant c.416C>T (p.Ala138=) which does not affect protein sequence. The full coding region of each clone was sequenced and confirmed to be the same as the intended construct before being cloned into destination vectors. Sequencing primers used are listed in Table 1.

**Table 1: Primers used to sequence clones used in this study**

| Primer name | Sequence | Usage notes |
|---|---|---|
| NPAS3CDS-1R-LL | CCCAAATGTGCTTCAAATAC | |
| NPAS3CDS-2F-LL | TCTGAAAATGAGGGACTTTGC | |
| NPAS3CDS-2R-LL | AGCGACACTCTCAGGCGTAG | |
| NPAS3CDS-3F-LL | CAGCTCAGCATCTTCCTCCT | |
| NPAS3CDS-3R-LL | AGCTATGGTGGCACTGGACT | |
| NPAS3CDS-4F-LL | TGCTACCACTTCATCCATGC | |
| NPAS3CDS-4R-LL | GTAGCGCTCCACCTTGATCT | |
| NPAS3CDS-5F-LL | AGTCCGAGAACAGCGAAGAC | |
| NPAS3CDS-5R-LL | ATGCTGTAGGGGGACTCGTT | |
| NPAS3CDS-6F-LL | TGTCCTCCCCCAACAGTG | |
| NPAS3CDS-6R-LL | CCCGTGGTGTAGACCCTCT | |
| NPAS3CDS-7F-LL | CAACTCCTTGCTGTACACTG | |
| ARNT-CDS-1R | ACGGCTTATAGGAGCCATCA | |
| ARNT-CDS-2F | TACGCATGGCAGTTTCTCAC | |
| ARNT-CDS-2R | GACCACCACGAAGTGAGGTT | |
| ARNT-CDS-3F | GGAAGGTCAGCAGTCTTCCA | |
| ARNT-CDS-3R | CATCCAGAGCCATTCTTGGT | |
| ARNT-CDS-4F | CTAGTGGCCATTGGCAGATT | |
| ARNT-CDS-4R | TAGCTGGCCAGTCCATCTCT | |
| ARNT-CDS-5F | AGCAGCAACAGCAAACAGAA | |
| ARNT-CDS-5R | GACATGGAGCTGAAGGAGGA | |
| ARNT-CDS-6F | TTCTGCAGGACAGATGTTGG | |
| BMAL1CDS1R | CTTCCCTCGGTCACATCCTA | |
| BMAL1CDS2F | TTAAGAGGTGCCACCAATCC | |
| BMAL1CDS2R | GACGAGGCAGCTGAGGTTAC | |
| BMAL1CDS3F | TTCCCCTCTACCTGCTCAAA | |
| BMAL1CDS3R | TCCAGGACGTTGGCTAAAAC | |
| BMAL1CDS4F | CGGAGTCGATGGTTCAGTTT | |
| BMAL1CDS4R | TCCTTGGTCGTTGTCAATCA | |
| BMAL1CDS5F | AATAGGCCGAATGATTGCTG | |
| HIF1ACDS1R | TTCACAAGGCCATTTCTGTG | |
| HIF1ACDS2F | TTTGAAAGCCTTGGATGGTT | |
| HIF1ACDS2R | TCAGATGATCAGAGTCCAAAGC | |
| HIF1ACDS3F | CAGTCGACACAGCCTGGATA | |
| HIF1ACDS3R | TCGTTGCTGCCAAAATCTAA | |
| HIF1ACDS4F | CAGCACGACTTGATTTTCTCC | |

| | | |
|---|---|---|
| HIF1ACDS4R | TGCTTCTGTGTCTTCAGCAAA | |
| HIF1ACDS5F | TTTTACCATGCCCCAGATTC | |
| HIF1ACDS5R | CCTTTTCCTGCTCTGTTTGG | |
| HIF1ACDS6F | GAAAGCGCAAGTCCTCAAAG | |
| TXNIP-prom-LL-F | AGCGCAACAACCATTTTCCC | |
| Hs.VGFprom-qPCR2LL-R | CTGGTCGGCTCTTGAATCTTTA | |
| M13F | GTAAAACGACGGCCAGT | pDONR F |
| M13R | CAGGAAACAGCTATGAC | pDONR R |
| T7S | TAATACGACTCACTATAGG | pCI-HA, pcDNA3.1-nV5/DEST F |
| pcI-R | GGTTTGTCCAAACTCATCAATGT | pcI-HA, pcDNA3.1-nV5/DEST R |
| EBV-R | GTGGTTTGTCCAAACTCATC | pHTN-CMV neo R |
| RVprimer3 | CTAGCAAAATAGGCTGTCCC | pGL4.10 F |
| GW primer1 | CACATTATACGAGCCGGAAGCAT | Gateway cassette |
| GW primer2 | CAGTGTGCCGGTCTCCGTTATCG | Gateway cassette |

### 2.2.4. Gateway conversion of pHTN-CMV-neo

The pHTN-CMV-neo HaloTag vector (Promega) was converted to a Gateway compatible vector using the Gateway conversion kit (Invitrogen). pHTN-CMV-neo (5 µg) was digested with 2 µL of FastDigest *Eco*RV (Thermo) in 1X FastDigest buffer for 5 minutes at 37°C, followed by deophosphorylation with FastDigest Alkaline Phosphatase (Thermo) for 10 minutes at 37°C and phosphatase inactivation at 65°C for 15 minutes. The linearized plasmid was run on a 0.8% agarose 1X TBE gel and the 6143 bp band was gel extracted using the QIAquick Gel Extraction Kit (QIAGEN). The digested vector backbone was quantified and 20 ng was added to a ligation reaction with 10 ng of the Gateway Cassette (reading frame B) using Quick Ligase (NEB) in 1X final Quick Ligase reaction buffer. The ligation was incubated overnight at room temperature.

The Gateway vector ligation (5 µL) was then transformed into One Shot *ccdB* survival chemically competent *E. coli* (Invitrogen) and transformants were plated onto

chloramphenicol-containing plates to select for the Gateway cassette. Single colonies were streaked on both chloramphenicol and ampicillin plates to confirm both resistance conferred by the Gateway cassette (chloramphenicol) and the pHTN-CMV-neo backbone (ampicillin). The orientation of the insert was confirmed by restriction digest and the sequence confirmed by sequencing using GW primer 1 and GW primer 2 listed in Table 1. Of note: kanamycin resistance was observed due to unexpected expression of the neomycin resistance (*neo*$^{R}$) cassette. As such for all clones recombined into this vector, kanamycin selection cannot be used to verify correct recombination; restriction digest must be used to screen for correct recombination.

### 2.2.5. Gateway cloning

Gateway cloning is a recombination based cloning method that allows for rapid movement of a clone from a single entry vector into multiple vectors for expression in different model systems and for different purposes. *NPAS3* isoforms 1 (933 aa) and 2 (901 aa) were cloned into the Gateway recombination based cloning system from the clones purchased from Origene, which were found to be non-expressing. The *NPAS3* coding sequences were subcloned into the Gateway system using primers indicated in Table 2. Both *NPAS3* transcript variants were cloned using the CDS cloning F and R primers. KapaHiFi was used for PCR amplification with the following reaction parameters: a 12.5 µL reaction was set up with 2X Kapa HiFi Hot Start Ready Mix diluted to 1X, and the addition of 50 ng of template DNA (pCMV6-AC-GFP-NPAS3 both 901 aa and 933 aa isoforms) 7.5 µM of forward and reverse Gateway cloning primers and 10% DMSO. Cycling as follows: 95°C for 5 minute initial denaturation

followed by 40 cycles of 98°C denaturation for 20 seconds, 60°C for 20 seconds annealing and 72°C extension for 3 minutes. The reaction was completed with a 5 minute incubation at 72°C. Functional domains were cloned from *NPAS3* transcript variant 1 using primers listed in Table 2. Location of domains was determined using ScanProsite tool and InterPro domain prediction tools using the clone of the coding sequence of NPAS3 transcript variant 1 as template (de Castro et al. 2006; Finn et al. 2017).

**Table 2: Gateway cloning primers used for *NPAS3* constructs**

| Primer name | Sequence |
|---|---|
| NPAS3 CDS cloning F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGGCGCCC ACCAAGCCCAGCTTTCAG |
| NPAS3_bHLH_GW_R | GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATGTGTTAGGTG GAGGGCCTTCCATTCG |
| NPAS3_PAS_GW_F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCGAATGGAAGGC CCTCCACCTAACACA |
| NPAS3_PAS_GW_R | GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAATGGGGGAGC TGTGCGATGTCC |
| NPAS3_TAD_GW_F | GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCTGCCGGAGAAA ACTTCCGAATCC |
| NPAS3 CDS cloning R | GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAGTCCTCCTTGC GCTCCAGAGTCTG |

Gateway BP recombination reaction to insert the PCR product into the pDONR Gateway entry vector was performed as follows: the purified PCR product was quantified and the input fmol of ends was calculated by the following formula:

$$ng = fmol \times length\ in\ basepairs \times (\frac{660fg}{fmol}) \times (\frac{1ng}{10^6 fg})$$

For the recombination reaction 150 ng of pDONR was used and insert concentration was calculated to ensure an equimolar concentration of recombination sites (equivalent to DNA ends). The DNA was diluted in TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and 2 µL of BP clonase was added to make a 10 µL reaction, which was incubated overnight at room temperature. The next day 10 µg of proteinase K was added to terminate the

reaction, which was then stored on ice until transformation into chemically competent *E. coli* DH5α, per the transformation section of the methods below.

For LR clonase reactions into Gateway destination (expression) vectors, 150 ng of destination vector was used and the amount of entry vector added to the reaction was calculated by the formula incubated above to ensure equimolar concentration of recombination sites. The plasmids were diluted to 8 μL in TE and 2 μL of LR clonase mixture was added, incubated overnight, and terminated by addition of proteinase K, as for the BP reaction. After transformation of the LR recombination products, transformant colonies were streaked onto both kanamycin and carbenicillin plates to test for resistance of kanamycin (not desired, pDONR entry plasmid) and ampicillin (desired, pDEST expression plasmid). Colonies only resistant to ampicillin were selected and used to inoculate an overnight culture to generate plasmid for sequence analysis.

### 2.2.6. Promoter constructs

For promoter constructs, both ChIP-seq and transcription factor binding prediction by ConTra v2.0 were used to identify the interval of interest. The ARNT ChIP-seq peaks from ENCODE experiment ENCSR155KHM were used to define the use of a 1 kb interval for transcription factor binding site prediction. ARNT binding sites were predicted using ConTra v2.0 analysis (Broos et al. 2011) with stringency of core = 0.95, similarity matrix = 0.85 was used, and the following ARNT complex binding sites were selected for analysis:

Ahr::Arnt (JASPAR_CORE_2016,MA0006.1,MA0006.1),

Arnt (JASPAR_CORE_2016,MA0004.1,MA0004.1),

ARNT::HIF1A (JASPAR_CORE_2016,MA0259.1,MA0259.1).

For *TXNIP*, the promoter region of interest was limited to 1000 bp upstream of the transcription start site of *TXNIP* as defined by transcript variant 1, NM_006472.1, supported by RNAseq data. The 5'UTR and intron 1 of *TXNIP* were also assessed for ARNT complex binding. For *VGF,* the promoter region assessed included the 1000 bp upstream of the transcription start site as defined by transcript variant 1 (NM_003378.3) and supported by RNAseq data. The 5'UTR and intron 1 of *VGF* were also analysed. The minimum interval to include the promoter proximal ARNT ChIP-seq peaks and ConTra v2.0 predicted ARNT complex binding sites and approximately 50nt of the 5'UTR was selected for synthesis. As these constructs were to be cloned by Gibson assembly, the synthesized DNA fragments included 20 nt overlapping the destination vector, pGL4.10, at both 5' and 3' ends of the construct for insertion into plasmid linearized by digestion with *Xho*I and *Hin*dIII. The adapter sequences were designed with the assistance of the NEBuilder Assembly tool (http://nebuilder.neb.com/, NEB). The sequence was submitted for synthesis by IDT DNA. Synthesized DNA fragments were rehydrated in TE and incubated for 20 minutes at 55°C to make a 10 ng/µL solution for Gibson Assembly cloning. For figures, ConTra v3.0 was used as it allowed for direct export of data to the UCSC genome browser, and identified identical regions relative ConTra v2.0 (Kreft et al. 2017).

### 2.2.7. Gibson Assembly: promoter constructs

Gibson Assembly was performed as described by the Gibson Assembly Cloning Kit (NEB) protocol. First, 2 µg of pGL4.10 was digested with 10 U each of *Xho*I and

*Hin*dIII (both NEB) in 1X NEBuffer 2.1 with 100 ng/µL BSA for 3 hours at 37°C. The digest was run on a 1% agarose 1X TBE gel, visualized and the expected 4208 bp band was extracted and purified using the QIAquick Gel Extraction Kit (QIAGEN) per the manufacturer's protocol and eluted in 30 µL of the kit elution buffer. DNA was quantified by Nanodrop. For promoter constructs, synthesized fragments were used in place of PCR fragments at a concentration of 10 ng/µL. Molarity was calculated using the formula:

$$ng = pmol \times length\ in\ basepairs \times (\frac{650pg}{pmol}) \times (\frac{1ng}{10^3 pg})$$

A three-fold molar excess of insert to vector was used, 0.045 pmol of insert and 0.015 pmol of linearized vector.

Fragments were diluted to 10 µL in dd$H_2O$ and 10 µL of 2X Gibson Assembly Master Mix was added to make a final 20 µL reaction. The reaction was incubated at 50°C for 15 minutes and then 2 µL of the reaction was added to a freshly thawed 50 µL aliquot of NEB 5-alpha Competent *E. coli* and incubated on ice for 30 minutes. Transformation reactions were heat shocked at 42°C for 30 seconds and then incubated on ice for 2 minutes. Room temperature SOC media was added to a final volume of 1 mL and cells were recovered at 37°C for 1 hour with shaking at 250 rpm. 100 µL of the recovered cells were streaked on pre-warmed LB plates supplemented with carbenicillin, which were inverted and incubated overnight at 37°C. Colonies were streaked out for single colonies on fresh LB plates supplemented with carbenicillin and incubated overnight. A fresh colony was selected to inoculate a 4 mL LB culture supplemented with ampicillin and incubated overnight at 37°C, with shaking at 250 rpm for DNA purification by miniprep the next day (QIAGEN). Sequences were confirmed using the

following primers: for the *VGF* promoter construct, RVprimer3 and VGF-prom-qPCR-R; for the *TXNIP* promoter construct, RVprimer3 and TXNIP-prom-qPCR2-F, sequences listed in Table 1.

### 2.2.8. Site-directed mutagenesis

Two variants of *NPAS3* were generated in the pDONR-*NPAS3*-933KX vector using site directed mutagenesis. Generation of c.1654G>C was achieved using the KapaHiFi kit with GC buffer (Kapa Biosystems). pDONR-*NPAS3*-933KX was used as template at 11 ng/µL, with 1X GC reaction buffer, 0.75 µL of 2 mM dNTP mix, 60 ng of each mutagenesis primer (*NPAS3*-G1654C-F 5'-CGGTGCTCTGGGCCCGATGCAGATCAA-3', *NPAS3*-G1654C-R 5'-TTGATCTGCATCGGGCCCAGAGCACCG-3', variant nucleotide underlined), and 0.5 µL of KapaHiFi polymerase for a final 25 µL reaction. Cycling conditions were as follows: initial denaturation at 95°C for 2 minutes, 18 cycles of 20 second denaturation at 98°C, 15 second annealing at 68°C and a 6 minute extension at 72°C; followed by 5 minutes at 72°C to complete the reaction.

Generation of c.2089G>A also used Kapa HiFi polymerase. The reaction conditions were as follows: 1X KapaHiFi Hot Start Ready Mix buffer (Kapa Biosystems) supplemented with 5% DMSO, 0.75 µM of each mutation containing primer (*NPAS3*-G2089A-F 5'- CCCGCAGGGCAGCGGCGGTGG-3', *NPAS3*-G2089A-R 5'-CCACCGCCGCTGCCCTGCGGG-3', variant nucleotide underlined) and 10 ng of pDONR-*NPAS3*-933KX template DNA. Cycling conditions were as follows: initial denaturation of 95°C for 2 minutes, 20 cycles of 98°C for 20 seconds, 65°C for 15

seconds and 72°C for 7 minutes. Reactions were completed with 5 minutes at 72°C. *Dpn*I (NEB) was added (0.25 µL) directly to the PCR reaction and incubated for 1 hour at 37°C. Of the mutagenesis reaction 2 µL were used to transform DH5α chemically competent *E. coli,* as per the protocol described in the transformation section.

### 2.2.9. Gibson Assembly mediated mutagenesis

For introduction of the c.910G>A mutation in both isoforms 1 and 2 of *NPAS3*, a 504 bp fragment of *NPAS3* (invariant between transcript variants 1 and 2) containing the mutated residue in between an *Avr*II site and a *Xho*I site with 20 bp adapters complementary to the 5' and 3' ends of the linearized *NPAS3* coding sequence was generated using the NEBuilder tool (accessed from http://nebuilder.neb.com). The destination vector, pDONR-*NPAS3*-901KX or pDONR-*NPAS3*-933KX was digested as follows: 500 ng of DNA was digested with 5 U of *Avr*II and *Xho*I (10 U total, both NEB) in 1X CutSmart buffer with 10 ng/µL BSA for 2 hours at 37°C. Digests were run on a 1% agarose 1X TBE gel and the 4890 bp band was cut out and extracted using the QIAquick Gel Extraction Kit (QIAGEN) and eluted in 30 µL of elution buffer. Fragments were quantified by Nanodrop. The synthesized fragment of *NPAS3* including the variant c.910G>A was resuspended in TE and incubated for 20 minutes at 55°C to ensure complete rehydration, resulting in a final concentration of 10 ng/µL for use as insert into the cut vector backbone. The Gibson Assembly was performed as per the NEB Gibson Assembly Cloning Kit protocol, as described in the Gibson Assembly: promoter constructs section with the following modification: a three-fold molar excess was used at 0.315 pmol of insert and 0.105 pmol of vector.

### 2.2.10. Transformation of chemically competent *E. coli* DH5α

DH5α chemically competent *E. coli* (Invitrogen) were used for transformations unless otherwise indicated. Cells were thawed on ice and DNA was added as indicated for each protocol, mixed by gentle swirling and incubated for 15-30 minutes on ice, followed by a 30 second heat shock at 42°C and 2 minutes on ice. LB broth was added to a final volume of 250 µL - 1000 µL and cells were allowed to recover for 1 hour at 37°C with shaking at 250 rpm. Different volumes of the transformation reaction were plated on pre-warmed LB agar plates with antibiotics, depending on the application, ranging from 10-100% of the reaction. Plates were inverted and incubated overnight at 37°C. Colonies were selected and streaked out on fresh LB antibiotic plates. Overnight cultures for plasmid purification were inoculated from the streaked single colonies.

### 2.2.11. Plasmid purification

For sequencing, recombination and molecular biology applications 4 mL LB liquid cultures with antibiotics added were inoculated with a single colony from an agar plate or a glycerol stock. Cultures were grown overnight at 37°C, 250 rpm. For the first preparation of a construct, 700 µL of the culture was reserved to make a glycerol stock with the addition of glycerol to 15% final concentration, and were stored at -80°C. The remainder of the culture was pelleted at 4696×g for 15 minutes at 4°C and the media decanted. Plasmids were purified using the Miniprep DNA kit (QIAGEN), eluted in 50 µL of elution buffer and quantified by Nanodrop.

For transfections, midipreps were performed to obtain high enough plasmid concentration with removal of Endotoxin A. A 1 mL LB + antibiotic starter culture was inoculated from single colonies or glycerol stocks and incubated for 6-8 hours at 37°C, 250 rpm. A 50 mL LB + antibiotic overnight culture were then inoculated with 500 µL of starter culture and grown overnight at 37°C, 250 rpm. Cultures were pelleted at 4696×g for 15 minutes at 4°C. Media was decanted and DNA was purified using the Midiprep Plasmid Plus Spin kit (QIAGEN) with the high yield protocol and eluted in 200 µL elution buffer. Plasmids were quantified by Nanodrop and diluted to make working stocks at 250 ng/µL for transfections to avoid multiple freeze-thaw cycles, and stored at -20°C.

### 2.2.12. Sequencing of constructs

Purified plasmid products were submitted for single-direction sequencing of the full-length coding region, or promoter region, by The Applied Genomics Centre (TAGC), an in-house sequencing facility. Sequencing read files were analyzed against a reference sequence using Sequencher (Genecodes). Constructs used were confirmed to be free of mutations unless otherwise indicated.

### 2.2.13. Cell culture: HEK 293T, U2OS

Cells were acquired from ATCC. Cells were kept in a humidified 5% $CO_2$, 37°C incubator, maintained in 75 $cm^2$ TC treated vented flasks (Sarstedt). HEK 293T and U2OS cells were maintained in 10 mL of Dulbecco's Modified Eagle Medium (DMEM), high glucose (Sigma M6406) supplemented with 10% FBS (Sigma F1051). Both lines

were maintained by subculturing every 2 days at 1:5 or every 3 days at 1:10, once cells reached 80% confluence. For subculturing media was aspirated, cells were washed once with PBS, then treated with Trypsin EDTA at 37°C for 2-10 minutes until released from the culture flask. Cells were then rescued with 5-10 mL of media supplemented with 10% FBS and subcultured at the indicated dilutions or plated for experiments. HEK 293T cells were also maintained in DMEM low glucose (Sigma M6429) as described above for corresponding experiments with low glucose.

For experimentation cells were plated at the following densities: $2.20\times10^6$ cells per 10 cm plate, $7.6\times10^5$ cells per 6 cm plate, $3.0\times10^5$ cells per 3.5 cm dish or well of 6-well plate or $5\times10^4$ cells per well of a 24-well plate.

### 2.2.14. Transfections

Cells were allowed to recover for 24 hours after plating, before transfection. Mirus TransIT-LT1 and Mirus TransIT Express (LT1 replacement, MirusBio) were used as follows. A 1.5 mL microcentrifuge tube was used to mix the transfection reagents. For a 10 cm plate, 4 µg of total transfected DNA was added to 500 µL of warmed serum-free growth medium and mixed by pipetting. Transfection reagent was added at a 3:1 v/m ratio (so here 12 µL of Mirus TransIT reagent) and mixed by pipetting. The transfection mixture was incubated for 30 minutes at room temperature and then gently mixed once before being added dropwise to cells, which were swirled to mix and returned to the incubator for 48 hours.

For smaller dishes transfection reactions were scaled down as follows: 6 cm dish, 250 µL media, 2 µg DNA, 6 µL TransIT reagent; 3.5 cm dish, 125 µL media, 1 µg DNA,

3 µL TransIT reagent. For luciferase experiments transfections were set up in triplicate with a 3.3X mastermix of 0.35 µg total DNA (50 ng pGL4.10-promoter-luciferase construct, 150 ng each driver/empty vector construct, 0.1 ng pGL4.7-TK-renilla Luciferase and 1.05 µL TransIT reagent per 1X reaction were added to serum free culture media to bring the final volume to 330 µL. The transfection mixture was incubated for 30 minutes and then 100 µL of mix was added dropwise to each well. Plates were gently swirled to mix and returned to the incubator for 48 hours. Transfections involving NPAS3 were performed using pcI-HA-*NPAS3* (HA-tagged), unless otherwise specified. Transfections involving ARNT were performed using pcDNA3.1-*ARNT* (untagged) unless otherwise specified. Transfections involving HIF1A and BMAL1 were performed using pcDNA3.1-nV5-*HIF1A* or *BMAL1* (both V5 tagged).

### 2.2.15. Total Protein Purification

Cells were harvested 48 hours after transfection, or at the indicated experimental time points. For harvesting, media was removed from dishes and the cells were washed twice in 1X PBS (10 mM $PO_4^{3-}$, 137 mM NaCl, 2.7 mM KCl, pH 7.4) before being collected in 1 mL of 1X PBS and pelleted by centrifugation at 1000×g for 5 minutes at 4°C. Cell pellets were frozen at -80°C before being processed. For total cell lysis, cells were thawed and directly lysed in Mammalian Lysis Buffer (MLB, 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Triton X-100 and 0.1% sodium deoxycholate) with freshly added protease inhibitor cocktail, diluted 1/200 (Sigma P8340), 1 mM PMSF, 1 mM DTT. HEK 293T cells were lysed in 250 µl MLB per 10 cm dish of cells input, while U2OS and SK-N-SH were lysed in 100 µL MLB per 10 cm dish. Cells in lysis buffer were sonicated

using the Sonic Dismembrator ultrasonic processor (Fisher FB-120) three times at 30% amplitude for 5 seconds on, 5 seconds off, and then incubated at 4°C for 30 minutes with rocking. Insoluble material was precipitated by centrifugation at 14 000×g for 5 minutes at 4°C and the lysate (supernatant) was collected. Lysates were stored at -80°C or diluted to 2.5 µg/µL (50 µg/20 µL) denatured for 5 minutes at 95°C in SDS sample buffer (125 mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 0.02% (v/v) β-mercaptoethanol, 0.005% bromophenol blue) and used directly or stored at -20°C.

### 2.2.16. Protein quantification

Proteins prepared as described previously were quantified using the BioRad Protein Assay (BioRad 500-0006) reagent diluted 1 in 5 in di$H_2O$ to a 1X final concentration. Samples (2 µL) were diluted in 1 mL 1X BioRad Protein Assay solution, vortexed briefly twice to mix and read on a GloMax Multi Jr Single Tube Multimode Reader (Promega) with the Absorbance Module (Promega model E6076) and 600 nm filter. The multimode reader was set to Photometer mode, blanked using diH2O and then used to measure absorbance. Protein concentration was calculated by interpolation on a standard curve processed each experiment using 0, 2, 5, 10, 20 and 30 µg of BSA added as 0-30 µL of 1 mg/mL BSA in di$H_2O$.

### 2.2.17. Immunoprecipitation

Once quantified, 300 µg of total protein lysates were used as input for immunoprecipitation. Protein A Dynabeads (Invitrogen) were washed three times in 1X PBS at room temperature and blocked in 1X PBS, 20% BSA for 1 hour at 4°C, with

rotation. Lysates were pre-cleared with 25 µL of Protein A Dynabeads and incubated for 90 minutes with rotation at 4°C. Lysates were diluted in immunoprecipitation (IP) buffer (150 mM NaCl, 50 mM Tris pH 8.0, 1% Triton X-100) to a final volume of 500 µL. Concurrent to the pre-clearing, blocked Protein A Dynabeads were pre-incubated with 2.5 µg of antibody (αHA, αNPAS3, see Table 3) diluted to 500 µL at 4°C for 90 minutes, with rotation. The supernatant was removed from pre-bound antibody::Protein A Dynabeads, and the pre-cleared lysate was added and incubated for 90 minutes, with rotation, at 4°C. The supernatant was removed from the beads and the IP reaction was washed five times with 500 µL of IP buffer for 10 minutes at 4°C. The final wash was completely removed and the protein was eluted from the beads by denaturation in 1X SDS-sample buffer (see western blot section for composition) and incubation at 95°C for 5 minutes. Whole IP reactions were loaded onto the gel alongside 30 µg of input protein for assessment of immunoprecipitation.

### 2.2.18. HaloTag pull downs

Before beginning, 200 µL of HaloLink resin was aliquoted and centrifuged at 800×g for 2 minutes at room temperature. Storage buffer was removed and the resin was washed 3 times in 800 µL of wash buffer (1X TBS, 20 mM Tris pH 7.5, 150 mM NaCl, + 0.05% IGEPAL CA-630) pelleting the resin at 800×g for 2 minutes between washes. Washed resin was kept in the final wash until lysates were ready to be applied.

48 hours post transfection of HaloTag fusion bait and prey proteins HEK 293T were washed twice with PBS and collected in 1 mL of 1X PBS in an Eppendorf tube and pelleted at 3000×g for 5 minutes at 4°C. PBS was aspirated and cell pellets were frozen at

-80°C for at least 10 minutes before being thawed on ice and resuspended in 300 µL of ice-cold mammalian lysis buffer (MLB, 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, 0.1% Na deoxycholate, 1mM DTT) supplemented with protease inhibitor cocktail lacking AEBSF (Promega G6521) and homogenized in a Dounce homogenizer by 25 passes of the B pestle. Lysates were cleared by centrifugation at 14 000×g for 5 minutes at 4°C and placed in a fresh Eppendorf tube.

For pull down, lysates were diluted with 700 µL of 1X TBS and an input sample of 10 µL was retained. Wash buffer was removed from the washed resin and the 1 mL of diluted lysate was applied to the resin and incubated for 30 minutes at room temperature, with rotation. After binding, the resin was pelleted at 800×g for 2 minutes and 10 µL of the unbound lysate was collected. The resin was washed four times with 1 mL of wash buffer, mixing by inversion, and pelleting at 800×g for 2 minutes each time. A final wash in 1 mL of wash buffer was applied and incubated for 5 minutes at room temperature with rotation. The resin was pelleted at 800×g for 2 minutes and the final wash removed. Proteins were eluted in 50 µL of SDS elution buffer (1% SDS, 50 mM Tris-HCl pH 7.5) for 30 minutes with shaking at room temperature. Eluates were collected and, alongside the input and unbound samples, 10 µL was processed for western blot analysis.

### 2.2.19. Western Blotting: Licor Protocol

Protein lysates prepared and quantified as previously described were diluted in a final 1X SDS-PAGE sample buffer (125 mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 0.02% (v/v) β-mercaptoethanol, 0.005% bromophenol blue) and denatured for 5 minutes at 95°C then chilled on ice, and condensate collected by brief centrifugation. Unless

otherwise noted, 50 µg of protein in 1X SDS-PAGE sample buffer were loaded per lane (10 lanes per gel) for electrophoresis on a 1 mm thick discontinuous 4% stacking (125 mM Tris-HCl pH 6.8, 0.1% SDS ) and 7.5-15% separating (375 mM Tris-HCl pH 8.8, 0.1% SDS) polyacrylamide gel in 1X SDS-PAGE running buffer (25 mM Tris-HCl, 192 mM glycine, 0.1% SDS) at 140 V for 60-75 minutes until the dye front reached the bottom of the gel. The PageRuler Plus Prestained protein ladder (Thermo 26619) was used as a size marker; 5 µL of ladder was run per lane for overexpression experiments, and 2.5 µL was run per lane for endogenous expression. Proteins were transferred to a nitrocellulose membrane through wet transfer with pre-chilled Towbin transfer buffer (20% methanol, 25 mM Tris, 192 mM glycine, 0.1% SDS) at either 110 V for 1 hour or 30 V for 18 hours at 4°C with stirring. Transfers were disassembled, rinsed once with diH2O, and blocked for 1-3 hours in undiluted LI-COR blocking buffer (LI-COR 927-40000).

For overexpression, blots were directly probed with primary antibodies diluted in a solution of 50% LI-COR blocking buffer in 1X PBS 0.05% Tween-20 (PBS-T) for 1 hour at room temperature or 16-20 hours at 4°C. For endogenous expression, blocked blots were rinsed twice with PBS-T before being probed with primary antibody diluted as described for overexpression for 16-20 hours at 4°C. Antibody dilutions used are listed in Table 3. Blots were washed four times with PBS-T for five minutes before being probed with secondary antibodies diluted per Table 3 in 50% Licor blocking buffer in PBS-T. Fully probed blots were then rinsed once in PBS-T, washed four times with PBS-T and then transferred to 1X PBS for scanning on the LI-COR Odyssey scanner using membrane settings with intensity 5.0 for both 600 nm and 800 nm channels for

overexpression and endogenous expression in HEK 293T and intensity 6.0 for SK-N-SH endogenous expression. Scans were processed using the LI-COR Image Studio software.

**Table 3: List of antibodies**

| Antibody target | Species | Conjugate | Clonality | Company | Product number | Western blot dilution | IF dilution |
|---|---|---|---|---|---|---|---|
| NPAS3 | guinea pig | - | polyclonal | n/a | n/a | 1 in 10 000 | 1 in 250 to 1 in 1000 |
| NPAS3-Nter | rabbit | - | polyclonal | ProSci | 4107 | 1 in 1000 to 1 in 5000 | 1 in 250 |
| NPAS3-IN | rabbit | - | polyclonal | ProSci | 4109 | 1 in 5000 | - |
| ARNT | rabbit | - | monoclonal | Cell Signaling Technologies | D28F3 | 1 in 10 000 | 1 in 1000 |
| HaloTag | mouse | - | monoclonal | Promega | G921A | 1 in 10 000 | 1 in 500 |
| HA | rabbit | - | polyclonal | Santa Cruz Biotechnology | SC-805 | 1 in 5000 | 1 in 250 |
| HA | mouse | - | polyclonal | Santa Cruz Biotechnology | SC-7392 | 1 in 5000 | 1 in 250 |
| V5 | Mouse | - | monoclonal | Invitrogen | S46-0705 | 1 in 5000 | - |
| AURKA | rabbit | - | monoclonal | Cell Signaling Technologies | D3E4Q | 1 in 1000 | - |
| NPAS1 | rabbit | - | polyclonal | ProSci | XAV-8519 | 1 in 5000 | - |
| TFIID | rabbit | - | polyclonal | Santa Cruz Biotechnology | SC-204 | 1 in 2500 | - |
| GAPDH | mouse | - | - | - | - | 1 in 10 000 | - |
| β-tubulin | mouse | - | monoclonal | Santa Cruz Biotechnology | SC-55529 | 1 in 5000 | - |
| Rabbit IgG | donkey | IRDye 680 | polyclonal | Life Technologies | A10043 | 1 in 25 000 | - |
| Mouse IgG | donkey | IRDye 680 | polyclonal | Life Technologies | A10038 | 1 in 25 000 | - |
| Guinea Pig IgG | donkey | IRDye 800 | polyclonal | LICOR | 926-32411 | 1 in 25 000 | - |
| Guinea Pig IgG | goat | AlexaFluor 594 | polyclonal | Invitrogen | A11076 | - | 1 in 1000 |
| Rabbit IgG | donkey | AlexaFluor 488 | polyclonal | Invitrogen | A21206 | - | 1 in 1000 |
| Mouse IgG | donkey | AlexaFluor 488 | polyclonal | Invitrogen | A21202 | - | 1 in 1000 |
| Mouse IgG | donkey | AlexaFluor 594 | polyclonal | Invitrogen | A21203 | - | 1 in 1000 |
| Mouse IgG | goat | Horseradish peroxidase | polyclonal | Thermo | 31431 | 1 in 10 000 | - |
| Rabbit IgG | goat | Horesradish peroxidase | polyclonal | Thermo | 32460 | 1 in 10 000 | - |

### 2.2.20. HaloCHIP

For ChIP experiments HEK 293T cells were transfected as indicated in the transfection protocol. Forty-eight hours after transfection a representative plate of cells

was counted, and the requisite number of experimental plates to achieve $1\times10^7$ cells per postive and negative HaloCHIP reaction were cross-linked for 10 minutes in 1% formaldehyde. Crosslinking was quenched with 0.125 M glycine and cells were washed twice with PBS. For each experimental condition $2\times10^7$ cells were collected and pooled with PBS, and pelleted at 1000×g for 5 minutes at 4°C. Cell pellets were resuspended in 2 mL of cytoplasmic lysis buffer (5 mM Tris-HCl, pH 8.0, 85 mM KCl, 1% IGEPAL CA-630) and incubated for 20 minutes followed by homogenization by 25 passes in a Dounce homogenizer with the B-pestle.

Nuclei were pelleted at 3000×g for 5 minutes at 4°C and resuspended at $1x10^7$ nuclear equivalents per 650 uL nuclear lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, 0.1% Na deoxycholate), sonicated for three cycles of 30 seconds on/off on an ice slurry at amplitude 6 (medium-high intensity). Following sonication, $MgCl_2$ was added to the lysis buffer to a 1 mM final concentration, then benzonase (Sigma E1014) was added at 250 U per $1x10^7$ nuclear equivalents and incubated 15 minutes at room temperature to enzymatically digest DNA (Pchelintsev, Adams, Nelson 2016). The reaction was quenched with addition of EDTA to a final concentration of 5 mM and a 1% input (6 µL) sample was reserved. Figure 2-2 shows the shearing of gDNA obtained by this protocol. Of the 1.3 mL of sheared chromatin input, 600 µL were used as input for each pull down reaction. For the non-specific pull down control, a HaloTag blocking reagent, which is covalently bound to the HaloTag rendering it unable to bind to the HaloTag resin, was added to half of the sample ("blocked") and incubated for 30 minutes at room temperature with rotation. The other half of the sample ("ChIP") was incubated along-side the blocked samples, without addition of blocking reagent.

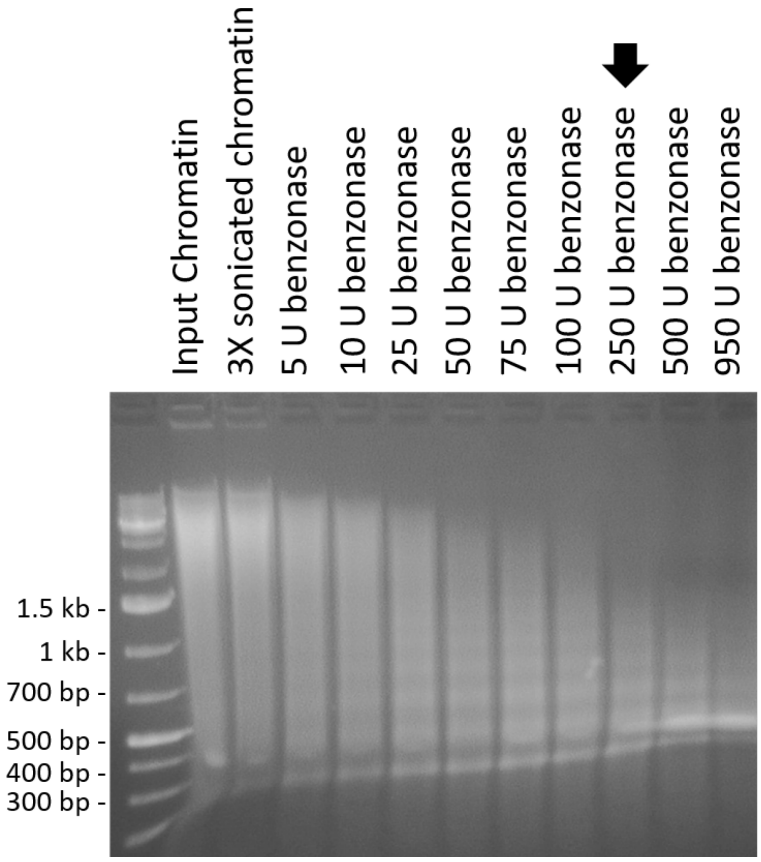



**Figure 2-2: Chromatin shearing obtained by the sonication and benzonase digestion protocol.**
Arrow indicates shearing conditions used for ChIP experiments.

Both blocked and ChIP samples were diluted to 2 mL in lysis buffer with 5 mM EDTA and applied to 150 µL prepared HaloTag resin (washed three times with 800 µL of TE + 0.01% IGEPAL), and incubated at room temperature for 2 hours with rotation. Samples were washed according to manufacturer's protocol (Promega HaloCHIP kit, 2 mL volume per wash): twice with 5 mM EDTA in nuclease free water, once with high salt wash buffer (50 mM Tris-HCl pH 7.5, 700 mM NaCl, 1% Triton X-100, 0.1% Na deoxycholate 5 mM EDTA), once with LiCl wash buffer (100 mM Tris-HCl pH 8.0, 500 mM LiCl, 1% IGEPAL CA-630, 1% Na deoxycholate) and twice more with nuclease-

free water. The buffer composition for the first three washes was supplemented with 5 mM EDTA relative to the manufacturer's protocol, as indicated in the composition above, in order to assure that the benzonase enzyme was not active and completely removed. Samples were eluted in 300 µL of elution buffer (300 mM NaCl, 10 mM Tris pH 8.0, 5 mM EDTA) at 65°C overnight. Again, 5 mM EDTA was added to ensure ChIP DNA stability. HaloCHIP DNA was purified using a QIAquick Gel Extraction clean up kit (QIAGEN), and eluted in 100 µL elution buffer. Enrichment of target DNA in ChIP samples over nonspecific background in blocked samples was assessed using qPCR or endpoint PCR.

### 2.2.21. ChIP PCR

PCR amplification of 1 µL of ChIP sample was performed using 0.25 U of GoTaq polymerase (Promega M300B) with green GoTaq reaction buffer used at a final 1X concentration. dNTPs were added to a final concentration of 50 µM, and primers to a final concentration of 1 µM (see Table 4 for primers used). Cycling conditions were as follows: initial denaturation at 95°C for 2 minutes, 30 cycles of 95°C for 30 seconds, 60°C for 15 seconds, 72°C for 15 seconds. The reaction was completed with a 5 minute incubation at 72°C. PCR reactions in green GoTaq reaction buffer were directly loaded onto a 1.5% agarose 1X TBE gel and run for 25-35 minutes at 125 V, then stained with 0.5 µg/mL ethidium bromide in ddH2O solution for 15-30 minutes, and destained for 30-60 minutes with ddH2O, if required. Gels were imaged on a Major Science UVDI CB-254/312 transilluminator imaging apparatus at 254 nm at various exposures.

**Table 4: Primers used for ChIP**

| Primer name | Sequence | Product size |
|---|---|---|
| TXNIPprom-qPCR2-LLR | AGGATCCCACTGACCCTAAA | 139 bp |
| TXNIPprom-qPCR2-LLF | CTGGCTAAGACTAGGCATGAAA | |
| TXNIPex1-qPCRLL-F | GTGATAGTGGAGGTGTGTGAAG | 120 bp |
| TXNIPex1-qPCRLL-R | CAGGTACTCCGAAGTCTGTTTG | |
| TXNIP-distCDS-LL-F | GAGTGTGGGTCCACCTTAGC | 115 bp |
| TXNIP-distCDS-LL-R | TGTATCACAACATGGGCGCT | |
| VGFprom-qPCR2LL-F | GCGTTGCTGAGTGGAATAGA | 105 bp |
| VGFprom-qPCR2LL-R | CTGGTCGGCTCTTGAATCTTTA | |
| VGF-intron1-LL-F | GTGTTTGCAACACCCCCATC | 92 bp |
| VGF-intron1-LL-R | AATACTCCGCTGTTCGTCCC | |
| VGF-distCDS-LL-F | TCAGTCCAAGTAGCGCCAAG | 93 bp |
| VGF-distCDS-LL-R | GGCTCTCCAGATTCACTCGG | |
| ANKRD37-prom-LL-F | CGGGAGACGTGTCAAACTCA | 71 bp |
| ANKRD37-prom-LL-R | AGAGGACTGGAGTAGGAGCG | |
| ANKRD37-distCDS-LL-F | GCTCCTGCTTTTGGCTTTACC | 136 bp |
| ANKRD37-distCDS-LL-R | ACAAAGCATACACAACCAGCA | |

### 2.2.22. HaloCHIP qPCR

qPCR was used to probe HaloCHIP DNA for *ANKRD37*. KAPA SYBR FAST Universal 2X qPCR master mix (5 µL per reaction) was mixed with 1 µL of 2 µM mix of forward and reverse primers (see Table 4 for primers used) and 2.5 µL of water per sample to create a master mix for all samples used. A triplicate master mix was made with 28.05 µL of primer master mix (8.5 µL per reaction) and 4.95 µL of HaloCHIP DNA (1.5 µL per reaction) and 10 µL was aliquotted into three wells of a 96-well plate for triplicate reactions of each condition. Samples were pelleted by brief centrifugation in a plate centrifuge and reactions were run in a BioRad CFX96 Touch using a two-step reaction: a 3 minute initial denaturation at 95°C followed by 40 cycles of 5 second

denaturation at 95°C and 25 second extension at 60°C. A melt curve to assess proper amplification was performed for every qPCR as follows: 10 second denaturation at 95°C followed by 5 second incubation at 0.5°C increments from 65°C to 95°C.

### 2.2.23. Luciferase

HEK 293T cells were plated at $5\times10^4$ cells per well of a 24-well plate in triplicate per luciferase condition assessed and allowed to recover for 24 hours before being transfected in triplicate, as described above. Forty-eight hours after transfection, cells were harvested using the Dual-Luciferase Reporter Assay System (Promega). Prior to harvesting cells, the kit reagents were brought to room temperature to equilibrate. At the time of luciferase assay media was removed from wells, cells were washed twice in 1X PBS for five minutes and then lysed with freshly diluted 1X passive lysis buffer for 30 minutes at room temperature. For *VGF* promoter experiments, 100 µL of passive lysis buffer was used. For *TXNIP* promoter experiments, 500 µL of passive lysis buffer was used. Luciferase Assay Reagent II was divided into 95 µL aliquots per reaction in individual 1.5 mL microcentrifuge tubes and briefly centrifuged. Stop-N-Glo reagent was diluted to 1X in Stop-N-Glo buffer.

Luminescence was measured using the GloMax Multi Jr Single Tube Multimode Reader (Promega) with the Luminescence Module, with Promega protocol DLR-0INJ. To read samples the lysed cell solution was added to an aliquot of Luciferase Assay Reagent II (20 µL for *VGF* promoter experiments, 2 µL for *TXNIP* promoter experiments), reporter driven luciferase luminescence was read, then 95 µL of Stop-N-Glo was added to terminate the reporter luciferase and initiate the control Renilla

luciferase expression for normalization. Renilla luciferase was then read and relative luminescence units (RLUs) calculated by the luminometer. Data were exported to Excel for statstical analysis and visualization.

### 2.2.24. Immunofluorescence

Coverslips were sterilized with 70% ethanol and placed in 6-well plates or 3.5 cm dishes and cells were plated at $3x10^5$ cells per dish in 2 mL of media. At the time of harvest cells were washed twice in 1X PBS, and fixed with 100 µL 2% paraformaldehyde buffered in 1X PBS pH 7.4 for 20 minutes at room temperature. The paraformaldehyde was removed, coverslips were rinsed once and washed twice for 5 minutes with 1 mL PBS-X (PBS with 0.05% Triton-X 100). Slides were blocked with 1 mL of 5% BSA in PBS-X for 15 minutes and then probed with primary antibodies diluted in block for 1 hour at room temperature. Antibodies were used at concentrations given in Table 3. Slides were then washed twice for 5 minutes with 1 mL PBS-X and probed with secondary antibody for 1 hour at room temperature in the dark. Slides were washed twice for 5 minutes with 1 mL PBS-X and finally incubated for 5 minutes with DAPI (2 µg/mL), diluted in PBS, and then mounted onto slides with ProLong Gold antifade reagent (Invitrogen), allowed to set overnight at room temperature and then sealed with nail polish. Slides were stored at 4°C until viewing.

### 2.2.25. Microscopy

Slides were visualized using a Leica DM RE fluorescent microscope. DAPI fluorescence was assessed using the CHROMA #31000 DAPI/Hoechst/AMCA filter,

AlexaFluor-488 by the CHROMA #41001 HQF FITC/BIODIPY/FLUO3/DiO filter and AlexaFluor-594 by the CHROMA #31002 TRITC filter. Images were captured using a Q imaging 32-0030B QICAM and Northern Elite Eclipse software (EMPIX). Scale bars were added using ImageJ (Schneider, Rasband, Eliceiri 2012).

### 2.2.26. RNA purification

Cells were harvested at indicated timepoints as follows: media was removed and cells washed twice with 1X PBS. The second wash was aspirated and then the RNeasy RNA purification kit (QIAGEN) RLT lysis buffer was added directly to the plate. Lysed cells were scraped and collected in a 1.5 mL microcentrifuge tube on ice and homogenized by 10 passes with a 21½ gauge needle. Homogenized samples were processed with the RNeasy RNA purification kit (QIAGEN) per manufacturer's protocols. eluted in 30 µL of nuclease free water and quantified by Nanodrop.

### 2.2.27. Two-step cDNA synthesis

cDNA was synthesized from 500 ng of RNA using the Quantitect RT kit (QIAGEN). RNA was diluted with water to 12 µL and 2 µL of 7X gDNA wipeout mix was added. Wipeout reactions were incubated in a 42°C heat block for 10 minutes to enzymatically remove gDNA. Reactions were placed on ice as a master mix of 1 µL of the included primer mix, 1 µL of the RTase enzyme and 4 µL of the 5X RT buffer per reaction was mixed. 6 µL of the RT master mix was added to the wipeout reaction and incubated in a 42°C heat block for 30 minutes, followed by a 3 minute denaturation at

95°C. cDNAs were diluted 1:25 in di$H_2O$ for input into qPCR reactions and stored at -80°C.

### 2.2.28. qPCR analysis of cDNAs

A primer-specific master mix was made using KAPA SYBR FAST Universal 2X qPCR master mix (5 µL per reaction, KapaBio) mixed with 1 µL of 2 µM mix of forward and reverse primers (see Table 5 for primers used). A triplicate master mix was made with 19.8 µL of primer master mix (6 µL per reaction) and 13.2 µL of diluted cDNA (4 µL per reaction) and 10 µL was aliquotted into three wells of a 96-well plate for triplicate reactions of each condition. Reactions were briefly centrifuged in a plate centrifuge and reactions were run in a BioRad CFX96 Touch using a two-step reaction: a 3 minute initial denaturation at 95°C followed by 40 cycles of 5 second denaturation at 95°C and 25 second extension at 60°C. A melt curve to assess proper amplification was performed for every qPCR as follows: 10 second denaturation at 95°C followed by 5 second incubation at 0.5°C increments from 65°C to 95°C. All qPCR data were normalized to three housekeeping genes (a combination of three of: *β-actin*, *HMBS*, *HPRT1*, *SDHA* and/or *YWHAZ*) from different pathways that were found to be stable among experimental replicates. qPCR data were analyzed using the BioRad CFX manager software and exported to Excel for visualization.

**Table 5: Primers used for RT-qPCR**

| Primer name | Sequence |
|---|---|
| TUBB3-PT-1 | CCTCCGTGTAGTGACCCTT |
| TUBB3-PT-2 | GGCCTTTGGACATCTCTTCAG |
| SOX1-PT-1 | CCACATCCTAATCTTGAGCCA |
| SOX1-PT-2 | CTGACGTCCACTCTCAGTCT |
| HMBS-qPCR-F-LL[a] | GGCAATGCGGCTGCAA |
| HMBS-qPCR-R-LL[a] | GGGTACCCACGCGAATCAC |
| SDHA-qPCR-F-LL[a] | TGGGAACAAGAGGGCATCTG |
| SDHA-qPCR-R-LL[a] | CCACCACTGCATCAAATTCATG |
| HPRT1-qPCR-F-LL[a] | TGACACTGGCAAAACAATGCA |
| HPRT1-qPCR-R-LL[a] | GGTCCTTTTCACCAGCAAGCT |
| ARNT-mRNA-qPCR-F | GAGAATTTCAGGAATAGTGGCCT |
| ARNT-mRNA-qPCR-R | CGAGTCTTAGCAGTAGCCTGG |
| ANKRD37-qRT-PCR-LLF | TAGGAGAAGCTCCACTACACA |
| ANKRD37-qRT-PCR-LLR | GCTGTTTGCCCGTTCTTATTAC |
| TXNIP-qRT-PCR-LLF | CGATAGTTTCGGGTCAGGTAAA |
| TXNIP-qRT-PCR-LLR | TTGGCTCTTCTCCACATGATAC |
| TXNIPex1-qPCRLL-F | GTGATAGTGGAGGTGTGTGAAG |
| TXNIPex1-qPCRLL-R | CAGGTACTCCGAAGTCTGTTTG |
| VGF-qRT-PCR-LLR | TCCTCCCTTGCACTCTCT |
| VGF-qRT-PCR-LLF | GACCTGCTGCTCCAGTATTT |
| NPAS3-3UTR-qRT-FLL | CCCAAACCGTGGTGGAATAA |
| NPAS3-3UTR-qRT-RLL | GGCCTTCTTCACCTCCTAAAC |
| HIST1H4H-qRTPCR-LLF | CAAGCGAATTTCTGGCCTTATC |
| HIST1H4H-qRTPCR-LLR | TTTGGCGTGCTCTGTGTAA |
| ATF5-qRTPCR-LLF | GTCTATGCCCGTCACATAACA |
| ATF5-qRTPCR-LLR | CCAGACAACCACCTGTAAGAA |
| DHCR24-qRTPCR-LLF | ACAGCATCAGGTGGGAAAG |
| DHCR24-qRTPCR-LLR | GGGATGAGTGGTTGGAGAAAT |
| ZBTB40-qRTPCR-LLF | TGGAGTTTCTGCTGGAAGTG |
| ZBTB40-qRTPCR-LLR | CCAGGCTTCAGGTAAGGAATAC |
| NCLN-qRTPCR-LLF | GCCTCAGCTTCCTCATCAATAG |
| NCLN-qRTPCR-LLR | ACCTCCCTCTCTGAGTTCCA |
| MAT2A-qRTPCR-LLF | TATCACCCAACGCTCCAAAG |
| MAT2A-qRTPCR-LLR | CATTGCCAGACAGAGGCTATAA |
| RNASE4-qRTPCR-LLF | CTCTGTCTCCTCAGCTCATTTC |
| RNASE4-qRTPCR-LLR | AGCCCAGCCTCATTCATTAC |
| ANG-qRTPCR-LLF | GGATAACTCCAGGTACACACAC |

| | |
|---|---|
| ANG-qRTPCR-LLR | CCGTCTCCTCATGATGCTTT |
| RPL37-qRTPCR-LLF | GGAGTACCACTGGAAACGTATG |
| RPL37-qRTPCR-LLR | CACTTAGCTAGCCACCTTACAC |
| ZNF581-qRTPCR-LLF | GAATCTGCGCCATCTTCCT |
| ZNF581-qRTPCR-LLR | CAGAGTGGGAAACGTGTTTATTG |
| USP49-qRTPCR-LLF | GACCTTTGCCTATGATCTCTCC |
| USP49-qRTPCR-LLR | CTCCCTCTGTGTTGTAGCAATAG |
| p21-CDKN1A-qRTPCR-F-LL | CGGAACAAGGAGTCAGACATT |
| p21-CDKN1A-qRTPCR-R-LL | AGTGCCAGGAAAGACAACTAC |
| p53-qRTPCR-F-LL | AGGGATGTTTGGGAGATGTAAG |
| p53-qRTPCR-R-LL | CCTGGTTAGTACGGTGAAGTG |
| BAX-qRTPCR-F-LL | GTCACTGAAGCGACTGATGT |
| BAX-qRTPCR-R-LL | CTTCTTCCAGATGGTGAGTGAG |
| BCL2-qRTPCR-F-LL | GGGAATCGATCTGGAAATCCTC |
| BCL2-qRTPCR-R-LL | CCCATCAATCTTCAGCACTCT |
| UNG-qRTPCR-F-LL[b] | TCTCCCTTGCCTTTATGGTG |
| UNG-qRTPCR-R-LL[b] | CACCCCAACATCTGTCACTG |
| PCNA-qRTPCR-F-LL[b] | GTGAACCTCACCAGTATGTC |
| PCNA-qRTPCR-R-LL[b] | CCAAGGTATCCGCGTTATC |
| CCNA2-qRTPCR-F-LL[b] | TAGATGCTGACCCATACCTC |
| CCNA2-qRTPCR-R-LL[b] | GATTCAGGCCAGCTTTGTC |
| CCNB1-qRTPCR-F-LL[b] | GCACCAAATCAGACAGATGG |
| CCNB1-qRTPCR-R-LL[b] | CGACATCAACCTCTCCAATC |
| AURKA-qRTPCR-F-LL[b] | AGGACCTGTTAAGGCTACA |
| AURKA-qRTPCR-R-LL[b] | GAGCCTGGCCACTATTTAC |
| B-Actin-qRTPCR-F-LL[c] | AAGCCACCCCACTTCTCTCTAA |
| B-Actin-qRTPCR-R-LL[c] | AATGCTATCACCTCCCCTGTGT |
| YWHAZ-qRTPCR-F-LL[d] | TCTGTCTTGTCACCAACCATTCTT |
| YWHAZ-qRTPCR-R-LL[d] | TCATGCGGCCTTTTTCCA |
| RELN-qRTPCR-F[e] | TGAGAGCCAGCCTACAGGA |
| RELN-qRTPCR-R[e] | TCGTTCCACATTCTGTACCAA |
| FMR1-qRTPCR-F[f] | TGGCTTCATCAGTTGTAGCAGG |
| FMR1-qRTPCR-R[f] | TCTCTCCAAACGCAACTGGTC |
| UBE3A-qRTPCR-F | TCAAGGCTTTTCGGAGAGGT |
| UBE3A-qRTPCR-R | TCCCTGGTATAGCCACCGTC |
| NOTCH1-qRTPCR-F | GGACGTCAGACTTGGCTCAG |
| NOTCH1-qRTPCR-R | ACATCTTGGGACGCATCTGG |
| NOTCH2-qRTPCR-F | ATCCCACAAAGCCTAGCACC |
| NOTCH2-qRTPCR-R | CCTTGTCCCTGAGCAACCAT |

| | |
|---|---|
| ARNT2-qRTPCR-F | CTCCTTCGGCCTCTTGTGTT |
| ARNT2-qRTPCR-R | GAGGGTAGGTCCCTCCATGT |

[a] Previously characterized in (Vandesompele et al. 2002).
[b] Previously characterized in (Aviner et al. 2015).
[c] Previously characterized in (Adesida et al. 2006)
[d] Generous gift from Dr A. Adesida
[e] Previously characterized in (Baek et al. 2015)
[f] Previously characterized in (Pretto et al. 2015)

#### 2.2.29. Statistics: molecular and cellular data

Statistical analysis of qPCR results was performed by the BioRad CFX manager software, all other statistical analysis was performed in Excel. For luciferase analysis, groups were compared by Student's *t*-tests, after running an *F*-test to compare relative variances. For qPCR graphs, error bars depict the standard deviation.

## 2.3 Assessment of endogenous NPAS3 function

### 2.3.1. One-step RT-PCR

For all qPCR experiments, two-step cDNA synthesis was performed, as described above. One-step RT-PCR was performed only for initial characterization of SK-N-SH cells as follows: The SuperScript III One-step RT-PCR kit (Invitrogen) was used to detect expression of *NPAS3* in SK-N-SH. In a 25 µL reaction the 2X complete reaction buffer was diluted to 1X, 10 µg of each primer was used (*NPAS3*-detect-F: TCAGGCACAGTCACTTGGAC and *NPAS3*-detect-R: AGATGCTGCTGTCATTGTCG, 861 bp product) with a range of 0-5% DMSO and 100 ng of SK-N-SH RNA was used as template, with water to make up the remainder of the

reaction volume. Cycling conditions were as follows: first the RT reaction was incubated at 60°C for 30 minutes, followed by 94°C denaturation. The PCR reaction followed immediately, with 40 cycles of 94°C denaturation for 15 seconds, 55.1°C annealing for 30 seconds, and 68°C extension for 1 minute. The reaction was completed with a 5 minute incubation at 68°C. PCR products were run on a 1% agarose 1X TAE gel. The expected 861 bp band was gel extracted with the QIAquick Gel Extraction kit (QIAGEN), eluted in elution buffer and submitted for sequencing with the *NPAS3*-detect F and R primers to confirm that the detected band is *NPAS3*.

### 2.3.2. Western blotting: chemiluminescent detection

This protocol was only used for initial characterization of SK-N-SH as an appropriate cell line for analysis of NPAS3 function. All other western blots were performed according to the above described LICOR protocol.

SK-N-SH cells were trypsinized per the usual protocol and the 5 mL of released cells in EMEM + 10% FBS were collected in a 15 mL conical tube at 1000×g, and washed twice with PBS. Cells were lysed in 500 µL of RIPA (50 mM Tris HCl pH 7.4, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 0.5% sodium deoxycholate) supplemented with 1X Mammalian Inhibitory Cocktail (Sigma) and 1 mM PMSF. Samples were homogenized by 20 passes through a 22 gauge needle and incubated on ice for 15 minutes, and for 20 minutes with rotation at 4°C. Insoluble material was pelleted by a 10 minute spin at 10 000×g at 4°C. The supernatant was quantified using the Coomassie Plus Protein Assay Reagent (Pierce) using a fresh standard curve made from 0-2 mg/mL BSA for interpolation of sample protein concentration. Absorbance at 595 nm was read

on a spectrophotometer. Samples were diluted to 30 µg in a final 1X Laemmli SDS-PAGE sample buffer (4% SDS, 20% glycerol, 120 mM Tris-HCl pH 6.8, 100 mM DTT, 0.01% (w/v) bromophenol blue), denatured for 5 minutes at 95°C and snap cooled on ice.

Samples were run on a discontinuous western blot with a 4% stacking gel (125 mM Tris-HCl pH 6.8, 0.1% SDS) and 6.5% separating polyacrylamide gel (375 mM Tris-HCl pH 8.8, 0.1% SDS) at 200 V for approximately 45 minutes in 1X SDS PAGE buffer (25 mM Tris-HCl, 192 mM glycine, 0.1% SDS) until the dye front reached the bottom of the gel. On each gel, 5 µL of PrecisionPlus Protein Kaleidoscope Prestained Protein Standard (BioRad) was loaded for a molecular weight marker. Proteins were transferred overnight onto nitrocellulose membranes in Towbin transfer buffer (15% methanol, 25 mM Tris-HCl, 192 mM glycine, 0.01% SDS) at 4°C. Membranes were blocked for 3 hours with 4% milk powder (w/v) in TBS-T (1X TBS 0.1% Tween-20), washed twice and probed overnight at 4°C with primary antibody (αGAPDH 1/10 000, αNPAS3 N-ter (ProSci) 1/1000) diluted in 4% milk power (w/v) in TBS-T (1X TBS 0.1% Tween-20). Blots were washed four times in TBS-T wash buffer and probed with secondary antibody (goat-anti-rabbit-horseradish peroxidase (Pierce chemical), goat-anti-mouse-horseradish peroxidase (Pierce chemical)) diluted in 4% milk powder (w/v) in TBS-T for 1 hour at room temperature. The blot was washed four times in TBS-T and chemiluminescence was assessed using the SuperSignal West Femto Max Sensitivity Substrate (Pierce chemical) as follows: luminol was diluted 1:1 with the peroxide reagent and applied to the dried nitrocellulose blot, wrapped in Saran wrap and incubated with X-ray film (Fuji) in an X-ray cassette for 1 minute (GAPDH) to 1 hour (NPAS3).

### 2.3.3. Cell culture: SK-N-SH and human neuroprogenitor cells

SK-N-SH cells were maintained in 75 $cm^2$ TC treated vented flasks (Sarstedt) with 10 mL of Minimum Essential Media-Eagle (EMEM, Sigma M0643) supplemented with 2.2 g/L sodium bicarbonate, pH adjusted to 7.3, filter sterilized and supplemented with 10% FBS. Cells were passaged every 4-7 days at 1:2 to 1:3. For subculturing, media was aspirated, cells were washed once with 1X PBS, then treated with trypsin EDTA at 37°C for 2-10 minutes until released from the culture flask. Cells were then rescued with 5-10 mL of media supplemented with 10% FBS and subcultured at the indicated dilutions or plated for experiments.

For experimentation, cells were plated at the following densities: $2.20\times10^6$ cells per 10 cm plate, $7.6\times10^5$ cells per 6 cm plate, $3.0\times10^5$ cells per 3.5 cm dish or well of 6-well plate or $5\times10^4$ cells per well of a 24-well plate.

ENStem-A human neuroprogenitor cells were purchased as a part of the human neuroprogenitor cell kit from Millipore and handled as per the kit instructions. Cells were plated on poly-L-ornithine and laminin coated TC treated dishes and maintained in ENStem-A neural expansion media (Neurobasal media plus B27 supplement and LIF) supplemented with 2 mM L-glutamine, penicillin/streptomycin and 20 µg/mL FGF-2. Media was replaced at least every 48 hours. When cells were 90-100% confluent, they were manually dissociated in fresh medium, pelleted, resuspended in fresh medium supplemented with FGF-2, and subcultured at 1:2.

### 2.3.4. Induction of circadian cycling

Circadian cycling was induced using a serum shock protocol described previously (Balsalobre, Damiola, Schibler 1998). Briefly, SK-N-SH cells were grown to 90% confluence and then shocked with 50% horse serum 50% EMEM for 2 hours (ZT 0-2), then washed twice with PBS, released into EMEM 10% FBS, and harvested at various timepoints for 48 hours for RNA analysis of markers of circadian cycling.

### 2.3.5. Hypoxia experiments

To assess the effects of mild hypoxia, cells were plated per the normal protocol, allowed to grow in normoxia (21% $O_2$, 5% $CO_2$, humidified atmosphere), transfected, and then incubated in a hypoxic incubator (3% $O_2$, 5% $CO_2$, humidified atmosphere) for 24 hours and harvested per the normal protocols. $CoCl_2$ was also used to mimic hypoxic gene regulation driven by stabilization of HIF1A (Yuan et al. 2003). Cells were cultured and transfected as usual and allowed to recover for at least 24 hours before being treated with 50-200 μM $CoCl_2$ for 24 hours and then harvested to assess stabilization of HIF1A and the resulting transcriptional regulation.

### 2.3.6. Growth curves

HEK 293T growth curves were generated as follows: HEK 293T cells were plated as per the usual protocol on 3.5 cm dishes and cultured in 2 mL of DMEM + 10% FBS (low or high glucose). At the indicated time points HEK 293T cells were manually dissociated from the dishes by pipetting into the growth medium and an aliquot was counted using a hemocytometer. Four quadrants were read per time point, unless total

cell number was less than 200, then 8 quadrants were read. Data were plotted in Excel and error bars represent standard deviation.

SK-N-SH growth curves were generated as follows: SK-N-SH cells were plated as per the usual protocol on 3.5 cm dishes and cultured in 2 mL of EMEM + 10% FBS. Media was either not replaced, or replaced every 48 hours. As SK-N-SH are highly adherent and resistant to manual dissociation, media was removed, cells were washed once in PBS, and incubated for 5 minutes with 0.5 mL trypsin. Cells were rescued with 1.5 mL of EMEM + 10% FBS, mixed by pipetting and counted and plotted as per the HEK 293T cell protocol.

### 2.3.7. shRNA constructs used

pLKO.1 GFP shRNA was a gift from David Sabatini (Addgene plasmid # 30323) (Sancak et al. 2008). The *NPAS3* specific shRNAs TRC76-78 were from MISSION® TRC-Hs 1.5 (Moffat et al. 2006)(Sigma). *NPAS3* shRNAs 613-616 were also obtained from Origene in a pRFP-C-RS vector. Sequences and full identifiers are listed in Table 6.

## Table 6 Identifiers and sequences of shRNAs used

| Target | shRNA ID | shRNA name | shRNA sequence | Backbone | Source |
|---|---|---|---|---|---|
| eGFP | 300323 | shRNA eGFP | GCAAGCTGACCCTGAAGTTCAT | pLKO.1 | Addgene [a] |
| NPAS3 | TRCN0000020976 | NPAS3 shRNA TRC76 | CCGGCCATCATTCGACTTACAATTACTCGAGTAATTGTAAGTCGAATGATGGTTTTT | pLKO.1 | The RNAi Consortium[b] (Sigma) |
| NPAS3 | TRCN0000020977 | NPAS3 shRNA TRC77 | CCGGGCTGTTAACTTCGTGGACGTTCTCGAGAACGTCCACGAAGTTAACAGCTTTTT | pLKO.1 | The RNAi Consortium[b] (Sigma) |
| NPAS3 | TRCN0000020978 | NPAS3 shRNA TRC78 | CCGGGAACCCATCAATTTCGACAATCTCGAGATTGTCGAAATTGATGGGTTCTTTTT | pLKO.1 | The RNAi Consortium[b] (Sigma) |
| scramble | TR30015 | scrambled shRNA | GCACTACCAGACTAACTCAGATAGTACT | pRFP-C-RS | Origene |
| NPAS3 | F1311613 | NPAS3 shRNA 613 | TCCTCGGAGACATCCGACTCTGAGTCAGA | pRFP-C-RS | Origene |
| NPAS3 | F1311614 | NPAS3 shRNA 614 | CATTTCCGAAACAGTCTCCATCTACCTAG | pRFP-C-RS | Origene |
| NPAS3 | F1311615 | NPAS3 shRNA 615 | ACAGCTTCGAGCACTCGGACTTTGAGAAC | pRFP-C-RS | Origene |
| NPAS3 | F1311616 | NPAS3 shRNA 616 | GGTATTACAGAGGACAACGAGAACTCCAA | pRFP-C-RS | Origene |

[a](Sancak et al. 2008)

[b] (Moffat et al. 2006)

### 2.3.8. Lentiviral particle production

Lentiviruses were produced based on the Addgene protocol available online (Addgene 2006) with some modification. HEK 293T cells were plated on a 6 cm tissue culture plate at $7.6\times10^5$ cells per plate and cultured with 5 mL of DMEM + 10% FBS. The day after plating the cells were transfected using the Mirus TransIT express transfection reagent protocol, where 250 µL of serum-free DMEM was used to mix 1 µg of pLKO.1 shRNA plasmid (Table 6) or pLIX_402_*NPAS3*-933 plasmid as well as 750 ng of psPAX2 packaging plasmid, and 250 ng of pMD2.G envelope plasmid. To this 6 µL of Mirus TransIT express reagent (MirusBio) was added and the transfection mixture was allowed to sit for 30 minutes at room temperature. The transfection mixture was added dropwise to the plated cells, plates were swirled gently to mix and then returned to the incubator. 15 hours post transfection the media was changed to fresh DMEM + 10% FBS, and returned to the incubator. On the second and third day after transfection the viral particle containing media was collected from the transfected cells and stored at 4°C. The pooled media was pelleted at 300×g and the cell-free supernatant was stored for transduction of target cells. For immediate usage, the viral mix was stored at 4°C, however several 1 mL aliquots were stored at -20°C for longer preservation.

### 2.3.9. Transduction of SK-N-SH

Optimal puromycin concentration for selection was performed on SK-N-SH plated at standard density, and allowed to grow to 80% confluence before replacement of media with a range of puromycin concentrations with replacement every 2 days. Cell

viability was assessed every day for 5 days, and 3 μg/mL was found to be the minimum concentration to kill 100% of cells. For transduction $2.2\times10^6$ SK-N-SH cells were plated on 10 cm plates and allowed to grow for 48 hours, until they reached 60-80% confluence. The media on these plates was changed to EMEM + 10% FBS supplemented with 8 μg/mL polybrene and then 1 mL of viral particle containing DMEM + 10% FBS was added dropwise to the plates and then the plates were returned to the incubator. Media was replaced 24 hours post transduction and cells allowed to grow another 24 hours (total 48 hours) when they were found to be fully confluent and were trypsinized, subcultured 1:3 in EMEM + 10% FBS supplemented with 3 μg of puromycin to select for transduced cells. Media was replaced every 4 days, or when cells required subculture, whichever came first. For shRNA expressing constructs, after the first subculture the ratio of splitting was reduced to 1:2 to enhance cell growth. The inducible pLIX_402-NPAS3-933 construct was subcultured at a ratio of 1:3 as growth remained stable. After 10 days of selection the puromycin dosage was decreased to 1.5 μg for maintenance. For experiments, cells were plated at standard densities indicated.

### 2.3.10. Generation of stably expressing constructs from transfections

For the shRNA constructs obtained from Origene, SK-N-SH cells were plated on 10 cm plates and allowed to grow to 60% confluence (48 hours) before being transfected with 4 μg of shRNA expressing plasmid, per the standard transfection protocol. Cells were allowed to recover for 48 hours and then media was replaced with fresh EMEM + 10% FBS supplemented with 3 μg/mL puromycin. Media was replaced every four days and cells allowed to expand into colonies for two weeks before being trypsinized, pooled

and expanded into a fresh flask. Selection was reduced to 1.5 μg/mL puromycin after two weeks and cells were subcultured at 1:2. For experiments cells were plated at the standard densities indicated.

### 2.3.11. Cytoplasmic/nuclear fractionation

When cells were ready for harvest, media was removed from dishes, cells were washed twice in 1X PBS before being collected in 1 mL of 1X PBS in a microcentrifuge tube, pelleted by centrifugation at 1000×g for 5 minutes at 4°C and then frozen at -80°C. Cells were thawed and resuspended in cytoplasmic lysis buffer (5 mM Tris pH 7.5, 85 mM KCl, 1% IGEPAL) with freshly added protease inhibitor cocktail (Sigma P8340), 1 mM PMSF, 1 mM DTT. HEK 293T cells were lysed in 200 μL per 10 cm dish, SK-N-SH cells were lysed in 100 μL per 10 cm dish. Cells were incubated in lysis buffer for 20 minutes on ice then homogenized by 25 passes of the B-pestle in a 2 mL Dounce homogenizer. Nuclei were pelleted by centrifugation at 3000×g for 5 minutes at 4°C, the supernatant was collected and saved (cytoplasmic fraction).

Nuclei were resuspended in Mammalian Lysis Buffer (see total protein purification section for constitution) with freshly added protease inhibitor cocktail (Sigma P8340), 1 mM PMSF, 1 mM DTT. 100 μL was used per 10 cm plate of HEK 293T nuclei, and 40 μL per 10 cm plate of SK-N-SH nuclei. Nuclei were incubated for 15 minutes on ice before being sonicated using the Sonic Dismembrator ultrasonic processor (Fisher FB-120) at 30% amplitude for 5 seconds on, 5 seconds off. For 40 μL volumes 3 cycles of sonication resulted in adequate lysis, for 100 μL volumes 6 cycles was required. Nuclear lysates were centrifuged at 14 000×g for 5 minutes at 4°C and the soluble lysate

(supernatant) was collected. Lysates were stored as described in the total protein purification section and western blots were run as described in the Licor western blot protocol.

# Chapter 3: Exonic Variation of *NPAS3* is associated with cognitive variation

## 3.1 Rationale

Although variation linked to and affecting *NPAS3* has been robustly associated with neuropsychiatric and neurodevelopmental disorders with common features of intellectual disability, the direct contribution of *NPAS3* to these disorders is unknown (Gonzalez-Penas et al. 2015; Kamnasaran et al. 2003; Kamnasaran et al. 2005; Pickard et al. 2008; Visser et al. 2010; Weber et al. 2011; Yu et al. 2014). Furthermore, as variants to *NPAS3* are present in the worldwide population, identification of the contribution of these variants to variation in normal intellectual function will provide insight into the role of *NPAS3* variation in normal and abnormal neurodevelopment and neuropsychological function (Lek et al. 2016; National Center for Biotechnology Information, National Library of Medicine 2017). To this end, we set out to genotype three coding *NPAS3* SNPs previously associated with schizophrenia (Macintyre et al. 2010) in a cross-sectional population of western Canadian high school students and young adults of pre-onset age for most major psychiatric disorders and assess cognitive function relative to *NPAS3* genotype. *COMT* rs4680, encoding p.V158M, a missense variant that affects enzymatic function and has been robustly associated with neuropsychiatric function (for review see (Witte and Floel 2012)), was also genotyped in the study population to address several additional hypotheses in the broader study, not specifically related to *NPAS3*.

## 3.2 Genotyping results

*COMT* rs4680 (c.472G>A, p.V158M) was genotyped using a method developed from a mini RFLP protocol outlined in Malhotra et al. 2002. The protocol was adapted for use on the QIAxcel capillary gel electrophoresis system to simplify calling of genotypes based on 86 bp and 68 bp restriction digest fragments on a PCR product overlapping rs4680. Representative data used to call genotypes is shown in Figure 3-1A. The protocol was used to successfully genotype *COMT* rs4680 in 86 of the 87 samples received. *COMT* rs4680 allele frequencies (G=0.506, A=0.494) were not found to differ from ExAC v0.3.1 allele frequencies, nor did the genotype frequencies vary from Hardy-Weinberg equilibrium (G/G=0.291, G/A=0.430, A/A= 0.279, $\chi^2$=1.24, $p$=0.54) (Lek et al. 2016). Analyses of association between *COMT* genotype and other measures under study as a part of this collaborative study were performed by other collaborators and the data are not presented here.

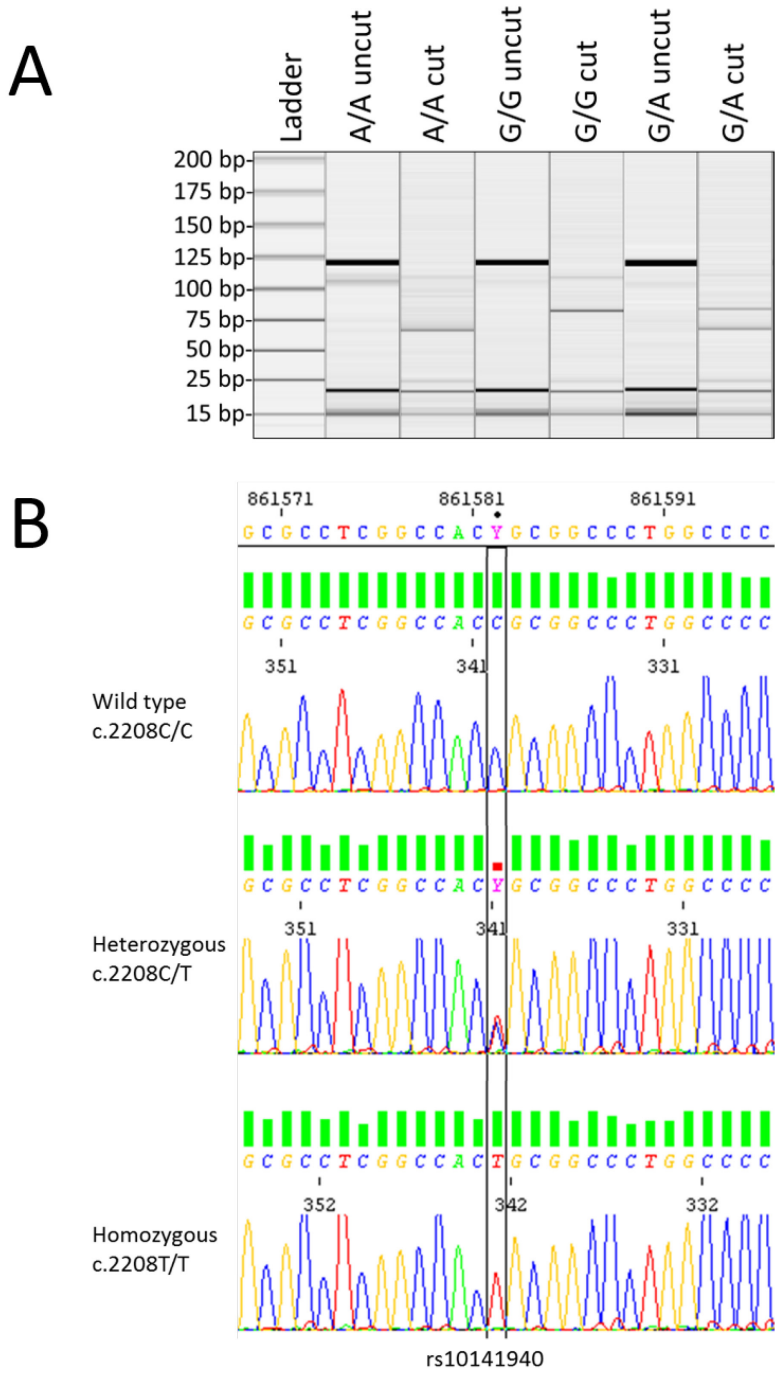



**Figure 3-1 Sample genotyping data for *COMT* and *NPAS3***
(A) QIAxcel capillary gel digital image demonstrating the 86 bp and 68 bp bands resolved for genotyping all three possible rs4680 genotypes. (B) Electropherograms representing all three genotypes of *NPAS3* SNP rs10141940 showing typical sequencing read quality used to genotype the three *NPAS3* SNPs.

Genotyping of *NPAS3* SNPs rs12434716 (c.1654G>C, p.Ala552Pro) and rs10141940 and rs10142034 (c.2208C>T and c.2262C>G, both synonymous) was performed using an established protocol developed in our lab which has previously been published (Macintyre et al. 2010). Sample genotyping data are shown in Figure 3-1B. The genotype for rs12434716 (p.Ala552Pro) was called for all 87 participants that submitted a saliva sample while the genotypes for the two linked SNPs rs10141940 and rs10142034 were able to be called definitively for 80/87 samples. The seven individuals not genotyped were due to a combination of difficulty obtaining a suitable PCR product, potentially due to sample quality, and/or inability to sequence through the highly repetitive GC-rich sequence proximal to these SNPs.

The allele frequencies of the genotyped SNPs (rs12434716: G=0.83, C=0.17; rs10141940 and rs10142034: C;C=0.84, T;G=0.16) did not vary from Hardy-Weinberg equilibrium (rs12434716: $\chi^2$=0.10, $p$=0.75; rs10141940 and rs10142034 $\chi^2$=0.0016, $p$=0.90) (Table 7). Comparison to the ExAC v0.3.1 release found that the population allele frequencies of all three SNPs did not vary from the European population frequencies (Lek et al. 2016). Remarkably, we also found that the allele frequencies of all three SNPs were not significantly different from those observed in a schizophrenia patient cohort (Macintyre et al. 2010). Similar to the previous study, we found these SNPs to co-segregate as a haplotype, likely owing to proximity. Recombination between the rs12434716 and rs10141940 (554 bp apart) SNPs was observed in two individuals (2.23%). No recombination was observed between rs10141940 and rs10142034 (54 bp apart).

**Table 7: Genotype and allele frequencies of the assessed SNPs relative to exome sequencing data of Europeans accessed from ExAC v0.3.1, (Lek et al. 2016) and previous healthy control (HC) and schizophrenia patient (SZ) samples (Macintyre et al. 2010).**

| dbSNP ID | Variant[a] | *N* | Genotype | | | Previous study HC | Previous study SZ | Allele | | Previous study HC | Previous study SZ | ExAC v0.3.1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **rs12434716** | c.1654g>c, p.Ala552Pro | 87 | G/G 68.9% *n*=60 | G/C 28.7% *n*=25 | C/C 2.30% *n*=2 | FE[b] *p*=0.092 | $\chi^2(2)=1.41$ *p*=0.288 | G 83.3% *n*=145 | C 16.7% *n*=29 | $\chi^2(1)=9.87$ *p*=0.002** | $\chi^2(1)=0.02$ *p*=0.288 | $\chi^2(1)=0.02$ *p*=0.879 |
| **rs10141940** | c.2208c>t | 80 | C/C 71.3% *n*=57 | C/T 26.3% *n*=21 | T/T 2.50% *n*=2 | FE[b] *p*=0.218 | $\chi^2(2)=2.24$ *p*=0.327 | C 84.4% *n*=135 | T 15.6% *n*=25 | $\chi^2(1)=5.05$ *p*=0.025* | $\chi^2(1)=0.02$ *p*=0.128 | $\chi^2(1)=0.11$ *p*=0.735 |
| **rs10142034** | c.2262c>g | 80 | C/C 71.3% *n*=57 | C/G 26.3% *n*=21 | G/G 2.50% *n*=2 | FE[b] *p*=0.218 | $\chi^2(2)=2.24$ *p*=0.327 | C 84.4% *n*=135 | G 15.6% *n*=25 | $\chi^2(1)=5.05$ *p*=0.025* | $\chi^2(1)=0.02$ *p*=0.128 | $\chi^2(1)=0.89$ *p*=0.346 |

[a] Variants are numbered based on NCBI transcript NM_001164749.1

[b] Fisher's Exact test (In-silico 2016)

* $p < 0.05$, ** $p < 0.01$

## 3.3 Sample characteristics

The study sample assessed here was found to be primarily European (66/87, 75.9%), with the remainder composed of Asian (7/87, 8.0%), Métis (5/87, 5.7%), and unknown (9/87, 10.3%) ethnicities. The average age of the population was 17.68 years (*SD*=1.49 years). The population was mostly female (68/87, 78.2%) and had an average estimated premorbid intellect (WRAT4 Reading *M*=102.07, *SD*=11.86).

## 3.4 *NPAS3* genotype is correlated with performance on a working memory task

Initial assessment of the differences in cognitive test scores by *NPAS3* genotype was performed using independent *t*-tests between homozygotes of the major allele as compared to heterozygotes (Table 8 and Table 9). Homozygotes for the minor allele were excluded from this analysis due to small sample size (*n*=2). No differences were observed between genotypes for premorbid intellect estimated by the WRAT4 reading subtest score, sustained attention on the CPT-0X or non-verbal working memory on the VPT. The general cognitive index of the SCIP (SCIP_GCI) was found to trend toward decreased cognitive function associated with the minor allele of the coding SNP rs12434716 (SCIP_GCI, $t(66.8)=1.87$, $p=0.066$; adjusted $R^2=0.018$, $F(1,83)=2.53$ $p=0.12$). This appears to be related to lower scores on the verbal working memory subtest of the SCIP in individuals heterozygous for the minor alleles of the *NPAS3* SNPs tested (for rs12434716: adjusted $R^2=0.094$, $F(1,83)=8.576$, $p=0.0044$; for rs14141940 and rs10142034 adjusted $R^2=0.08895$, $F(1,76)=8.518$, $p=0.0046$). No differences were

observed between *NPAS3* genotypes for SCIP subtests of verbal list learning or delayed recall, verbal fluency, memory, or processing speed.

## 3.5 Summary of Findings

Through this study I have successfully genotyped three SNPs of *NPAS3* in a cross-sectional cohort of western Canadian students and found that the genotype and allele frequencies observed do not differ from those observed in the worldwide population. *NPAS3* genotype was found to be associated with reduced performance on a working memory task. This finding represents the first neuropsychological outcome that is specifically associated with *NPAS3* genotype, providing insight into how variation in *NPAS3* affects the psychological function of individuals, both healthy and those with neuropsychiatric or neurodevelopmental disorders, that carry these variants.

**Table 8: Cognitive test scores relative to *NPAS3* coding SNP rs12434716 genotype**

| | rs12434716 genotype | | | |
|---|---|---|---|---|
| **Cognitive test** | **G/G *n*=60** | **G/C *n*=25** | **C/C *n*=2** | ***N*=87** |
| WRAT4_R[a] | 102.33±1.65 | 101.68±2.01 | 99.00±3.00 | 102.07±1.17 |
| CPT-D[ab] | 0.83±0.16 | 0.61±0.08 | 0.87±0.01 | 0.77±0.09 |
| VPT[a] | 12.10±0.20 | 12.12±0.37 | 11.00±1.00 | 12.08±0.42 |
| VLT-I[c] | 23.32±0.45 | 23.04±0.56 | 24.00±3.00 | 23.25±0.53 |
| VLT-D[c] | 7.93±0.26 | 7.60±0.27 | 7.00±1.00 | 7.82±0.34 |
| WMT[c] | 19.95±0.34 | 18.24±0.38** | 21.50±0.50 | 19.49±0.51 |
| VFT[c] | 16.02±0.52 | 15.20±0.65 | 22.50±0.50 | 15.93±0.28 |
| VMT[c] | 12.53±0.27 | 12.40±0.40 | 13.00±1.00 | 12.51±0.39 |
| SCIP GCI[c] | 79.75±1.22 | 76.48±1.25 | 88.00±1.00 | 79.00±0.94 |

[a] WRAT4_R: Wide Range Achievement Test, reading subtest, CPT-D: Connors Continuous Performance Test, VPT: Visual Patterns Test,
[b] Conner's CPT was not completed by one G/G individual
[c] Screen for Cognitive Impairment in Psychiatry (SCIP) subtests: VLT_I: verbal list learning-immediate recall, VLT_D: verbal list learning-delayed recall, WMT: verbal working memory, VFT: verbal fluency, VMT: visuomotor tracking, SCIP_GCI: general cognitive index score of all SCIP subtests
** G/G vs G/C, $p < 0.01$.

**Table 9: Cognitive test scores relative to *NPAS3* exonic SNP rs10141940 and rs10142034 genotypes**

| | rs10141940 and rs10142034 genotype | | | |
|---|---|---|---|---|
| **Cognitive test** | **C/C;C/C n=57** | **C/T;C/G n=21** | **T/T;G/G n=2** | ***N*=80** |
| WRAT4_R[a] | 102.23±1.77 | 103.48±2.01 | 99.00±3.00 | 102.48±1.22 |
| CPT-D[a] | 0.72±0.12 | 0.57±0.10 | 0.87±0.01 | 0.68±0.09 |
| VPT[a] | 11.96±0.20 | 12.00±0.41 | 11.00±1.00 | 11.95±0.43 |
| VLT-I[b] | 22.95±0.45 | 23.14±0.63 | 24.00±3.00 | 23.03±0.55 |
| VLT-D[b] | 7.84±0.27 | 7.57±0.30 | 7.00±1.00 | 7.75±0.35 |
| WMT[b] | 19.74±0.34 | 17.95±0.37** | 21.50±0.50 | 19.31±0.54 |
| VFT[b] | 15.95±0.55 | 15.43±0.75 | 22.50±0.50 | 15.98±0.30 |
| VMT[b] | 12.42±0.28 | 12.57±0.47 | 13.00±1.00 | 12.48±0.40 |
| SCIP GCI[b] | 78.89±1.26 | 76.67±1.38 | 88.00±1.00 | 78.54±0.98 |

[a] WRAT4_R: Wide Range Achievement Test, reading subtest, CPT-D: Connors Continuous Performance Test, VPT: Visual Patterns Test
[b] Screen for Cognitive Impairment in Psychiatry (SCIP) subtests: VLT_I: verbal list learning-immediate recall, VLT_D: verbal list learning-delayed recall, WMT: verbal working memory, VFT: verbal fluency, VMT: visuomotor tracking, SCIP_GCI: general cognitive index score of all SCIP subtests
** C/C;C/C vs C/T; C/G, $p < 0.01$.

# Chapter 4: NPAS3 and ARNT interact to regulate target genes

## 4.1 Rationale

NPAS3, as a bHLH-PAS transcription factor family member, is predicted to interact with ARNT as an obligate heterodimer (Brunskill et al. 1999; Pieper et al. 2005). NPAS3 has three predicted domains, common to bHLH-PAS transcription factors: a bHLH domain, shown to play a role in DNA binding and protein::protein interactions, tandem PAS domains, involved in PAS protein interaction and environmental signalling, and the C-terminus, predicted to be a transactivation domain (Brunskill et al. 1999). None of these domains had been functionally characterized. In order to characterize the function of NPAS3, I assessed whether NPAS3 can interact with ARNT in human cells and, if an interaction is detected, to identify which domains are critical for this interaction. Furthermore, I set out to assess whether NPAS3 is able to regulate previously identified NPAS3-regulated genes, if ARNT is involved in this regulation and whether the observed regulation is direct. Deletion constructs of NPAS3 were developed to characterize the function of NPAS3 domains relative to their predicted roles. Finally, three variants of *NPAS3* identified in humans were introduced into *NPAS3* clones to assess their relative function in assays developed to test wild-type NPAS3.

## 4.2 Development of an NPAS3 antibody in guinea pigs

Commercially available αNPAS3 antibodies were found to be unsuitable for functional assessment of NPAS3, due to the presence of secondary bands of sizes not predicted by the major transcripts of *NPAS3* indexed on NCBI. Because of these factors, generation of a custom NPAS3 antibody was undertaken in guinea pigs. As no cell line known not to express NPAS3 (e.g. cell lines with deletion of *NPAS3*) was available, expression of tagged constructs were used for assessment of antibody suitability. Three guinea pigs were injected with the NPAS3 oligopeptide indicated in Figure 2-1 and serum was used to identify NPAS3 immunoreactivity in one of the three guinea pigs that was selected for further analysis (data not shown). HA-tagged NPAS3 isoforms were expressed in U2OS human osteosarcoma cell lines and protein lysates were collected. NPAS3 antibody, or NPAS3 immunoreactive serum, and HA-tag antibodies were used on a western blot of 50 μg of total cell lysate. An approximately 130 kDa band was detected for both NPAS3 isoforms that was not present in the untransfected cells (Figure 4-1). Analysis using the LI-COR Odyssey scanner, comparing two wavelengths, allowed for probing with the novel NPAS3 antibody and HA antibody at one time with secondary antibodies to guinea pig and rabbit, respectively, conjugated to IRdye-800 nm (green) and AlexaFluor-680 nm (red). Direct comparison of the bands detected by the two antibodies was observed by a yellow signal where fluors overlap (Figure 4-1A). The antibody was found to specifically detect the HA-tagged NPAS3 protein. Furthermore, the NPAS3 antigenic serum was used to probe a blot along side commercial NPAS3 and NPAS1 antibodies, showing that the signal detected by the custom antibody was not from NPAS1, and a 100 kDa band was found to be co-detected by the commercial NPAS3

antibody (Figure 4-1B,C). As these initial screens were positive for detection of NPAS3, investigation of whether the antibody can detect native NPAS3 *in situ* was undertaken.

Immunoprecipitation of expressed HA-NPAS3 isoforms 1 (933 aa) and 2 (901 aa, formerly isoform 1) was performed in U2OS cells. Both an HA-antibody and the NPAS3 antibody were compared to confirm that the immunoprecipitated band is the tagged clone. The antibody immunoprecipitated the same size band as the input, demonstrating the antibody's ability to interact with native protein (Figure 4-1D). Finally, immunofluorescence using the NPAS3 antibody on SK-N-SH neuroblastoma cells demonstrated nuclear localization that did not co-localize with NPAS1 (Figure 4-1E). Based on these data, the antibody was found to specifically detect NPAS3 and suitable for use in functional assessment of NPAS3.

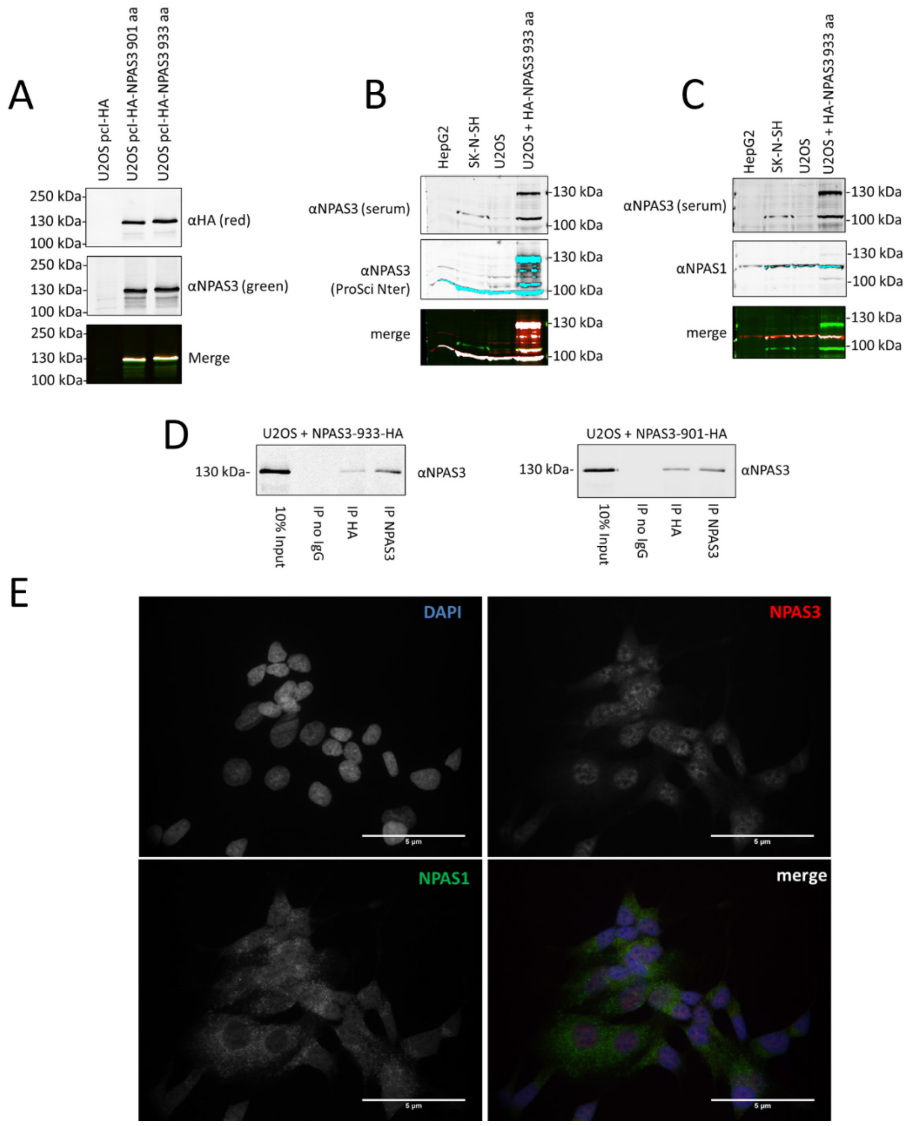



**Figure 4-1: Validation of the NPAS3 antibody**
(A) Western blots showing detection of HA-tagged clones of NPAS3 expressed in U2OS cells. Merge demonstrates the bands detected by the HA and NPAS3 antibodies are the same. (B-C) Western blot of untransfected HepG2 and SK-N-SH as well as U2OS transfected with an empty vector or HA-NPAS3-933 aa expressing construct where the unpurified NPAS3 antigenic serum (green) was used to co-probe the blot with either ProSci NPAS3 N-ter (B) or NPAS1 (C) antibodies (red) to determine whether the expression is specific to NPAS3 and whether it co-probes with a commercial NPAS3 antibody. (D) Western blot of IP demonstrating that the NPAS3 antibody can interact with native HA-NPAS3. (E) Immunofluorescence demonstrating nuclear localization of signal from the NPAS3 antibody (red) that does not co-localize with signal from an NPAS1 antibody (green). 1000X magnification, scale bar = 5 μm.

## 4.3 NPAS3 and ARNT interact

ARNT is presumed to be the obligate heterodimeric partner of NPAS3 (Pieper et al. 2005). In order to determine whether NPAS3 and ARNT interact *in vivo*, the HaloTag system (Promega) was used. The HaloTag is an approximately 34 kDa tag that is enzymatically active and able to catalyze the creation of a covalent bond between the tagged molecule and HaloLink Resin, allowing for high affinity precipitation of complexes and high stringency of washes. HaloTag-ARNT and NPAS3 were co-expressed in HEK 293T cells and total cell lysates were applied to the HaloLink Resin using the Mammalian Pull down kit (Promega) protocol. NPAS3 and ARNT were found to interact robustly (Figure 4-2A). The HaloTag expressed in isolation was not able to pull down NPAS3, confirming that the interaction was specific to ARNT.

As the HaloTag-ARNT construct is of a splice isoform with a two amino acid deletion relative to isoform 1 (this variation is C-terminal to the bHLH and PAS domains known to be involved in bHLH-PAS protein dimerization and not expected to affect interaction) the interaction was confirmed using an independent clone of ARNT isoform 1. HaloTag-NPAS3 (both the 933 aa isoform 1 and 901 aa isoform 2) were generated and used to pull down co-expressed ARNT isoform 1, confirming that NPAS3 and ARNT interact when co-expressed (Figure 4-2B). As ARNT was detected as endogenously expressed in HEK 293T cells, HaloTag-NPAS3 was singly expressed and found to be able to interact with endogenous ARNT, supporting the specificity of this interaction (Figure 4-2C). When HaloTag-ARNT was expressed singly, endogenous NPAS3 expression was not detected in the input (data not shown). As such the reciprocal single-expression endogenous pull down could not be interpreted.

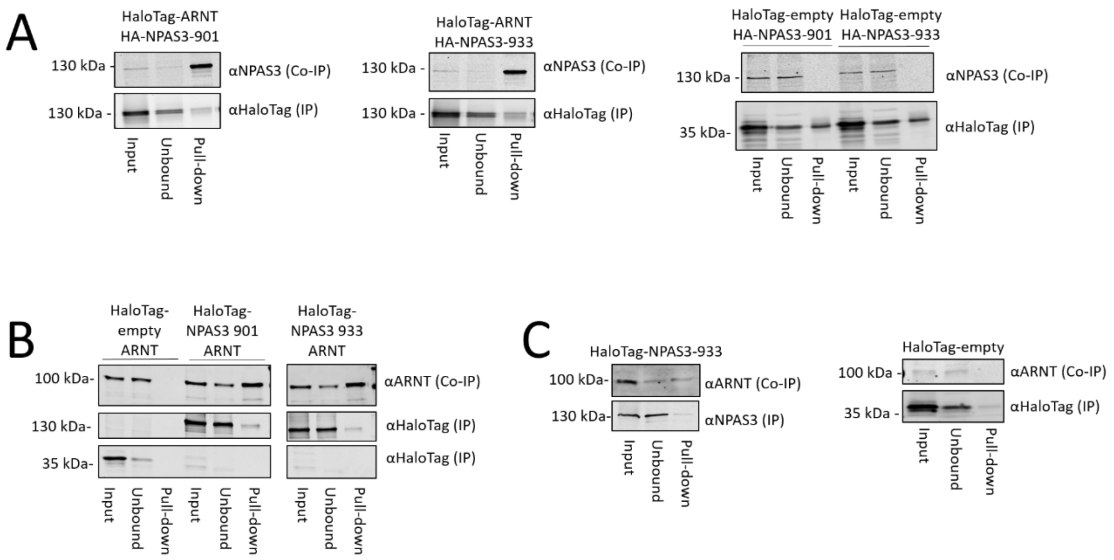



**Figure 4-2: NPAS3 and ARNT physically interact.**
Western blots of HaloTag pull downs demonstrating that (A) HaloTag-ARNT interacts with co-expressed NPAS3, (B) HaloTag-NPAS3, both 933 aa and 901 aa isoforms, are able to interact with ARNT isoform 1 and (C) HaloTag-NPAS3 is able to interact with endogenous ARNT.

## 4.4 NPAS3 and ARNT co-regulate common targets

As NPAS3 and ARNT were found to directly interact, potential regulatory targets of NPAS3::ARNT were identified through a screen of known targets of both. Direct targets of NPAS3 had not been detected at the time of this study. A microarray study assessing differential regulation of genes in response to expression of NPAS3 in HEK 293 cells has been published previously, however whether the observed regulatory effect is direct is unknown (Sha et al. 2012). ARNT ChIP-seq data was published from K562 cell lines by the ENCODE project phase 3 (experiment ID ENCSR155KHM) (Rosenbloom et al. 2013), demonstrating direct targets. These unbiased genome-wide and

transcriptomic data were compared to identify potential direct targets of the NPAS3::ARNT heterodimer.

Encode ChIP-seq data for ARNT was viewed on the UCSC genome browser and the top 50 genes found previously to be up- and down-regulated by NPAS3 expression were screened for ARNT ChIP-seq peaks with the following criteria: peaks must be present in both replicate ChIP-seq experiments and identified in pooled, conservative and optimal peak calling algorithms, resulting in a score out of 5 (Kent et al. 2002; Rosenbloom et al. 2013; Rosenbloom et al. 2015). Thirteen genes were identified from this screen and are listed in Table 10. *VGF* was added to the screen despite the peak in its promoter not reaching significance in either conservative or optimal peak calling algorithms (score 3/5), due to the deeper characterization performed in the original microarray paper and neuronally-relevant functions (Sha et al. 2012).

**Table 10: Potential NPAS3 and ARNT co-targets screened in this study.**

| Gene name | ENCODE ARNT ChIP-seq peaks (X/5) | Fold regulation by NPAS3 (Sha et al. 2012) | Other identifiers | Functions |
|---|---|---|---|---|
| *VGF* | 3 | 2.92 | | Dendritic arborization, proneurogenic (Severini et al. 2008; Zhang et al. 2013a) |
| *HIST1H4H* | 5 | 2.05 | Histone H4 | G1-S transition, replication licencing factors (Saade et al. 2009) |
| *ATF5* | 5 | 1.83 | | Cell cycle progression G1-S, stress response, cAMP signalling, SVZ neurogenesis (Ishihara et al. 2015; Torres-Peraza et al. 2013) |
| *DHCR24* | 5 | 1.75 | Seladin-1 | Neuroprotective, oxidative stress, inflammation (Hernandez-Jimenez et al. 2016) |
| *ZBTB40* | 5 | 1.74 | | Unknown |
| *NCLN* | 5 | 1.70 | Nicalin | Nicalin-NOMO complex, nodal signalling (Haffner et al. 2007) |
| *MAT2A* | 5 | 1.67 | | S-adenosylmethionine synthesis, hypoxia (Liu et al. 2011) |
| *USP49* | 5 | -2.10 | | Deubiquitinase (Zhang et al. 2013c) |
| *ANG* | 5 | -2.16 | RNASE5 | Alternate ORF transcribed from *RNASE4* (Li et al. 2013) |
| *ZNF581* | 5 | -2.24 | | Unknown |
| *RPL37* | 5 | -2.33 | | p53 pathway, MDM2, MDMX (Daftuar et al. 2013) |
| *RNASE4* | 5 | -2.40 | | Neuroprotective during oxidative stress (Li et al. 2013) |
| *TXNIP* | 5 | -2.57 | | Oxidative stress response, inflammation (Yoshihara et al. 2014) |
| *ANKRD37* | 5 | -4.49 | | Hypoxia (Benita et al. 2009; Galbraith et al. 2013) |

In order to confirm that these genes are regulated by NPAS3 and/or ARNT, NPAS3 and ARNT were expressed in HEK 293T cells and RNA was collected. The experiment was conducted with media replacement 24 hours prior to harvest and also where media was not replaced, in order to assess whether the observed gene regulation was related to culture conditions, such as depletion of nutrients in media or serum. qPCR was used to quantify the expression of the potential co-target genes listed in Table 10. Of the 14 genes screened only two were identified as differentially regulated, *TXNIP* and *VGF* (Figure 4-3). *VGF* was identified as being up-regulated by co-expression of *NPAS3* and *ARNT*, consistent with the regulation observed in the original microarray study. *TXNIP* was identified as being down-regulated by expression of *NPAS3* and *ARNT*, but only in cells without media exchange 24 hours prior to harvest. Interestingly, I did not observe any differential regulation of genes in cells singly transfected with *NPAS3*, contrary to what was observed in the original study. These data demonstrate that NPAS3 and ARNT can cooperate to affect the regulation of two genes, which provide a means for assessment of the mechanism by which NPAS3 and ARNT affect gene regulation. All further experiments were performed with media unchanged to ensure that both *VGF* and *TXNIP* regulation by NPAS3 and ARNT should be observed, unless otherwise indicated.

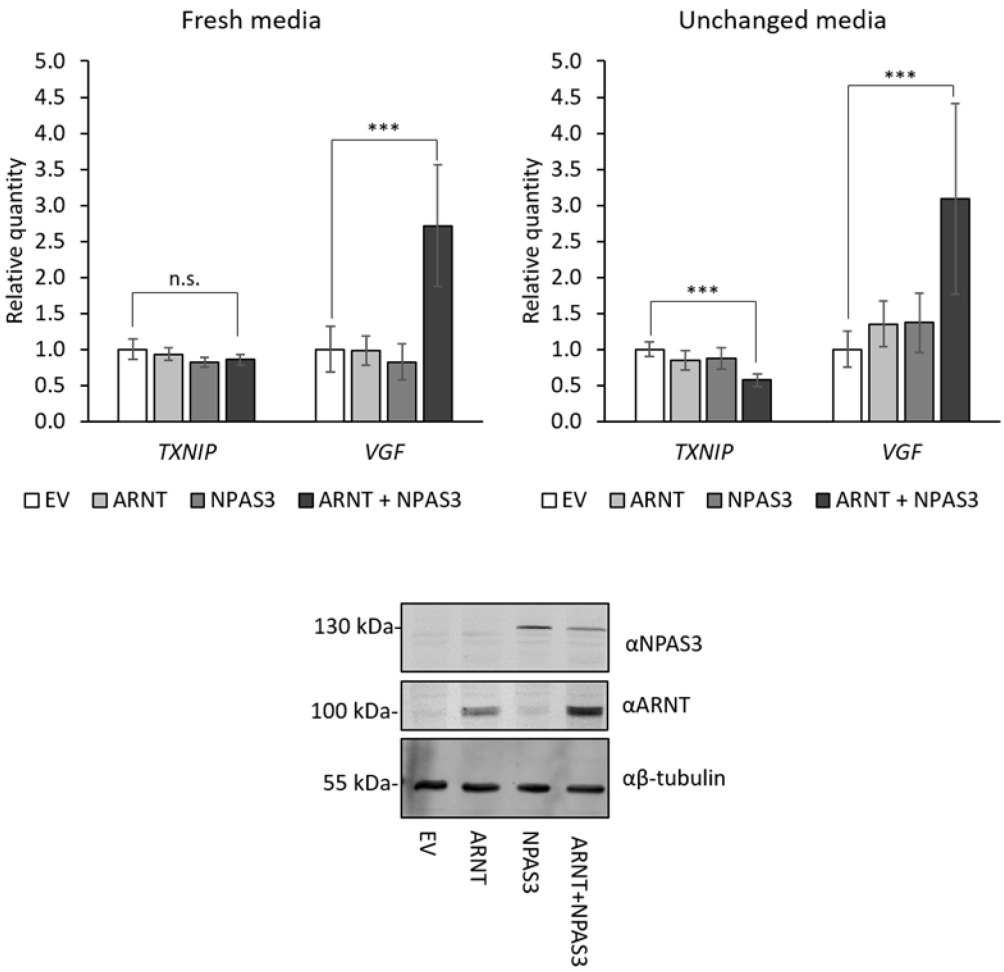



**Figure 4-3: NPAS3 and ARNT regulate expression of *TXNIP* and *VGF*.**
(A) qPCR data showing that NPAS3 and ARNT cause 2.5-3-fold upregulation of *VGF* when expressed in HEK 293T cells, regardless of media condition. *TXNIP* was found to be down-regulated in HEK 293T cells cultured without media exchange after transfection, but not when media was exchanged. (B) Western blot showing expression of NPAS3 and ARNT constructs. *p<0.05, **p<0.01, ***p<0.001

## 4.5 NPAS3 and ARNT bind to regions flanking the *TXNIP* transcription start site

To determine whether the observed upregulation of *TXNIP* by NPAS3 and ARNT is a direct effect, a modified chromatin immunoprecipitation protocol was performed. HaloResin was used to capture sheared crosslinked chromatin of cells expressing HaloTag-NPAS3, in the presence and absence of ARNT, as well as HaloTag-ARNT in the presence and absence of NPAS3. Three regions of *TXNIP* were assessed: the promoter, exon 1 and a region distal to the ARNT ChIP-seq peaks but within the locus (Figure 4-4A). NPAS3 was found to directly bind the promoter region of *TXNIP*, suggesting that the promoter region is the target of NPAS3 (Figure 4-4B). ARNT was found to specifically bind to exon 1, with high background at the promoter precluding conclusions. No specific signal was observed at the distal site where ARNT is not predicted to bind with high affinity, demonstrating that chromatin shearing is efficient and the specificity of signals observed. These data demonstrate that both NPAS3 and ARNT directly bind regions proximal to the transcription start site of *TXNIP*, supporting a role for NPAS3 in the regulation of *TXNIP* expression.

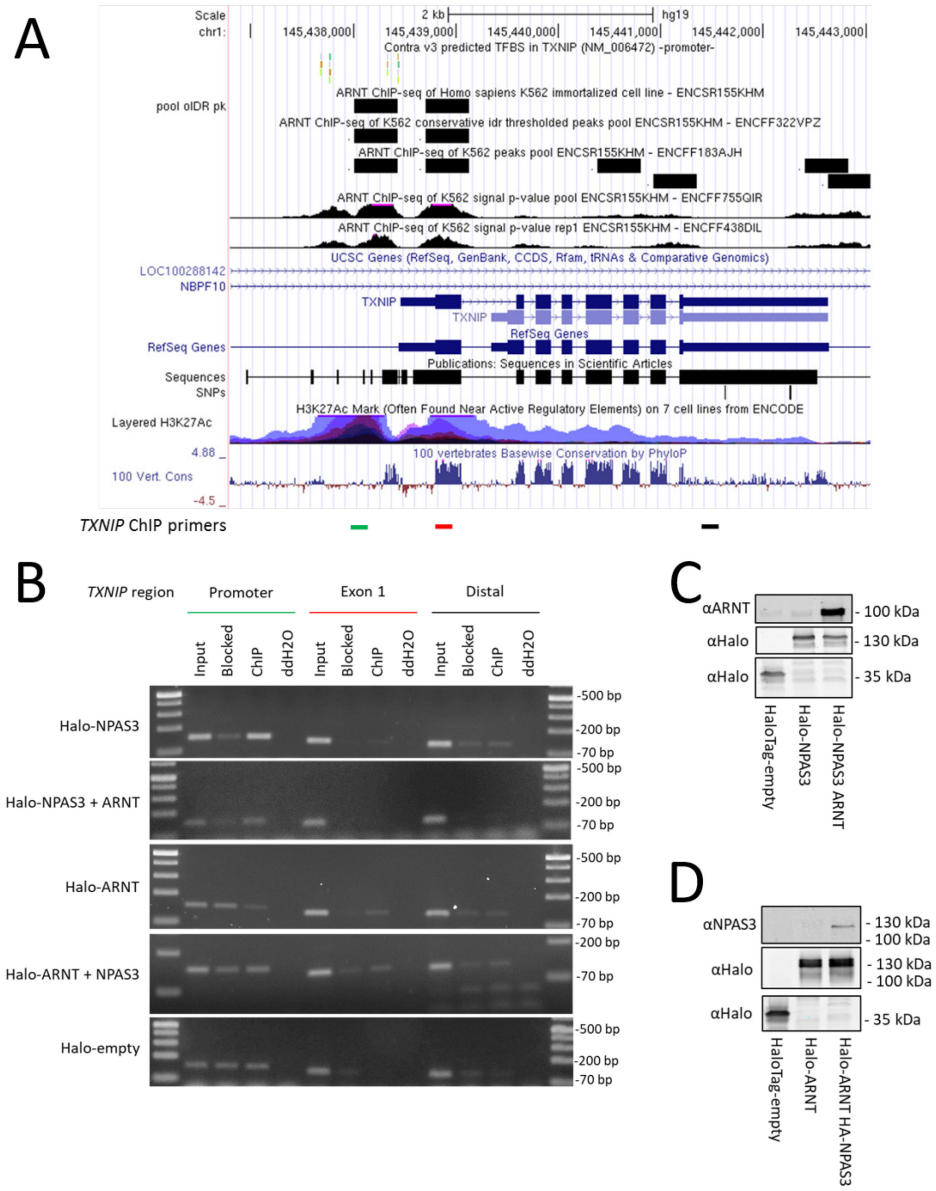



**Figure 4-4: HaloCHIP demonstrating NPAS3 and ARNT binding proximal to the *TXNIP* transcription start site**
(A) UCSC genome browser view of the *TXNIP* locus on human genome build hg19 with ENCODE ARNT ChIP-seq peak data and primers used for HaloCHIP PCR assessment of binding. Accessed from http://genome.ucsc.edu. ConTra v3.0 data indicating the location of predicted ARNT binding sites based on position weight matrices from the JASPAR database, key: binding sites predicted for green (MA0259.1) HIF1A::ARNT, gold (MA0004.1) ARNT, and blue (MA0006.1) AhR::ARNT. (B) HaloCHIP results on the *TXNIP* locus, demonstrating direct binding by NPAS3 and ARNT. (C-D) Western blots showing expression of all constructs used for HaloChIP.

## 4.6 NPAS3 and ARNT cooperatively regulate the *TXNIP* promoter

As NPAS3 was found to bind directly to the proximal promoter region of *TXNIP*, a promoter construct was designed for synthesis that encodes the 867 bp upstream of the transcription start site of *TXNIP* as defined by NCBI transcript variant 1 (NM_006472.1, -867 = Chr1:145,438,461 hg19) to 68 base pairs into the 5'UTR to ensure that the true transcription start site was included in the construct to drive luciferase reporter gene expression (Figure 4-5). A clone of the selected 950 bp region of the *TXNIP* promoter could not be synthesized without mutations, despite several attempts, likely owing to the repetitive G/C rich regions that had previously also limited primer design in the region for PCR analysis and also resulted in a lack of quality assurance by the manufacturer (IDT DNA). A construct was used with three variants relative to the UCSC genome browser build hg19, relative to the transcription start site (Chr1:145,438,461) are as follows: -246hetA/C (Chr1:145,438,215hetA/C), -222ΔC (Chr1:145,438,239ΔC) and +49T>C (Chr1:145,438,510T>C). As these variants were found to be outside of the potential ARNT binding sites predicted by ConTra v2 (Broos et al. 2011) and in regions of relatively low conservation, the construct was used for luciferase analysis with the variants (Figure 4-5).

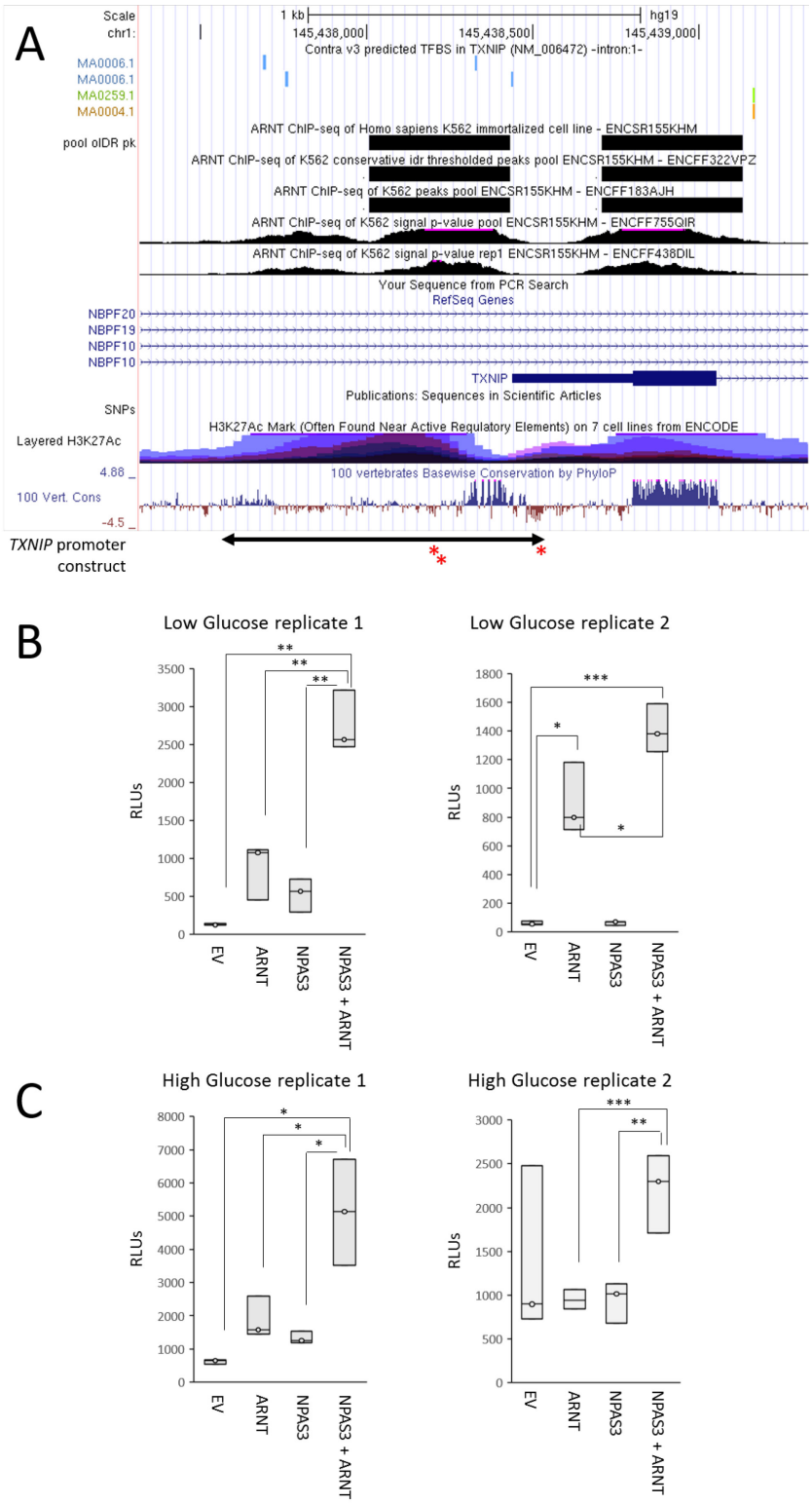



**Figure 4-5: NPAS3 and ARNT are able to regulate expression driven by the promoter region of *TXNIP*.**

Figure 4-5 Figure legend: (A) UCSC genome browser view of the *TXNIP* locus on human genome build hg19 with ENCODE ARNT ChIP-seq peak data. Accessed from http://genome.ucsc.edu. Approximate interval of the TXNIP promoter construct is indicated by the double arrow. ConTra v3.0 data indicating the location of predicted ARNT binding sites based on position weight matrices from the JASPAR database, key: binding sites predicted for green (MA0259.1) HIF1A::ARNT, gold (MA0004.1) ARNT, and blue (MA0006.1) AhR::ARNT. Red asterisks: variants in gene block which are not localized to the predicted ARNT binding sites or in regions of high conservation. (B) Two replicates of luciferase results of HEK 293T cells cultured in low glucose media and transfected with pGL4.10-TXNIPpromoter in the presence and absence of NPAS3 and/or ARNT. Firefly luciferase activity was normalized against activity of a co-transfected renilla luciferase driven by the constitutive TK promoter to result in relative luciferase units (RLUs). (C) Two replicates of the same luciferase assays performed in high glucose media. n=3 for each replicate. Top and bottom bars indicate the 25$^{th}$ and 75$^{th}$ centile, center bar indicates median value. *p<0.05, **p<0.01, ***p<0.001

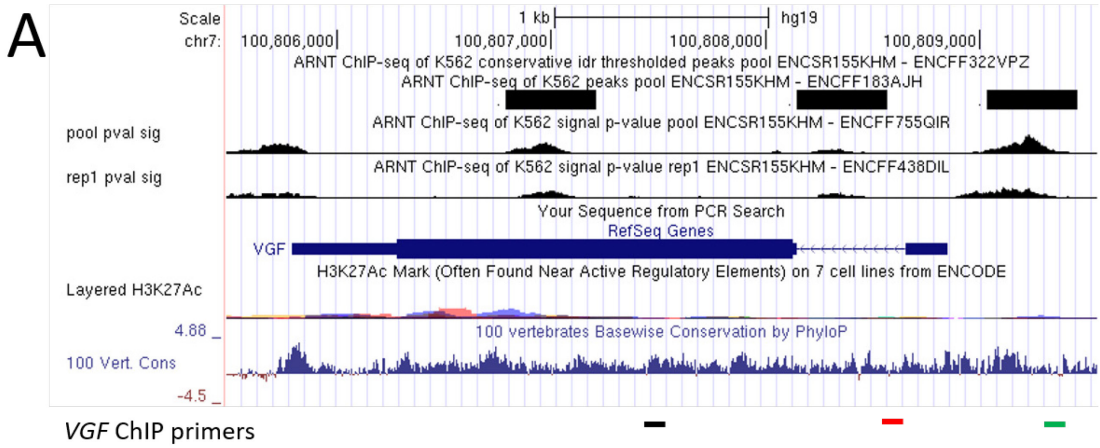
Luciferase analysis of this promoter was performed in both low and high glucose conditions, to confirm that the promoter region was normally responsive to glucose (Figure 4-5B,C relative EV RLU mean values: low glucose = 58.8±14.9, 131.2±11.8; high glucose = 1366.4±966.5, 624.7±68.6). High concentrations of glucose have been shown to activate transcription driven by a region of the *TXNIP* promoter contained within our construct (the minimal glucose responsive region is unmutated in this construct). Furthermore, I have found *TXNIP* to be expressed at a higher level in high glucose than low glucose in HEK 293T cells (see section 5.3). These data suggest that our promoter construct is behaving as expected for the *TXNIP* proximal promoter.

NPAS3 and ARNT were found to coordinately regulate the expression driven by the *TXNIP* promoter region. However, expression driven by this promoter construct was activated contrary to the expected effect given the previous result that NPAS3 and ARNT cause repression of *TXNIP* (Figure 4-3). Despite the opposite regulatory effect than

observed from the endogenous *TXNIP* promoter, these data demonstrate that NPAS3 and ARNT are able to regulate expression of *TXNIP,* in part through direct binding within the interval -867 bp upstream to +68 bp from the transcription start site. These data suggest that the sequences required for repression are not coded for within this construct, however, they support the finding that NPAS3 and ARNT in combination directly regulate the expression of *TXNIP,* at least partly through the proximal promoter. Owing to issues with construction of reporters of *TXNIP* promoter driven expression, further experimentation was not pursued.

## 4.7 NPAS3 directly binds the *VGF* promoter

HaloCHIP samples were also probed for the *VGF* locus as I found *VGF* to be upregulated by NPAS3 and ARNT. The location of primers used to identify the region(s) that NPAS3 and/or ARNT bind at the *VGF* locus are shown in Figure 4-6. NPAS3 was found to directly bind proximal to the promoter of *VGF* (Figure 4-6). Signal was also observed on the distal site, and this secondary peak is not entirely unexpected and may represent independent binding at a downstream potential ARNT ChIP-seq peak, seen in Figure 4-6A. ARNT ChIP-seq peaks were observed across the *VGF* region preventing design of primers >1 kb distal to a potential ARNT ChIP-seq peak. HaloCHIP performed with HaloTag-ARNT resulted in positive ChIP signals at all probed regions at the *VGF* locus, suggesting that ARNT can bind to multiple sites proximal to *VGF* (Figure 4-7). Taken together, these data demonstrate that both ARNT and NPAS3 directly bind proximal to the promoter of *VGF* and may also bind to a site distal to the promoter, however the distal HaloCHIP-positive site was not characterized further.



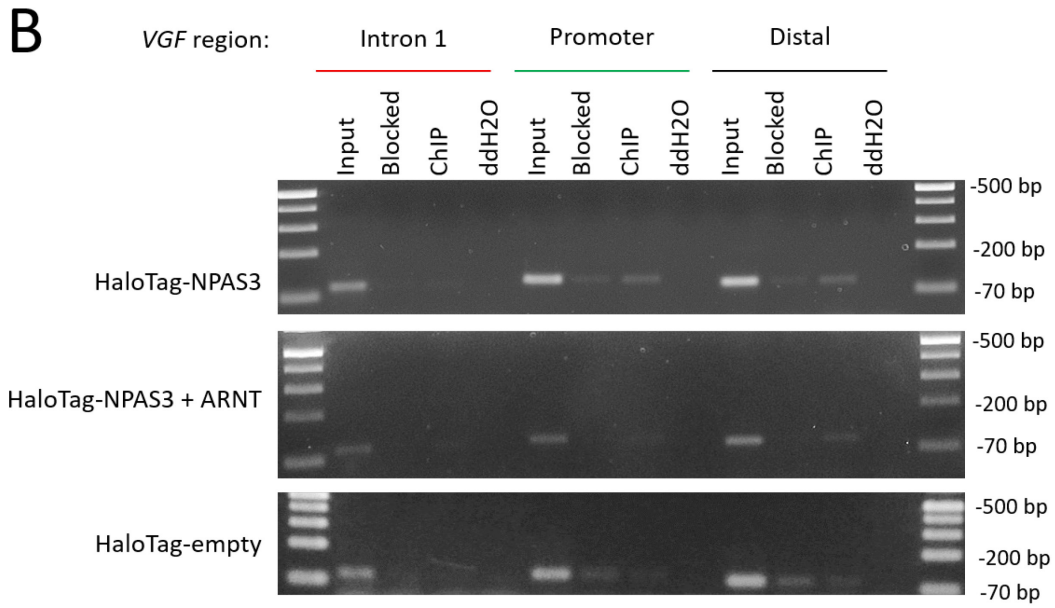




**Figure 4-6: HaloCHIP PCR demonstrating NPAS3 directly binding to the promoter region of *VGF*.**
(A) UCSC genome browser of the human genome build hg19 showing the *VGF* locus with ENCODE ARNT ChIP-seq data and primers used for assessment of binding in HaloCHIP experiments. Accessed from: http://genome.ucsc.edu. (B) Results of HaloCHIP with HaloTag-NPAS3 with and without co-expression of ARNT.

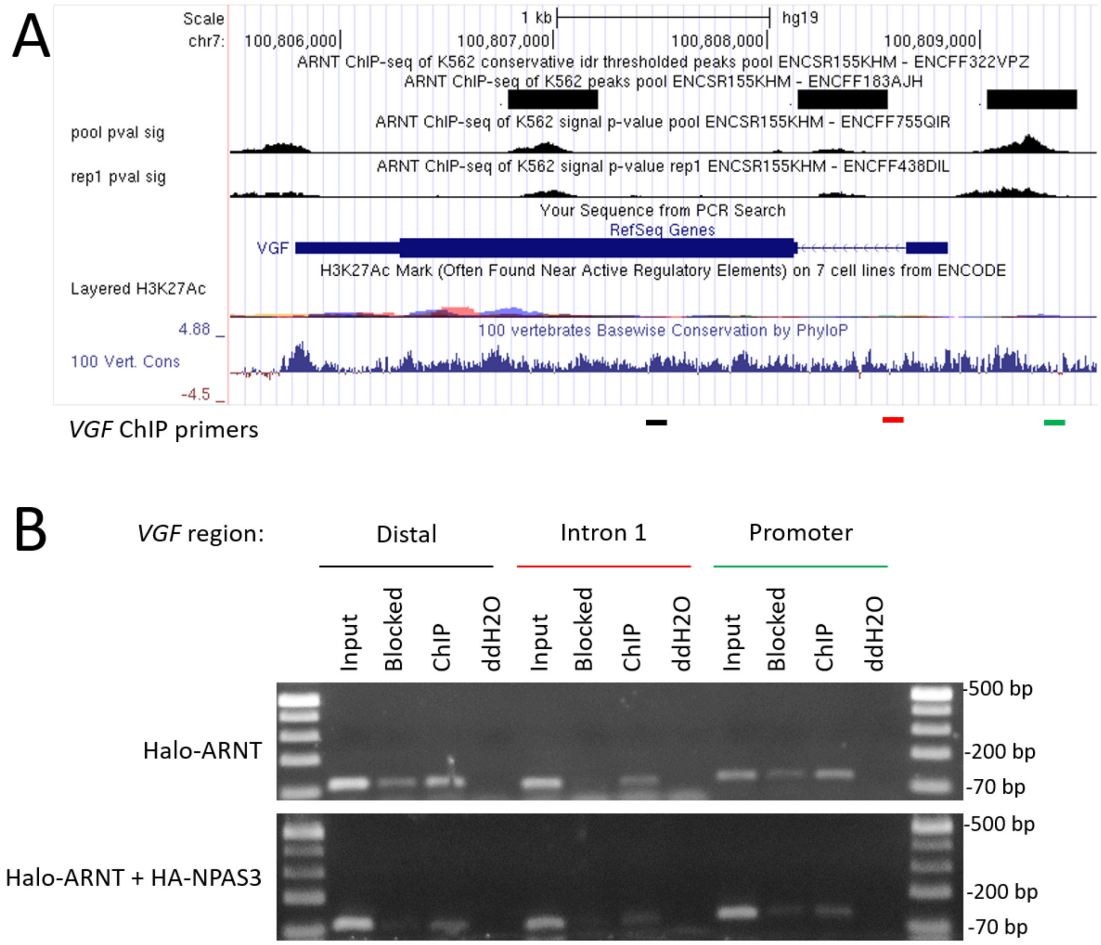



**Figure 4-7: HaloCHIP data demonstrating that ARNT directly binds the *VGF* locus.**
(A) UCSC genome browser of the human genome build hg19 showing the *VGF* locus with ENCODE ARNT ChIP-seq data and primers used for assessment of binding in HaloCHIP experiments. Accessed from: http://genome.ucsc.edu. (B) Results of HaloCHIP with HaloTag-ARNT, with and without co-expression of NPAS3.

## 4.8 Regulation of the *VGF* promoter by NPAS3 and ARNT

As both NPAS3 and ARNT were found to directly bind the region directly upstream of the *VGF* transcription start site, the interval under the ENCODE ARNT ChIP-seq peak was cloned upstream of a promoterless luciferase gene in pGL4.10 for luciferase reporter analysis (Figure 4-8). The luciferase reporter driven by this promoter was found to be expressed and both NPAS3 and ARNT were found to be able to individually activate luciferase expression (Figure 4-8). Both isoforms of NPAS3 were similarly able to activate expression of the reporter driven by this promoter. Co-expression of NPAS3 933 aa with ARNT resulted in a further increase in activity of this promoter, supporting the conclusion that they act cooperatively to regulate expression driven by the *VGF* promoter. These data demonstrate that NPAS3 and ARNT are able to cooperatively regulate the expression of *VGF* through its proximal promoter. Combined with the HaloCHIP data, these data support the conclusion that NPAS3 and ARNT directly bind the promoter region of *VGF* to activate its expression.

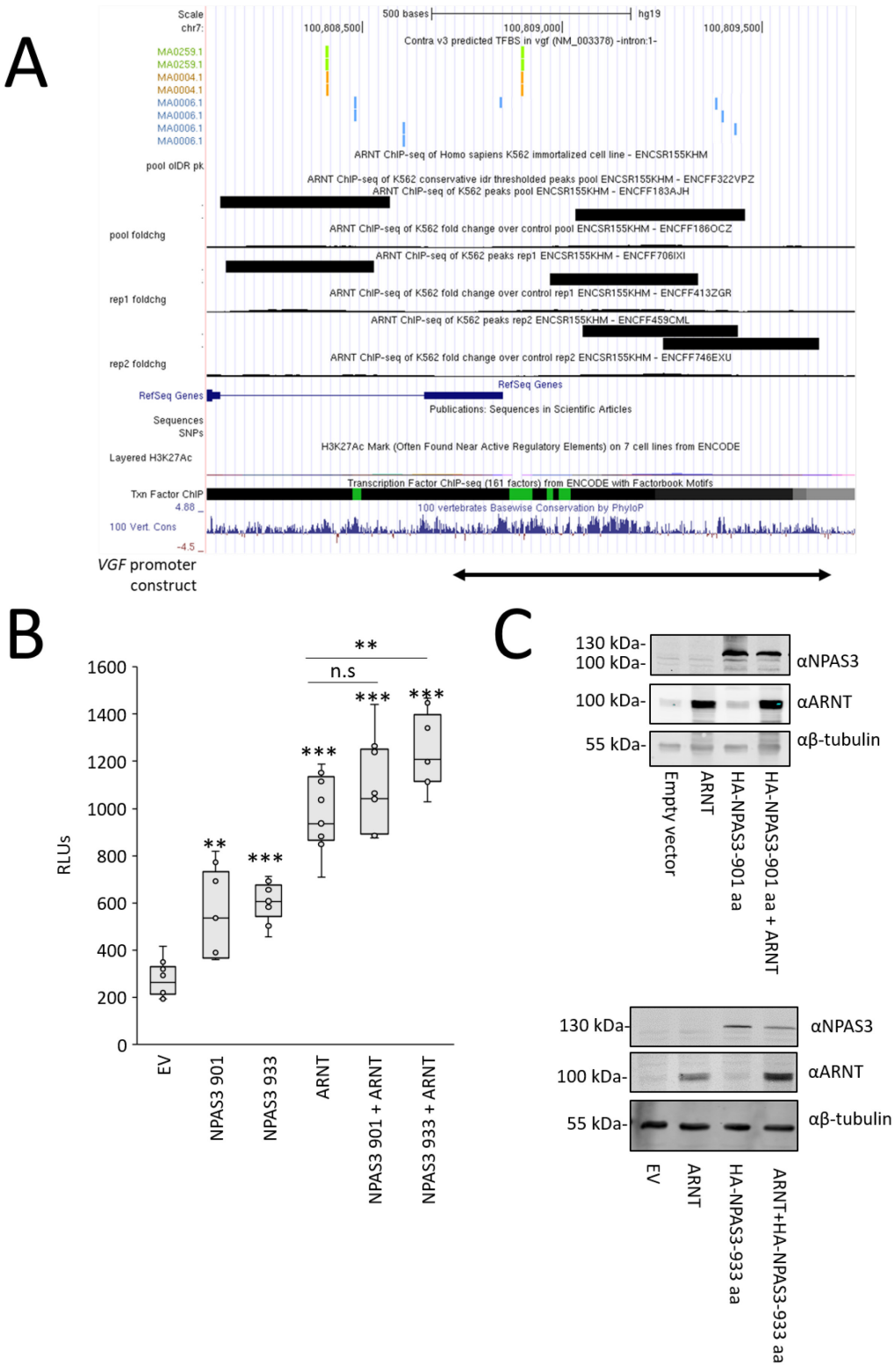



**Figure 4-8: Luciferase assessment of NPAS3 and ARNT regulation of a reporter driven by the *VGF* promoter.**

Figure 4-8 Figure legend: (A) Approximate interval of the *VGF* promoter construct is indicated by the double arrow. ConTra v3.0 data indicating the location of predicted ARNT binding sites based on position weight matrices from the JASPAR database, key: binding sites predicted for green (MA0259.1) HIF1A::ARNT, gold (MA0004.1) ARNT, and blue (MA0006.1) AhR::ARNT. Accessed from http://genome.ucsc.edu (B) Two isoforms of NPAS3 expressed with and without ARNT were expressed along with a luciferase reporter driven by 750 bp proximal to the *VGF* transcription start site. Firefly luciferase activity was normalized against activity of a co-transfected renilla luciferase driven by the constitutive TK promoter to result in relative luciferase units (RLUs). Top and bottom bars indicate the 25$^{th}$ and 75$^{th}$ centile, center bar indicates median value. Whiskers indicate data extremes. Asterisks over top of box and whiskers indicated relative activation over EV. $n$=6 *$p$<0.05, **$p$<0.01, ***$p$<0.001 (C) Western blots showing expression of all constructs used to drive luciferase reporter expression.

## 4.9 Functional assessment of NPAS3 variants identified in the human population

In order to determine their effects on NPAS3 function in the individuals who carry them, three coding variants of NPAS3 were functionally assessed:c.910G>A (p.Val304Ile), c.1654G>C (p.Ala552Pro) and c.2089G>A (p.Gly697Ser). All variants were found to be expressed with no obvious secondary bands according to western blots (Figure 4-9A,B). Two NPAS3 variants were tested for interaction of their translated proteins with ARNT: p.Val304Ile and p.Ala552Pro given their previous association with neuropsychiatric disorder and, for p.Val304Ile, its localization between sequences coding for the PAS interaction motifs of NPAS3. p.Gly697Ser was not assessed due to its lack of association with disorder, as well as its relative localization to the transactivation domain of the translated protein, which are not predicted to be required for interaction with ARNT. Both p.Val304Ile and p.Ala552Pro were found to interact normally with ARNT

(Figure 4-9C). Immunofluorescence showed that all protein variants were able to localize to the nucleus and no differences in localization were noted (Figure 4-10).

Variants were tested for relative gene regulatory function using the *VGF* promoter driven luciferase reporter vector, when expressed individually and co-expressed with ARNT. Both transactivation domain variants (p.Ala552Pro and p.Gly697Ser) were found to normally activate *VGF* promoter driven expression relative to the wild-type NPAS3. With co-expression of ARNT, the p.Ala552Pro variant acted normally, however the p.Gly697Ser had apparent reduction in transactivation ability relative to wild-type NPAS3 co-expressed with ARNT, despite apparently normal transactivation function observed when singly transfected (Figure 4-11A).

Assessment of the p.Val304Ile variant was carried out with both the 933 aa and 901 aa isoforms of NPAS3. In both isoforms this variant was found to transactivate *VGF* promoter driven expression normally, relative to the wild-type of both isoforms, and cooperative transactivation with co-transfection with ARNT was also observed, further supporting the HaloTag pull down data demonstrating that NPAS3 p.Val304Ile is able to interact with ARNT (Figure 4-11B).

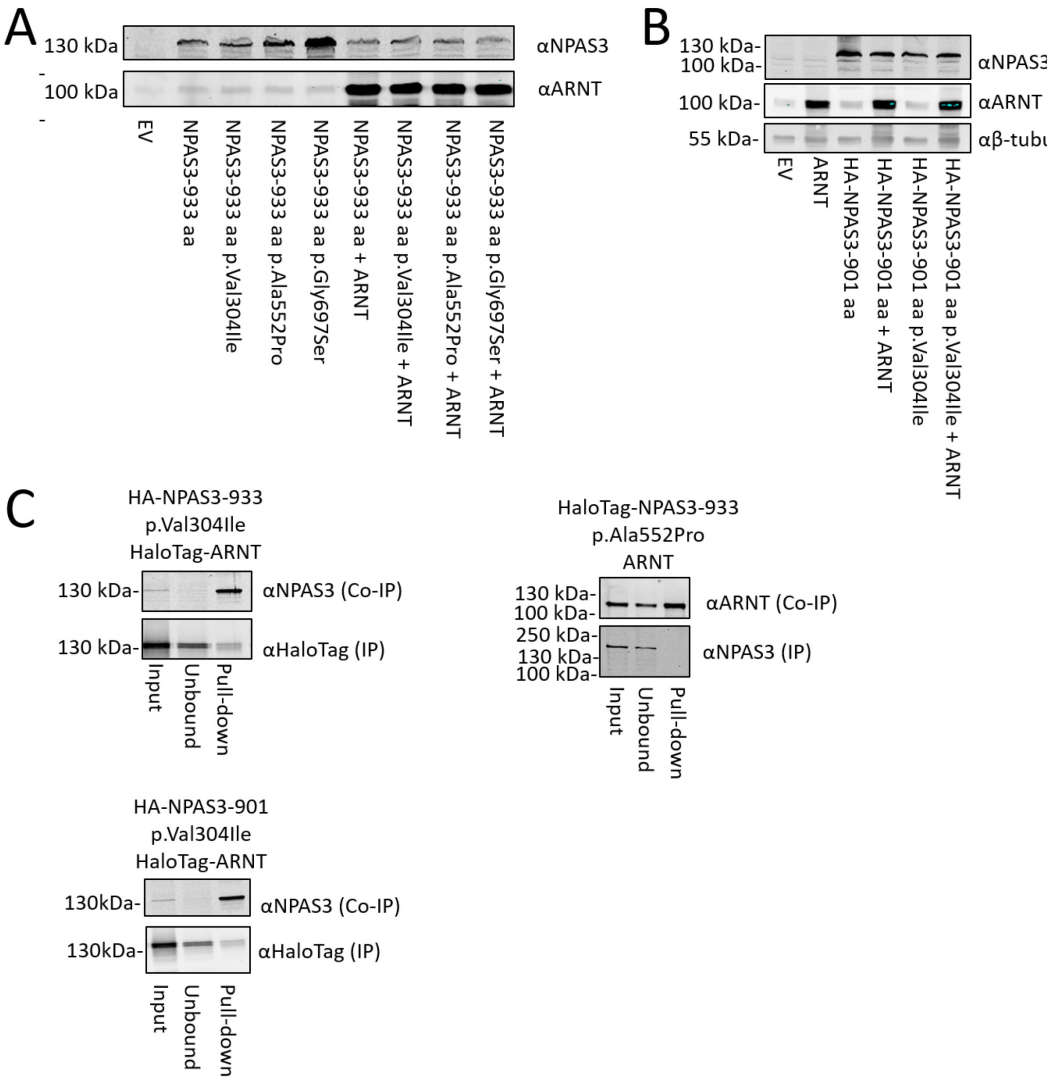



**Figure 4-9: NPAS3 variants are normally expressed and able to interact with ARNT.**

(A) Western blots showing expression of wild type NPAS3 933 aa isoform 1 and all variants tested in this isoform in the presence and absence of ARNT. (B) Western blot showing expression of the wild type NPAS3 901 aa isoform 2 and the variant tested in this isoform in the presence and absence of ARNT. (C) HaloTag pull down data showing that the PAS domain proximal variant p.Val304Ile is normally able to interact with ARNT, as well as the previously schizophrenia-associated variant p.Ala552Pro.

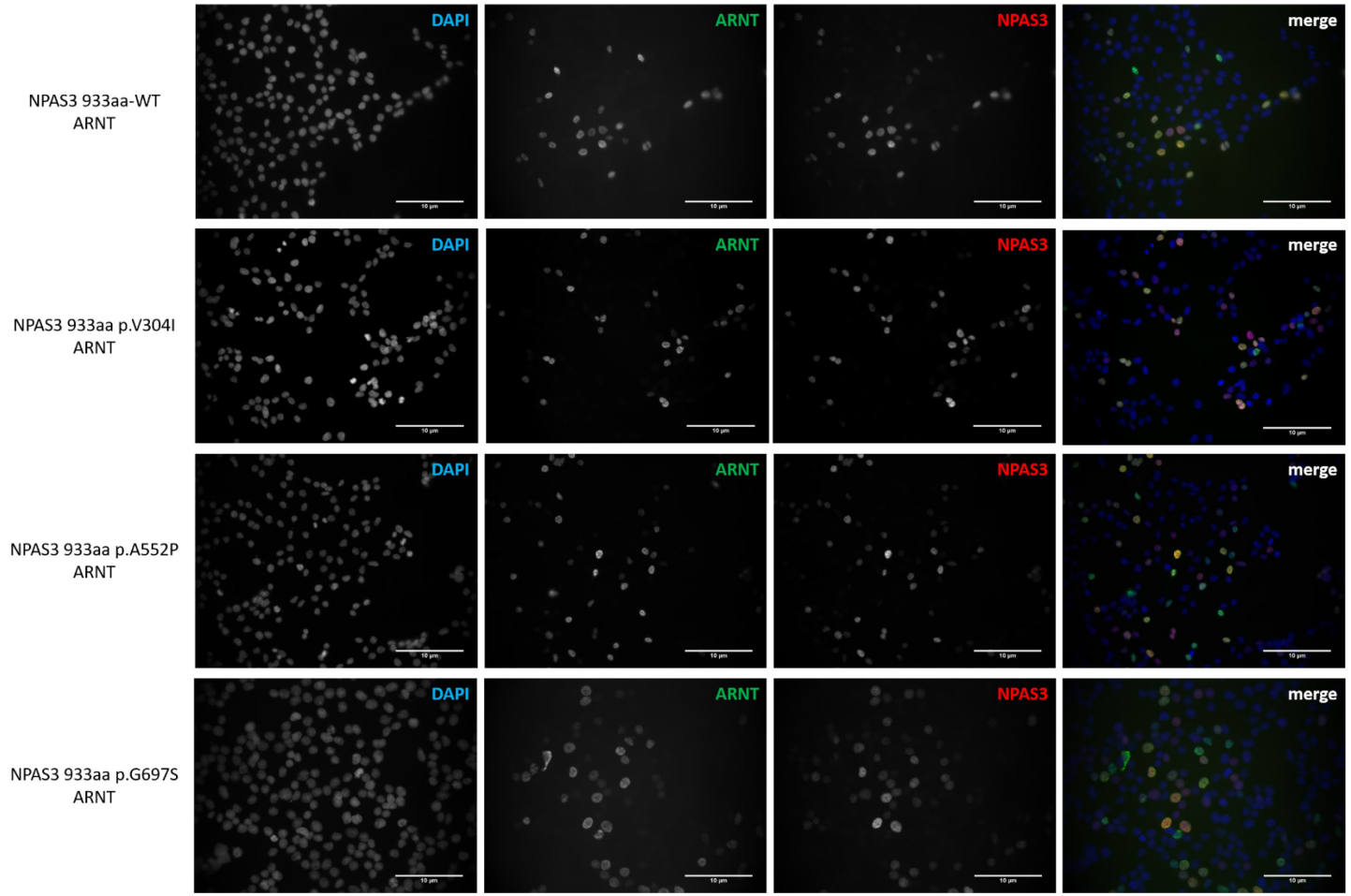



**Figure 4-10: Immunofluorescence of variants demonstrating nuclear localization**
HEK 293T expressing NPAS3 isoform 1 (933 aa) with variants p.Val304Ile, p.Ala552Pro, and p.Gly697Ser and ARNT isoform 1, showing predominantly nuclear localization. Blue: DAPI, green: ARNT, red: NPAS3. 400X magnification, scale bar = 10 μm.

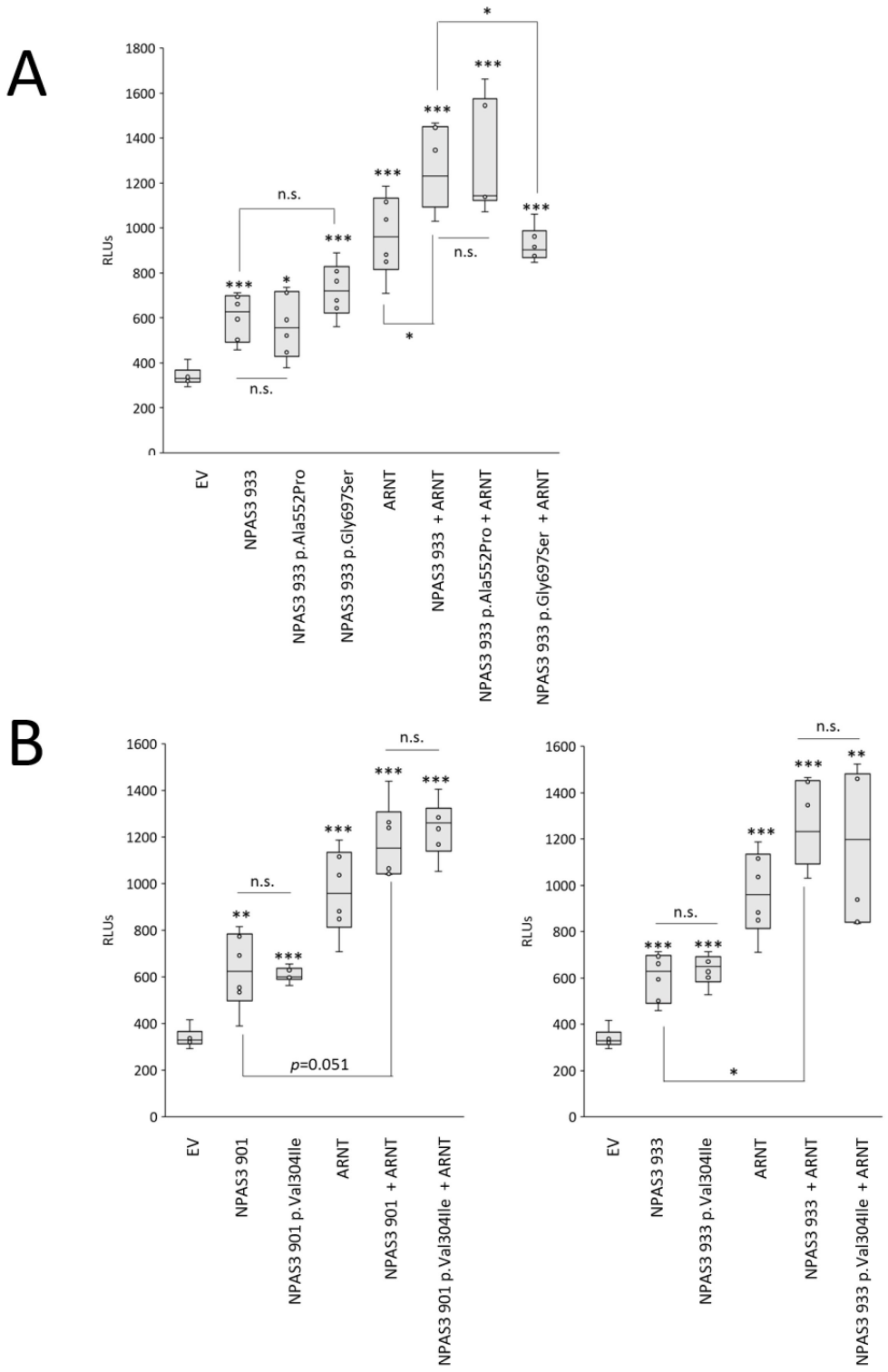



**Figure 4-11: NPAS3 variants are able to activate expression driven by the *VGF* promoter.**

Figure 4-11 figure legend: (A) *NPAS3* variants identified by our group were expressed with and without ARNT co-transfected with a luciferase reporter driven by 750 bp proximal to the VGF transcription start site. (B) Assessment of the PAS domain variant p.Val304Ile by luciferase in two isoforms of NPAS3. Firefly luciferase activity was normalized against activity of a co-transfected renilla luciferase driven by the TK promoter to calculate RLUs. Top and bottom bars indicate the 25$^{th}$ and 75$^{th}$ centile, center bar indicates median value. Whiskers indicate data extremes. n=6 *$p$<0.05, **$p$<0.01, ***$p$<0.001. Asterisks over the tops of box and whisker plots indicated relative activation over EV.

## 4.10 Analysis of the functional role of NPAS3 domains

In order to assess the domains required for the interaction between NPAS3 and ARNT, deletion constructs were generated which divided NPAS3 (isoform 1) into three functional domains: the bHLH DNA-binding domain (bHLH), the PAS interaction domain (PAS), and the C-terminal predicted transactivation domain (TAD, Figure 1-1). These HA-tagged domains were expressed in HEK 293T cells with HaloTag-ARNT and total lysates were applied to the HaloResin to determine which domains of NPAS3 interact with ARNT. No domain in isolation was found to be able to interact with ARNT (Figure 4-12). Both the bHLH and PAS domains were found to be required for the interaction between NPAS3 and ARNT and the predicted transactivation domain was found to be unnecessary.

Assessment of the NPAS3 domains by immunofluorescence provides insight into their contribution to subcellular localization. NPAS3 domain constructs were transfected in the presence and absence of ARNT, and NPAS3 localization was read by a second reader blinded to the presence of ARNT and the identity of the domains. Constructs were scored as nuclear based on the absence of cytoplasmic signal; if cytoplasmic signal was

observed, localization was scored as nuclear=cytoplasmic. The localization data shown in Figure 4-13 and Figure 4-14 are consistent with the expected results from literature and bioinformatic analysis. NPAS3 localization was found to be primarily nuclear and this nuclear localization was enhanced by co-expression with ARNT (Figure 4-13, Figure 4-14A). A predicted nuclear localization sequence is present in the putative transactivation domain of NPAS3 (Macintyre et al. 2010). Almost complete nuclear localization of constructs containing the transactivation domain was observed regardless of ARNT expression, supporting the function of the predicted nuclear localization sequence (Figure 4-14A).

Constructs not containing the transactivation domain were not predominantly nuclear in all conditions (Figure 4-14A). The bHLH domain was found to be equally distributed between nuclear and nuclear=cytoplasmic localization, when expressed in isolation. However, co-expression of *ARNT* resulted in decreased nuclear localization. Of note, a nuclear export sequence (NES) has been predicted in the region of the bHLH domain construct (residues 1-125, NetNES predicted NES from residues 88-98, Figure 4-14B) (la Cour et al. 2004). The localization of constructs including the PAS domain were found to be enriched in the nucleus with co-expression of ARNT. The bHLH-PAS domain was equally commonly found to be primarily nuclear or distributed equally between the nucleus and cytoplasm when expressed in isolation, however *ARNT* expression resulted in increased nuclear localization, supporting the data that these two constructs can functionally interact. Furthermore, these data indicate that interaction between NPAS3 and ARNT affects the localization of NPAS3. The PAS domain in isolation was not observed to have specific nuclear localization unless *ARNT* was

expressed, suggesting that some level of functional interaction with ARNT may be preserved although it was not pulled down by the HaloTag pull down protocol. This may indicate that this interaction is transient/low affinity, or signal was too low to be detected (Figure 4-14A, Figure 4-12).

NPAS3 domains were expressed in the presence and absence of ARNT to drive expression of the *VGF* promoter driven luciferase reporter in order to assess relative transactivation function of each domain of NPAS3. ARNT individually was able to activate *VGF* expression, as expected. Of all *NPAS3* constructs tested, only constructs encoding the transactivation domain of NPAS3 were found to be able to activate expression of the *VGF* promoter driven reporter (Figure 4-15 A). When co-transfected with *ARNT*, only full-length *NPAS3* was found to be able to cooperatively activate expression. The bHLH domain only and bHLH-PAS domain constructs had no effect on luciferase reporter expression when expressed in isolation. However, when co-expressed with *ARNT*, they caused relative down-regulation of luciferase reporter expression relative to the singly-transfected *ARNT* constructs potentially due to their ability to interact with ARNT. This may result in a heterodimer which is not able to activate reporter expression due to the absence of the C-terminal transactivation domain. Loss of the bHLH domain resulted in relative decrease in the ability of NPAS3 to co-activate the *VGF* promoter when co-expressed with ARNT, supporting its role in the interaction between NPAS3 and ARNT and in DNA binding specificity.

Expression of the individual domain constructs to drive promoterless (empty pGL4.10) luciferase expression was performed to confirm that the effects I observed with the *VGF* and *TXNIP* promoters were specifically driven by the promoter constructs.

Expression of *NPAS3* in the presence and absence of ARNT did not drive expression of the promoterless reporter, demonstrating specificity of the observed regulatory effect to the promoter regions tested. Intriguingly, the transactivation domain when expressed without the bHLH domain was able to weakly activate expression of luciferase non-specifically, relative to full-length NPAS3 (Figure 4-15 B). These data support the role of the C-terminal half of NPAS3 as a potent transactivation domain, and the bHLH domain in conferring sequence specificity of the regulatory action of NPAS3.

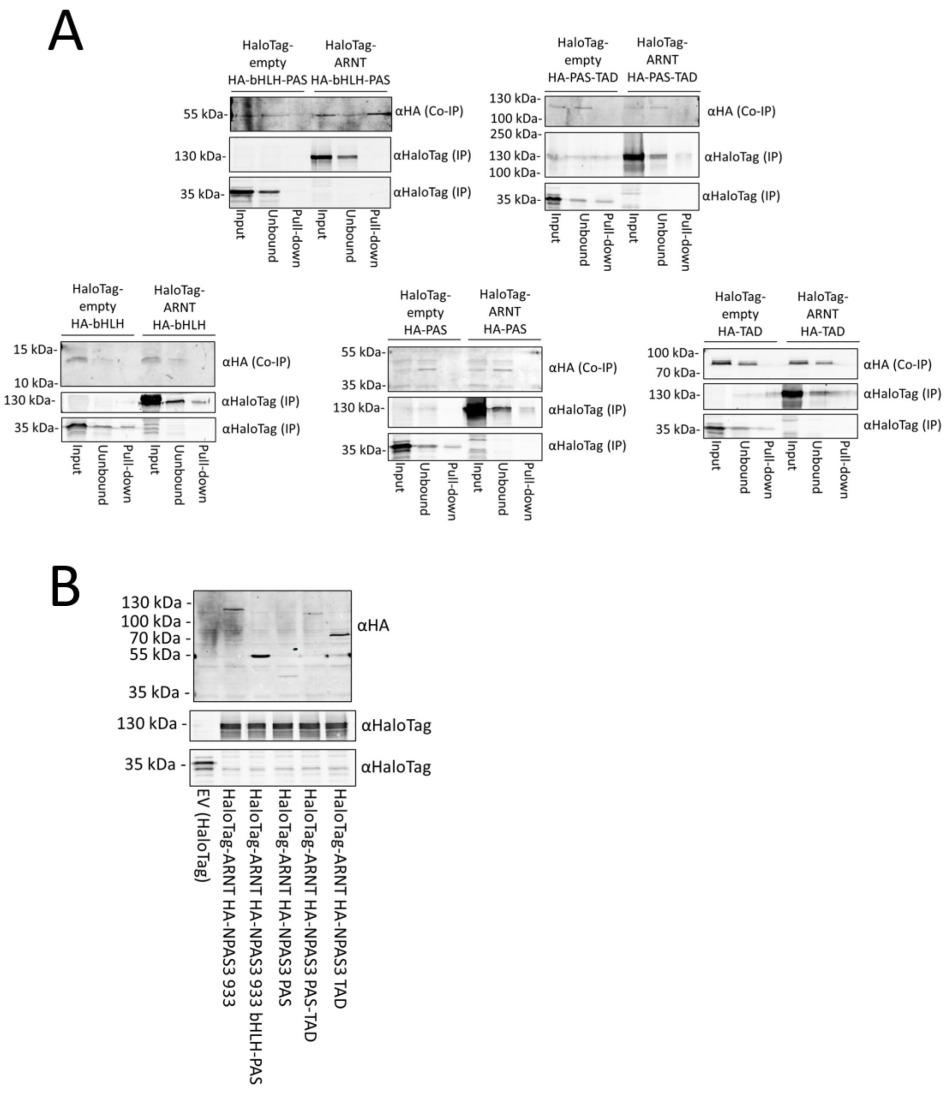



**Figure 4-12: Both the bHLH and PAS domains are required for interaction of NPAS3 with ARNT.**
(A) HaloTag fused ARNT pull down of NPAS3 domain constructs. (B) Western blot demonstrating expression of NPAS3 domain constructs. Expected sizes: full length: 101 kDa, bHLH-PAS 49.4 kDa, PAS 35.8 kDa, PAS-TAD 98.2 kDa, TAD 53.2 kDa. Not shown: bHLH, 13 kDa.

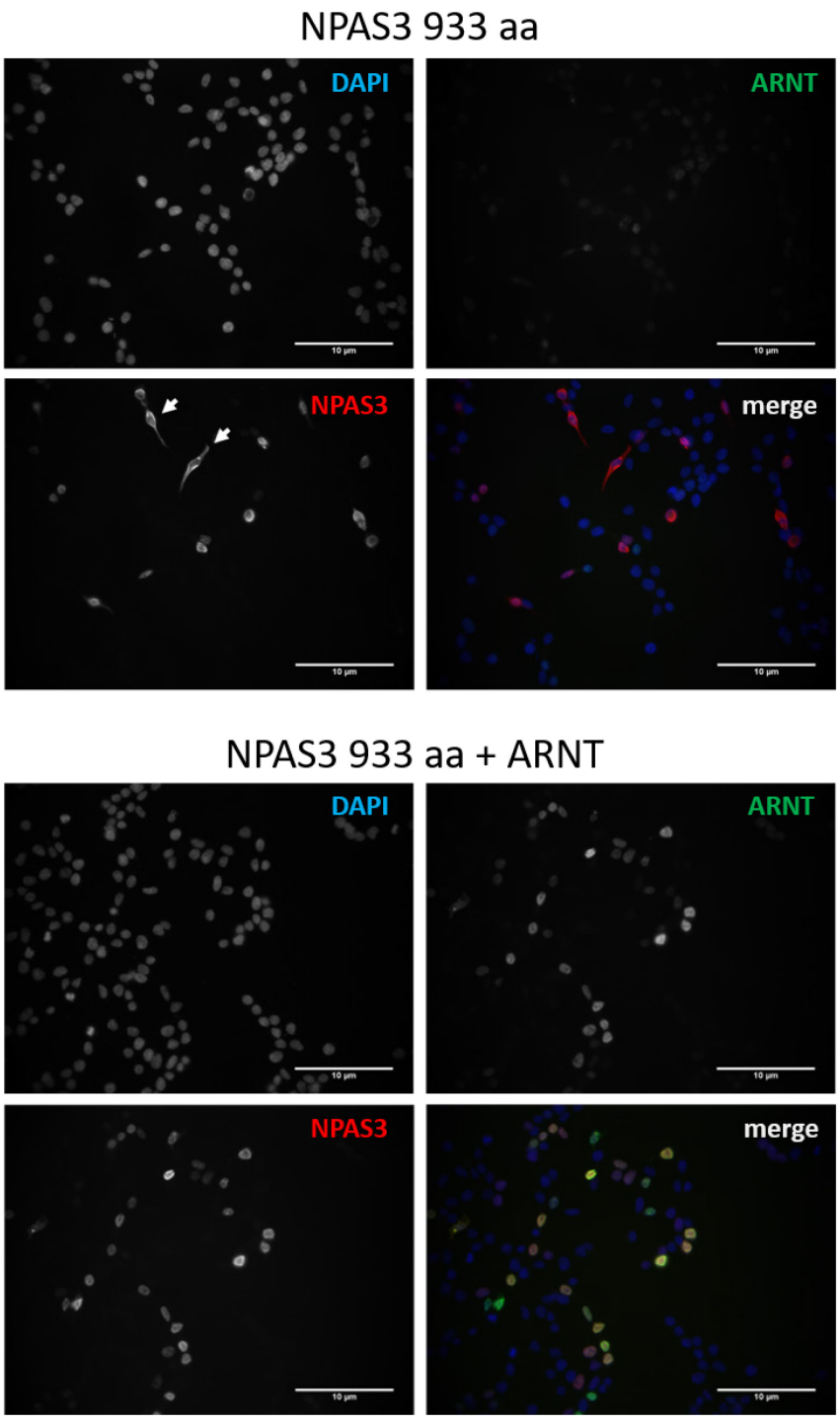



**Figure 4-13: NPAS3 localization is affected by ARNT expression.** (A) Immunofluorescence data demonstrating that NPAS3 is primarily localized to the nucleus when expressed in HEK 293T cells. Co-transfection with ARNT enhances nuclear localization observed, reducing the number of cells with cytoplasmic signal (white arrows). 400X magnification, scale bar: 10 µm.

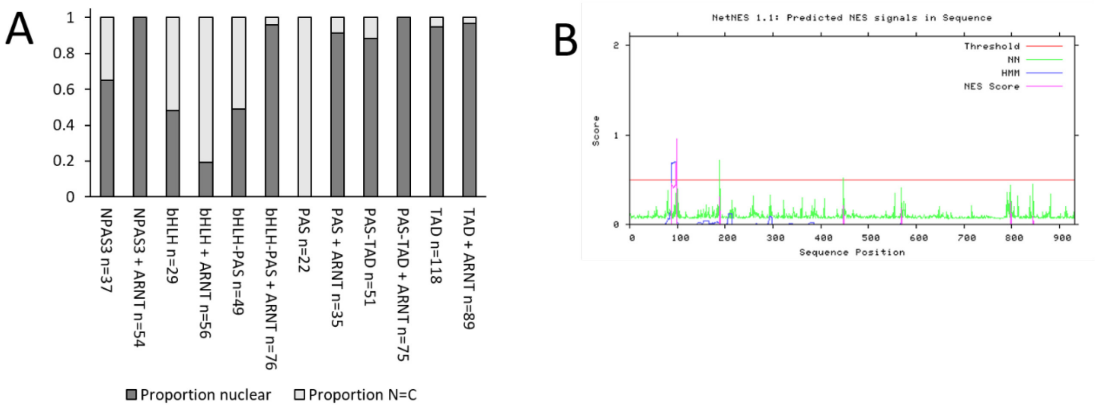



**Figure 4-14: Quantification of factors affecting NPAS3 localization.** (A) Quantification of immunofluorescence images taken of each NPAS3 domain construct expressed in isolation or with ARNT co-expression. Localization was scored as either entirely nuclear or equally distributed between the nucleus and cytoplasm (N=C). Sample sizes are indicated in X-axis labels. (B) NetNES 1.1 prediction of nuclear export signals in the 933 aa isoform of NPAS3 in the region encoded by the bHLH domain (residues 1-125). Accessed from: http://www.cbs.dtu.dk/services/NetNES/

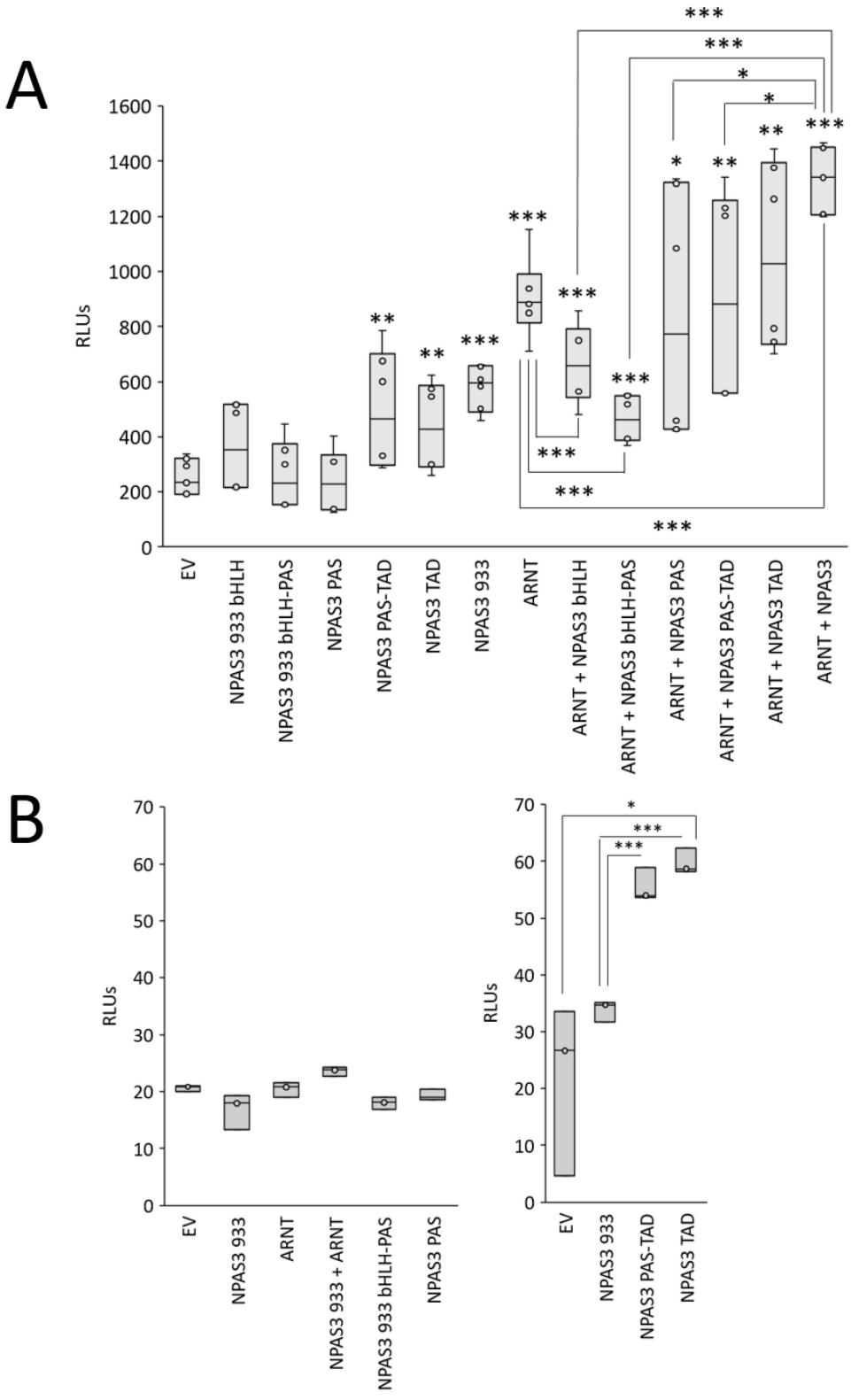



**Figure 4-15: Luciferase assessment of NPAS3 domain function.**

Figure 4-15 figure legend: (A) Luciferase results of NPAS3 domains expressed individually or with ARNT. NPAS3 domains were expressed with and without ARNT co-transfected with a luciferase reporter driven by 750 bp proximal to the VGF transcription start site. Firefly luciferase activity was normalized against activity of a co-transfected renilla luciferase driven by the constitutive TK promoter to calculate RLUs. Asterisks over the top of box and whisker plots indicate relative activation over EV, n=6. (B) Effect of NPAS3 domains on promoterless luciferase (empty pGL4.10) expression, n=3. Top and bottom bars indicate the 25$^{th}$ and 75$^{th}$ centile, center bar indicates median value. Whiskers indicate data extremes, *$p$<0.05, **$p$<0.01, ***$p$<0.001

## 4.11 Summary of results

Through these experiments I have developed an antibody that can specifically detect the expression of NPAS3. Fourteen genes were identified as potential co-targets of NPAS3 and ARNT (Table 10) and two of these genes, *VGF* and *TXNIP*, were found to be differentially regulated by co-expression of NPAS3 and ARNT. The ability of NPAS3 and ARNT to regulate expression of these genes by their proximal promoters was confirmed by luciferase reporters driven by the *VGF* and *TXNIP* promoters. Direct binding of the promoter regions of both *VGF* and *TXNIP* was also detected by HaloCHIP.

Three variants of *NPAS3* identified in the human population were characterized for their functional significance using the assays developed to characterize the wild-type NPAS3::ARNT complex. The previously schizophrenia-associated variants, c.910G>A (p.Val304Ile) and c.1654G>C (p.Ala552Pro) were found to be functionally identical to the wild-type NPAS3 isoforms. The rare population variant c.2089G>A (p.Gly697Ser) was found to have normal transactivation function, however reduced ability to cooperatively activate gene regulation when co-expressed with *ARNT* was observed.

Finally, the domains of NPAS3 were functionally characterized using the luciferase assays developed for the characterization of NPAS3- and ARNT-mediated gene regulation. Our data are consistent with the predicted role of the bHLH domain as a DNA-binding domain critical for the specificity of regulation by NPAS3 that, alongside the PAS domain, is necessary for interaction between NPAS3 and ARNT. Further, luciferase studies demonstrate that the C-terminal half of NPAS3 does have inherent transactivation function.

# Chapter 5: Cellular factors in NPAS3 regulation and function

## 5.1 Rationale

During experimentation to assess the gene-regulatory function of the NPAS3::ARNT heterodimer, I observed differential gene regulation dependent on culture conditions. Furthermore, several potential targets of NPAS3 have been shown to be regulated in response to environmental stressors. As such, I sought to identify cellular context that might influence the gene regulatory output of NPAS3. To this end, I looked at several stimuli relevant to identified NPAS3 target genes that are known to affect regulation by ARNT. To this end, I also assessed gene regulation driven by HIF1A, the hypoxia stabilized bHLH-PAS heterodimeric partner of ARNT (Wang et al. 1995), as well as BMAL1, a bHLH-PAS protein whose expression, localization and regulatory effects are regulated in a circadian pattern (Lee et al. 2008; Rey et al. 2011; Spengler et al. 2009; Tamaru et al. 2009). For experiments assessing the function of endogenous NPAS3, focus was placed on environmental stimuli with known effects on *TXNIP* expression to determine if NPAS3 affects the regulation of *TXNIP* in response to these stimuli.

## 5.2 Hypoxia

Several of the potential co-targets of NPAS3 and ARNT listed in Table 10 have functions in hypoxia. *ANKRD37* was found to be the most robustly regulated gene by

NPAS3 and to have very strong ARNT ChIP-seq peaks flanking the transcription start site (Sha et al. 2012). Furthermore, *ANKRD37* has been shown to be up-regulated as a part of the canonical HIF1A-mediated hypoxia response (Benita et al. 2009). HaloCHIP samples were probed by qPCR for NPAS3 binding at the *ANKRD37* promoter. As can be seen in Figure 5-1, ARNT binds with high affinity to the *ANKRD37* promoter and while NPAS3 is able to bind the *ANKRD37* promoter when expressed individually, co-expression with *ARNT* causes increased affinity for NPAS3 to the *ANKRD37* promoter. These data demonstrate that NPAS3 and ARNT are able to directly bind the *ANKRD37* promoter.

In my original screen of potential ARNT and NPAS3 co-targets, *ANKRD37* was not found to be differentially regulated (Figure 5-2). This experiment was performed using ambient oxygen conditions (normoxia, 21% oxygen), consistent with the experimental set up used in the microarray study which found *ANKRD37* expression to be repressed by NPAS3 (Sha et al. 2012). As I found NPAS3 and ARNT to bind to the *ANKRD37* promoter in normoxia without effecting expression, and *ANKRD37* is regulated in response to hypoxia, we pursued the effects of NPAS3 and ARNT on its regulation in hypoxia (Benita et al. 2009).

Two models of hypoxia were tested to determine if the lack of regulation of *ANKRD37* by ARNT and NPAS3 is due to the oxygen environment. HEK 293T cells were transfected with *NPAS3* and/or *ARNT*, allowed to recover for 24 hours and then transferred to a hypoxic (3% oxygen) incubator for 24 hours, or cells were allowed to recover for 46 hours and incubated for 2 hours in hypoxia (3% oxygen). In this condition I observed the expected rapid down-regulation of *TXNIP* expression by hypoxia and

NPAS3 and ARNT was found to down-regulate *TXNIP* in normoxia, indicating that the experimental conditions were adequate to both induce hypoxic cellular responses and differential regulation by NPAS3 and ARNT. *ANKRD37* expression was not observed to be up-regulated in the 3% oxygen culture, nor was any significant down-regulation of *ANKRD37* observed with expression of *NPAS3* and *ARNT* (Figure 5-2B).

Down-regulation of *TXNIP* by hypoxia is caused by metabolic response to hypoxic environment, whereas *ANKRD37* induction has been shown to be caused by stabilization of HIF1A (Benita et al. 2009; Chai et al. 2011). The down-regulation of *TXNIP* in hypoxia has been shown to be transient, lasting less than 24 hours before expression is returned to normal levels and eventually increasing due to HIF1A stabilization (Baker et al. 2008; Chai et al. 2011). As we observe consistent down-regulation of *TXNIP* at 2 and 24 hours, as well as no induction of *ANKRD37*, this may suggest that incubation in 3% oxygen is not sufficient for stabilization of HIF1A in our hands. In order to specifically probe *ANKRD37* regulation with response to HIF1A, HIF1A was stabilized by addition of 100 μM $CoCl_2$, where stabilization was observed by western blot (Figure 5-2D). *NPAS3*, *HIF1A* and *ARNT* were singly and co-expressed in all combinations and cells were treated with $CoCl_2$ for 24 hours to stabilize HIF1A. *ANKRD37* was not found to be significantly up-regulated with $CoCl_2$ treatment, with or without co-expression of HIF1A (Figure 5-2C). Furthermore, the trend of the effect of *NPAS3* expression was up-regulation, rather than the down-regulation observed in the original study identifying *ANKRD37* as a NPAS3-regulated gene. These data suggest a confounding variable whereby the gene regulatory response of our cells to hypoxia by means of HIF1A stabilization is not consistent with the literature, but the metabolic

response appears intact. As such, I have found that the promoter region of *ANKRD37* is a direct target of NPAS3::ARNT, however, the regulatory consequences require further exploration.

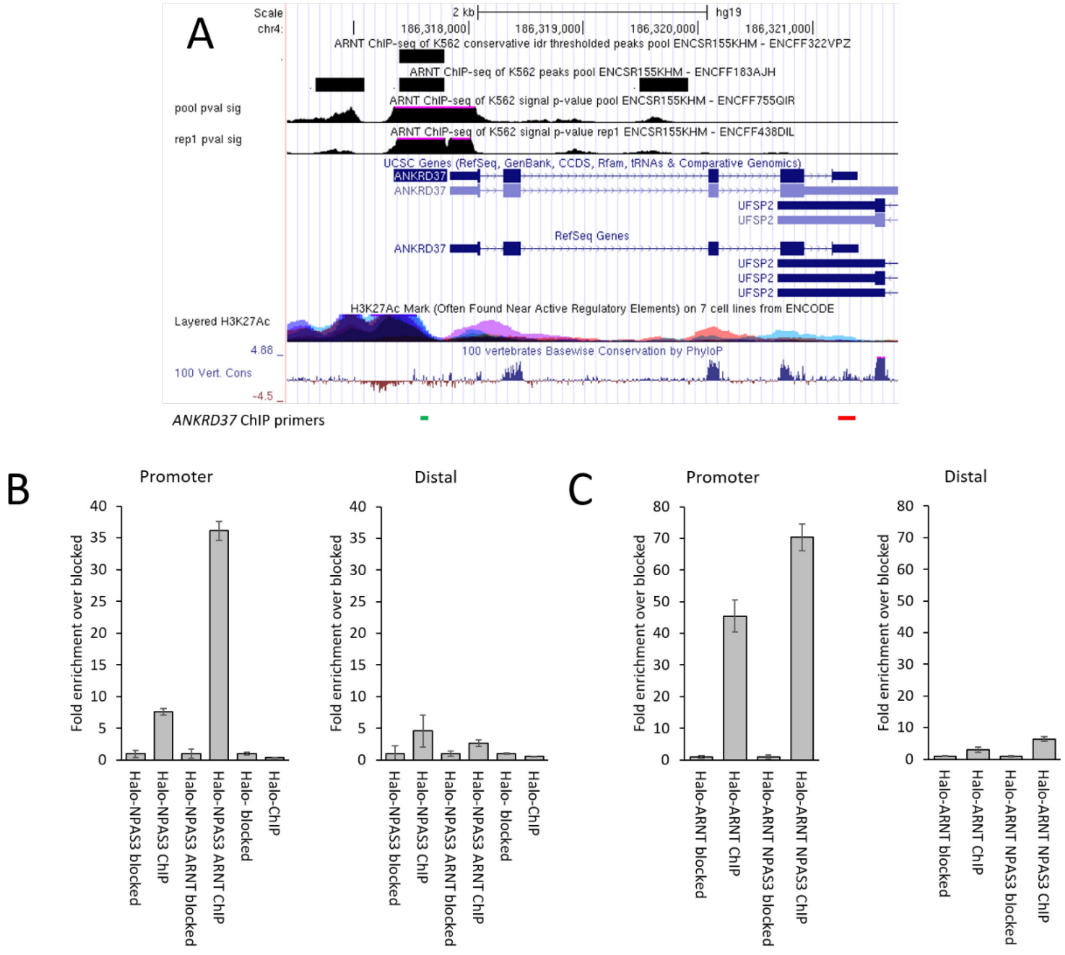



**Figure 5-1: ARNT and NPAS3 bind directly to the *ANKRD37* promoter.**
(A) UCSC genome browser data for the *ANKRD37* locus including ARNT ChIP-seq data used to design primers to assess binding by ChIP at the *ANKRD37* promoter. Available from www.genome.ucsc.edu. (B) Halo-NPAS3 ChIP qPCR data for the *ANKRD37* proximal promoter in normoxia. (C) Halo-ARNT ChIP qPCR data for the *ANKRD37* locus in normoxia. Error bars indicate standard deviation.

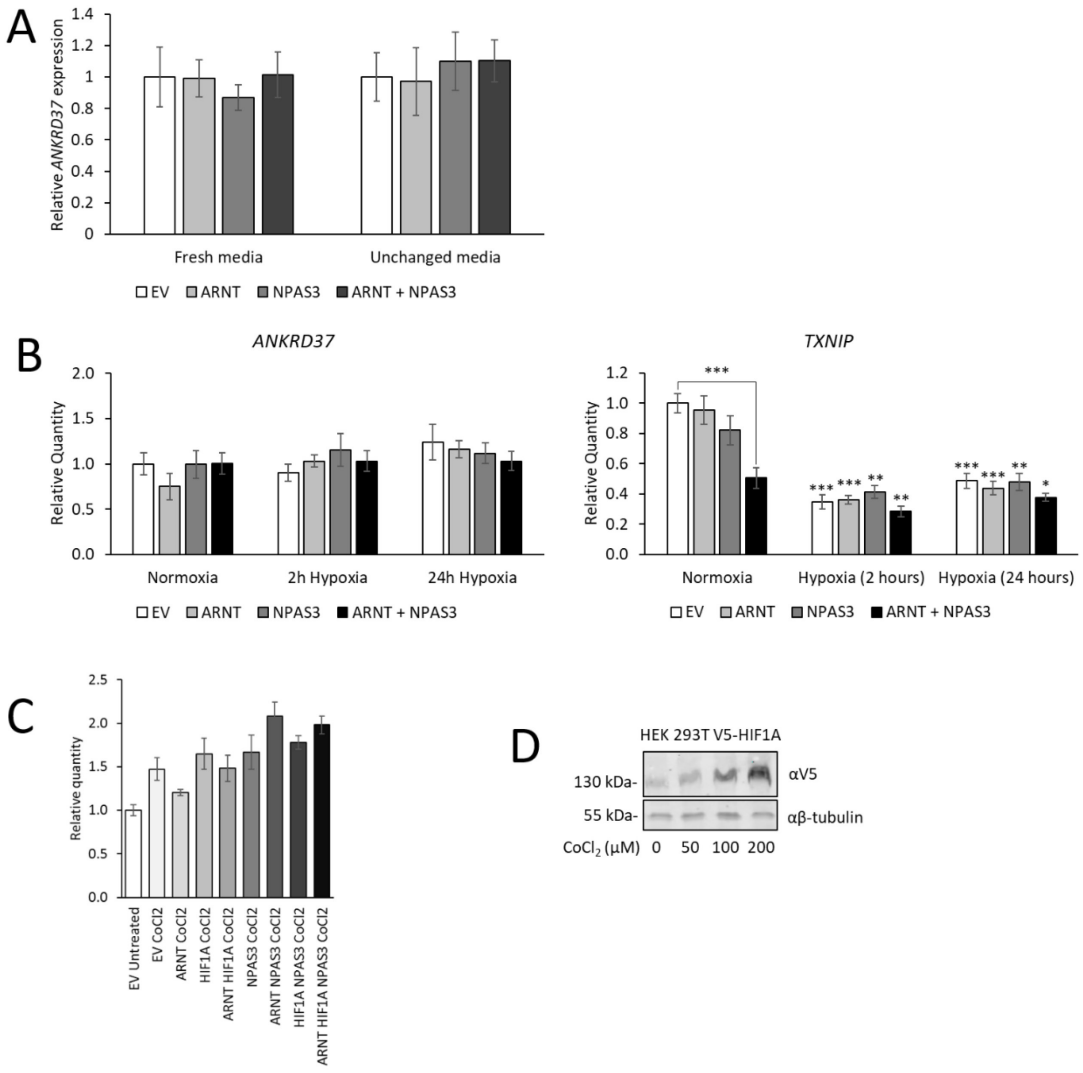



**Figure 5-2: *ANKRD37* is not differentially regulated by NPAS3, HIF1A or ARNT in any hypoxia-like condition tested.**
(A) qPCR data showing no effect of *ARNT* and *NPAS3* expression and media change on *ANKRD37* expression. (B) qPCR demonstrating no increase in ANKRD37 expression by culture in hypoxia (3% O2) or response to NPAS3 or ARNT expression, despite observation of the expected down-regulation of *TXNIP* relative to cells expressing the same construct(s) in normoxia. Asterisks above hypoxia data indicate significant down-regulation respective to the same condition cultured in normoxia. (C) qPCR data showing no significant effect of *ARNT*, *HIF1A* or *NPAS3* overexpression with 100 μM CoCl2 induced stabilization of V5-tagged HIF1A to mimic hypoxic gene regulation, despite stabilization of HIF1A indicated by the western blot in (D). Error bars indicate standard deviation. *$p$<0.05, **$p$<0.01, ***$p$<0.001

## 5.3 Glucose, a driver of *TXNIP* expression

*TXNIP* expression is differentially regulated in low and high glucose culture by carbohydrate response elements (ChoREs) localized within the ARNT ChIP-seq peak and proximal to the region that NPAS3 and ARNT directly bind, as assayed by HaloCHIP (Figure 4-4) (Yu and Luo 2009). Furthermore, in the reporter construct by which we found NPAS3 and ARNT are able to affect *TXNIP* promoter driven regulation, the glucose response elements are encoded and functional (Figure 4-5). As such, I assessed the effects of environmental glucose concentration on the endogenous expression of *TXNIP* and *NPAS3* in HEK 293T cells. qPCR analysis demonstrated that *TXNIP* expression was up-regulated in high (25 mM) glucose relative to low (5.5 mM) glucose, and reporter expression driven by the *TXNIP* promoter construct was also found to be higher in high glucose, as compared to low glucose (Figure 5-3, Figure 4-5 A vs B). *NPAS3* expression was found to be up-regulated after 96 hours of culture in both low and high glucose. *TXNIP* expression was found to be down-regulated after 48 hours in low glucose, with over 2-fold down-regulation at 96 hours. These observations are consistent with the increase in *NPAS3* expression and its known role as a repressor of *TXNIP* expression. In high glucose conditions *TXNIP* expression was found to be markedly up-regulated after 72 hours in culture, along with increased expression of *NPAS3*, which is in disagreement with the expected repression if NPAS3 is involved in regulating endogenous *TXNIP*.

Differential regulation of *TXNIP* by NPAS3 may be a result of differential protein expression and/or varying nuclear localization of NPAS3 in high and low glucose conditions. To assess whether NPAS3 protein is differentially expressed or localized to

the nucleus, cytoplasmic/nuclear fractionation followed by western blotting was performed. Multiple bands were detected in the endogenous nuclear lysates, complicating interpretation of data. The expected size of NPAS3 (933 aa isoform) based on protein sequence is 101 kDa. However, the estimated size from western blot analysis of expressed *NPAS3* constructs is closer to 130 kDa, as such, bands present in the range of 100-130 kDa were assessed for relative expression. *NPAS3* showed no differences in levels at 96 hours in both low and high glucose, nor in localization (Figure 5-3 C). NPAS3 was consistently detected exclusively in the nuclear fraction. ARNT was consistently observed in both the cytoplasm and nucleus, with predominantly cytoplasmic localization. Visual inspection of the cells suggested a potential difference in growth rate, however, growth curves in low and high glucose conditions were found not to be significantly different (Figure 5-3D).

As the relationships between *NPAS3* and *TXNIP* expression are opposed in low (*NPAS3* and *TXNIP* expression discordant) and high (*NPAS3* and *TXNIP* expression concordant) glucose, environmental variables were considered as potential confounders affecting their expression. Cell density has been found to affect the localization of AhR and the regulation of AhR::ARNT target promoters, where AhR is exported from the nucleus in response to cell::cell contact (Ikuta, Kobayashi, Kawajiri 2004). Further, ARNT localization to the nucleus has been shown to be stabilized by nuclear localization of heterodimeric partners (Chilov et al. 1999). Consistent with this finding, I have observed that nuclear localization of NPAS3 is enhanced by expression of ARNT (Figure 4-13Figure 4-14A). In the initial glucose experiments, ARNT was localized predominantly in the cytoplasm in both high and low glucose conditions at confluence

(Figure 5-3C). As such, I explored the effect of cell density for effects on *NPAS3* expression, NPAS3 and ARNT localization, and regulation of *TXNIP.* Effects of glucose on secreted factors were also considered, as increased *TXNIP* expression in response to high glucose has been shown to cause activation of the NLRP3/NALP3 inflammasome and cause secretion of pro-inflammatory molecules, such as the cytokines IL-1β and IL-6, and increased rates of apoptosis (Gao et al. 2015; Zhou et al. 2010). These soluble factors may, in turn, affect the expression of *NPAS3* and *TXNIP*.

In order to assess the contribution of soluble factors independently of cell density, media conditioning experiments were performed to determine if the regulation of *NPAS3* and *TXNIP* over time in culture is related to either cell contact or soluble factors or both. Media from 96 hour cultured HEK 293T cells in low or high glucose were collected, cells removed by centrifugation, and media were directly applied to cells plated for 24 hours in the same glucose condition and incubated for another 24 hours to allow for cellular response, but not for the cells to reach confluence or for the observed regulation of *NPAS3* or *TXNIP* starting at 72 hours (Figure 5-3B) to occur. In low glucose conditions *NPAS3* was found to be up-regulated at 96 hours in culture, when cells were confluent, and only weakly affected by treatment of sub-confluent cells with media from confluent cells for 24 hours (Figure 5-4). A minor effect of conditional media was observed on *TXNIP* expression in subconfluent cells, however no strong effect was observed in the low glucose condition, consistent with previous findings.

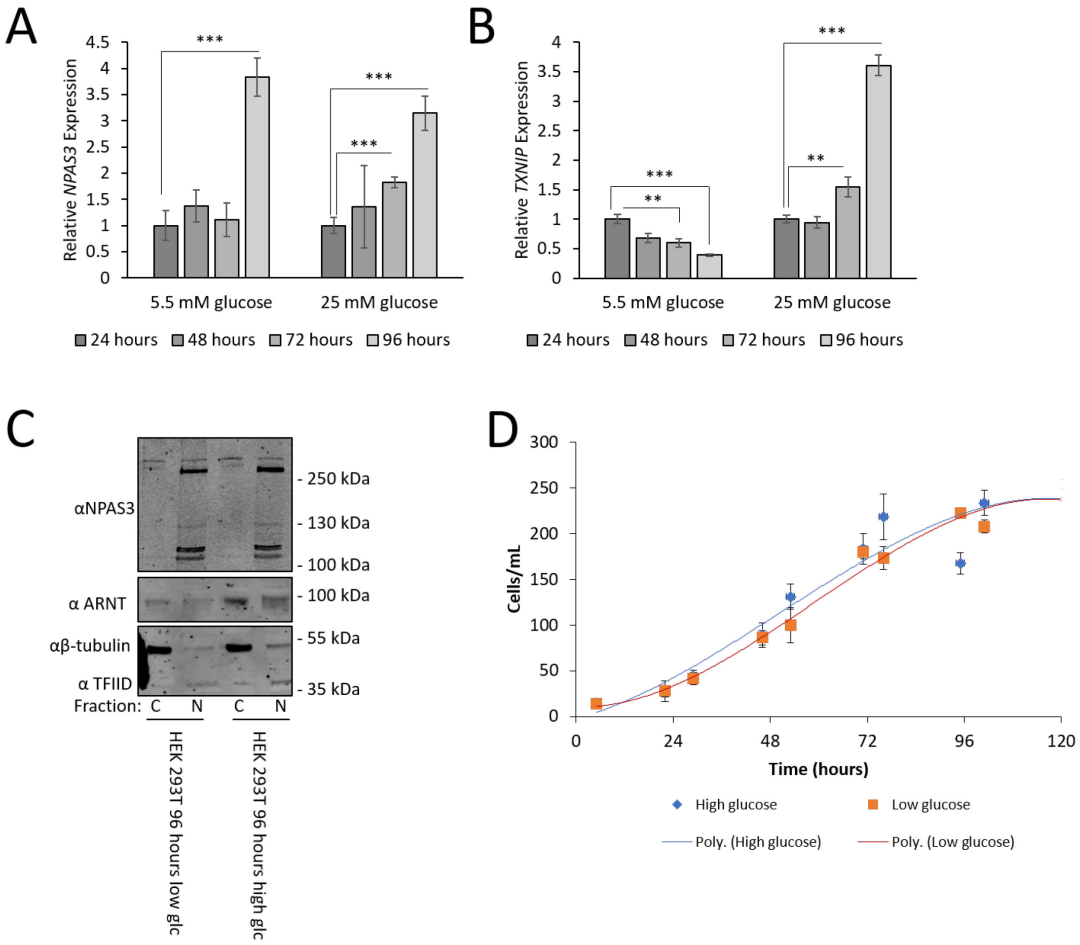



**Figure 5-3: *TXNIP* and *NPAS3* expression are differentially regulated in low and high glucose.**
(A) qRT-PCR showing the expression of *NPAS3* increasing from 24-96 hours in both low (left) and high (right) glucose. (B) qRT-PCR data showing the expression of *TXNIP* over time in low glucose media decreases, and increases in high glucose media. (C) Western blot showing that NPAS3 protein expression in low and high glucose are the same, however ARNT protein is increased in high glucose. (D) Growth curve of HEK 293T cells demonstrating no significant difference in growth rate or cell density. *$p$<0.05, **$p$<0.01, ***$p$<0.001

In high glucose conditions, *NPAS3* was up-regulated at 96 hours when cells were confluent, and also at 48 hours in culture after 24 hours treated with media harvested from cells cultured for 96 hours, suggesting that, in the high glucose condition, soluble factors affect the regulation of *NPAS3*. Similarly, *TXNIP* was up-regulated to the same level in the confluent 96 hour culture, as well as the sub-confluent 48 hour culture treated with media from HEK 293T cells grown in high glucose conditions, showing that its regulation is primarily affected by soluble environmental factors.

Western blots of nuclear and cytoplasmic lysates did not detect any difference in NPAS3 or ARNT protein expression or localization in any condition (Figure 5-4B). These data suggest that *NPAS3* mRNA expression is regulated by soluble factors released under certain conditions, as well as cell density. However, as differences in protein expression were not observed, conclusions cannot be made about whether endogenous NPAS3 protein is differentially expressed, or whether it affects the observed regulation of *TXNIP*.

As these data demonstrate regulation of *NPAS3* at the mRNA level, but it is unclear if NPAS3 protein is differentially expressed in HEK 293T cells, the line of experimentation was switched to SK-N-SH neuroblastoma cells. SK-N-SH were selected as they are of a neural lineage and thus more representative of the biological system involved in neurogenesis and the known functions of NPAS3.

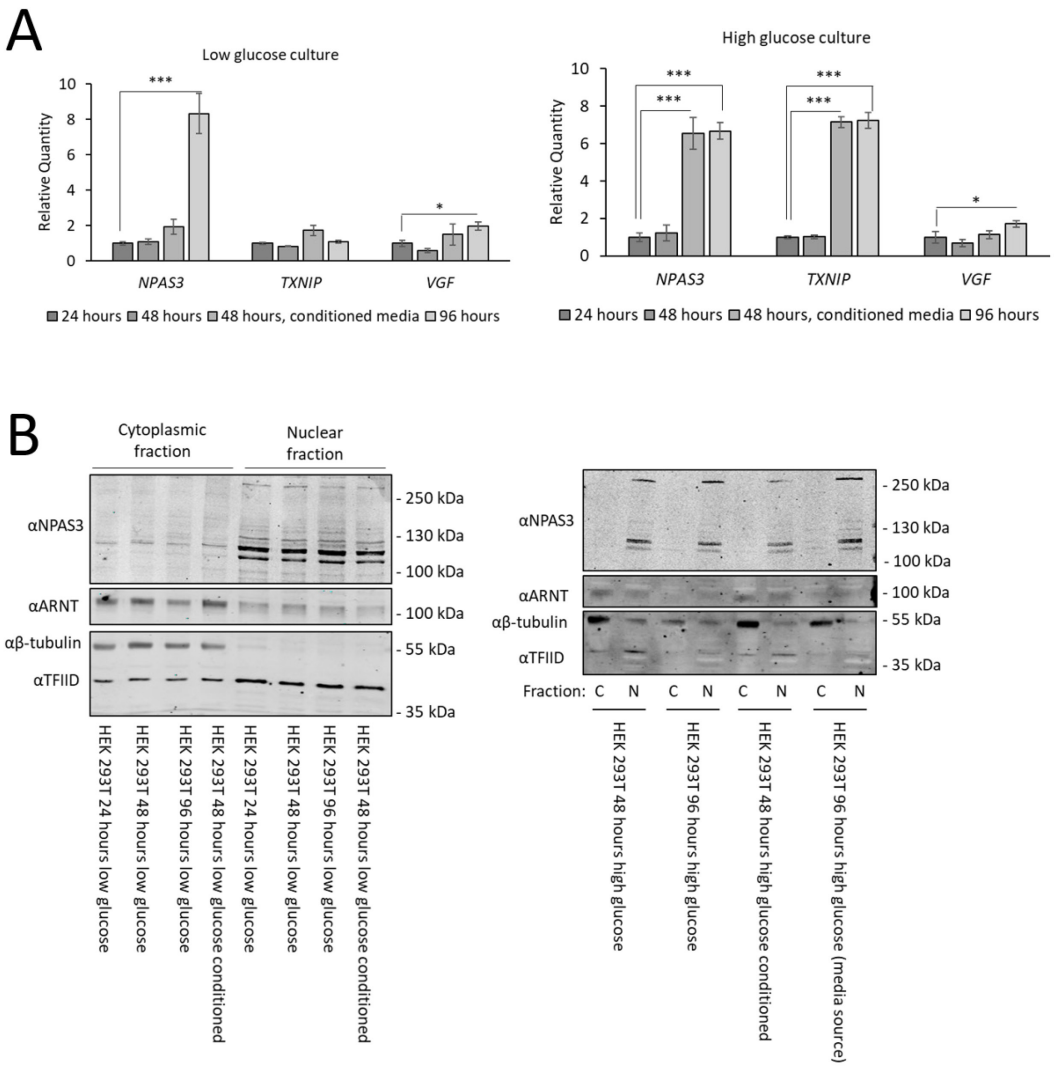



**Figure 5-4: NPAS3 is regulated by both soluble and cell density related signals.**
(A) qRT-PCR showing expression of *NPAS3* increasing from 24-96 hours in both low (left) and high (right) glucose. Media taken from 96h plated cells ("conditioned") and added to cells at 24 hours for 24 hours resulted in non-confluent cells at high glucose up-regulating *NPAS3* expression. (B) Western blot showing that mRNA expression does not lead to a coordinate increase in protein expression or change in localization. **p*<0.05, ***p*<0.01, ****p*<0.001

## 5.4 Characterization of SK-N-SH as an *NPAS3* expressing cell line

SK-N-SH cells were found to express NPAS3 at the mRNA level by RT-PCR (Figure 5-5A), which was confirmed by Sanger sequencing. To assess whether NPAS3 protein is expressed, two independent cultures of SK-N-SH cells were assessed by western blotting using the ProSci αNPAS3 N-ter antibody. This antibody detected protein at 100 kDa as well as a band at approximately 130 kDa, both may be NPAS3 (Figure 5-5B). These bands are not fully consistent with the expected size from the antibody product information sheet (Figure 5-5C), and with our western blot data of expressed NPAS3 constructs using our custom NPAS3 antibody (Figure 4-3). Immunofluorescence analysis using the commercial antibody demonstrated both cytoplasmic and nuclear localization (Figure 5-5D). As SK-N-SH cells were found to be positive for NPAS3 at both the RNA and protein level, they were selected for further analysis of the function of endogenously expressed *NPAS3*.

## 5.5 Culture conditions

As *NPAS3* was found to be differentially regulated at the mRNA level in HEK 293T cells over time in culture, the expression of *NPAS3* in response to growth factors identified as of interest in HEK 293T cells were explored in SK-N-SH cells at analagous time points in culture, based on a comparison of growth curves (Figure 5-3D, Figure 5-6B). Cells were assessed with and without media replacement (Figure 5-6). SK-N-SH cells were first assessed at 120 hours in culture, when they are entering stationary phase/quiescence based on the growth curves, and confirmed by decreased expression of G2/M phase markers of proliferative cells (*CCNB1* and *AURKA*, Figure 5-6C) (Aviner et

al. 2015). SK-N-SH cells at 120 hours did not appear to be apoptotic, given the relative increase in anti-apoptotic *BCL2* expression over the pro-apoptotic *BAX* (Minn, Hafele, Shalev 2005). Neither the nuclear localization, nor amount of NPAS3 protein were found to be affected by media change at this time point, although in the first replicate experiment a band was observed in the unchanged media 120 hour nuclear lysate and not the fresh media 120 hour nuclear lysate (Figure 5-6A, ~120 kDa band). In this replicate, the expression of *TXNIP* was found to be decreased, and *VGF* was found to be increased relative to the fresh media condition, where this band is absent (Figure 5-6D). This result was initially suggestive of NPAS3 acting at this time point. However, a second replicate of this experiment resulted in no 120 kDa protein being detected and no difference in expression of *TXNIP* or *VGF*. These data do not exclude the possibility that NPAS3 may be present and acting in SK-N-SH cells at this time point, however, the number and variability of bands detected by the NPAS3 antibody prevents full interpretation of these results.

In order to try to reduce the number of confounding variables, such as the media and serum components, instead of media replacement at 96 hours, cells were treated with mouse recombinant bFGF (FGF2) and heparin. bFGF was chosen because *NPAS3* expression has been shown to be critical for FGF-mediated adult neurogenesis in mice (Pieper et al. 2005). As can be seen in Figure 5-7, no response was observed in *NPAS3* mRNA, or protein expression. The response of the cells was assayed by the increase in the proliferative markers *CCNB1* and *AURKA*, suggesting that the cells were stimulated to grow.

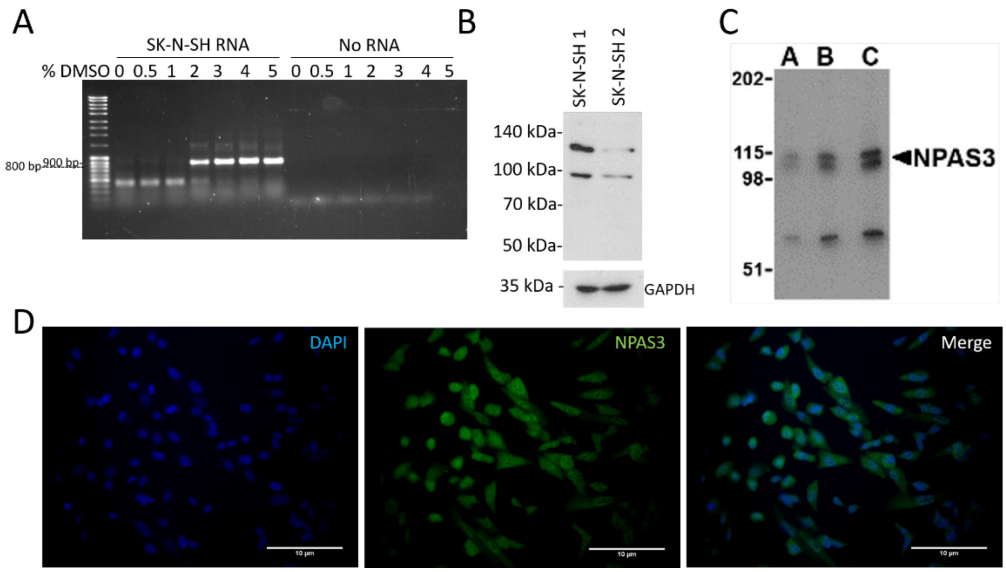



**Figure 5-5: SK-N-SH is an *NPAS3*-expressing cell line.**
(A) Agarose gel showing one-step RT-PCR performed on SK-N-SH cell total RNA preps, where the expected 861 bp band was detected. (B) Western blot on two SK-N-SH cell lysates using the NPAS3 N-ter antibody (ProSci) which detects bands at 100 kDa and 130 kDa relativeto the pattern in the technical insert in (C) where lanes A-C are 0.5, 1 and 2 µg/ml SK-N-SH cell lysate. (D) Immunofluorescence of SK-N-SH cells with the NPAS3 N-ter antibody (ProSci) showing both cytoplasmic and nuclear signal. 400X magnification, scale bar: 10 µm.

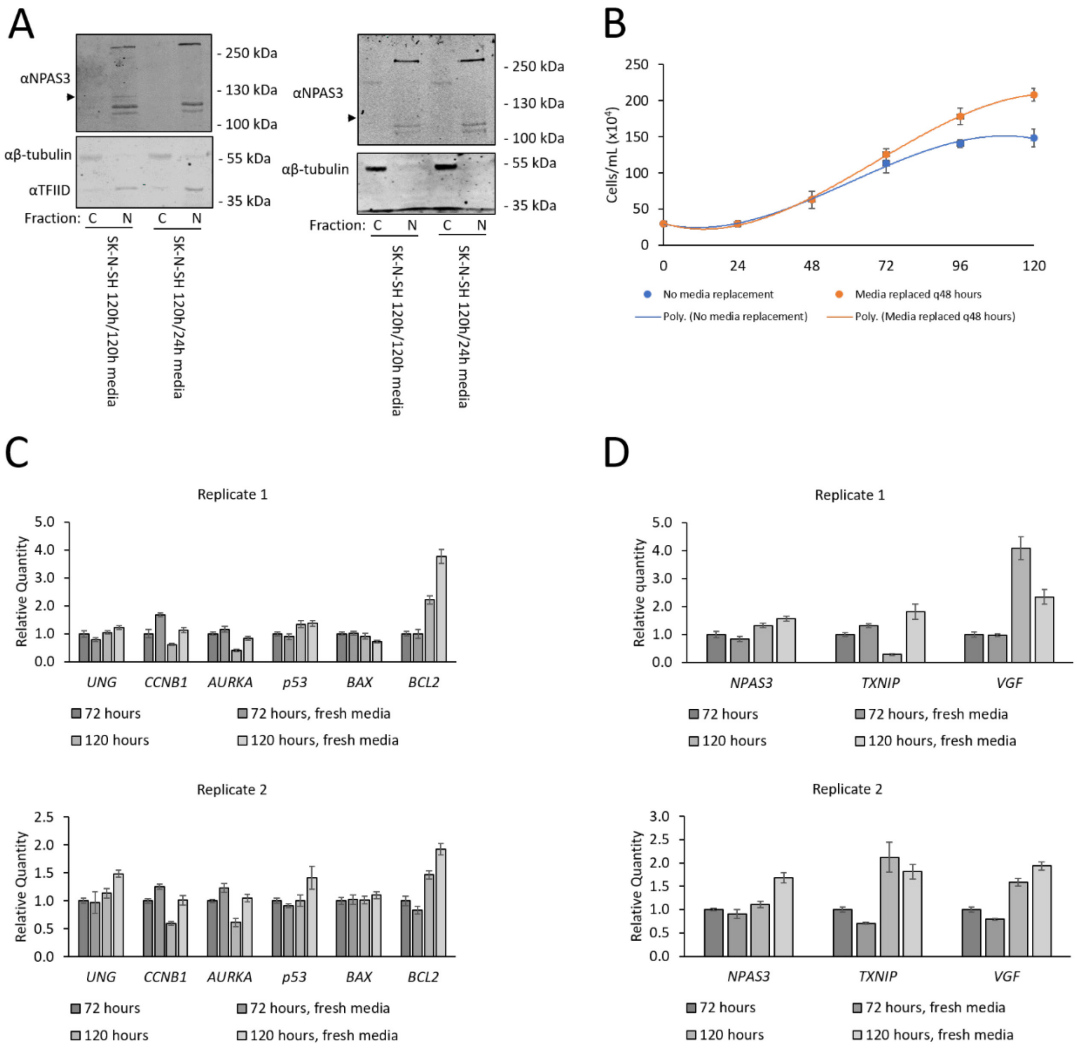



**Figure 5-6: SK-N-SH growth parameters**
(A) Western blots of nuclear and cytoplasmic lysates of SK-N-SH cells with fresh media and unchanged media. Arrowhead indicates band running at the approximate size of the NPAS3 933 aa clone. (B) Growth curve of SK-N-SH cells with and without media replacement every 48 hours. (C) qPCR data showing that as cells become confluent markers of G2/M (*CCNB1*, *AURKA*) decrease, and anti-apoptotic marker *BCL2* increase suggesting they are becoming quiescent. (D) qPCR data showing that *NPAS3* expression is weakly increased in quiescent cells (120 h) as compared to mid-log phase cells (72 h). Downstream targets of NPAS3 were found to respond differently between replicates.

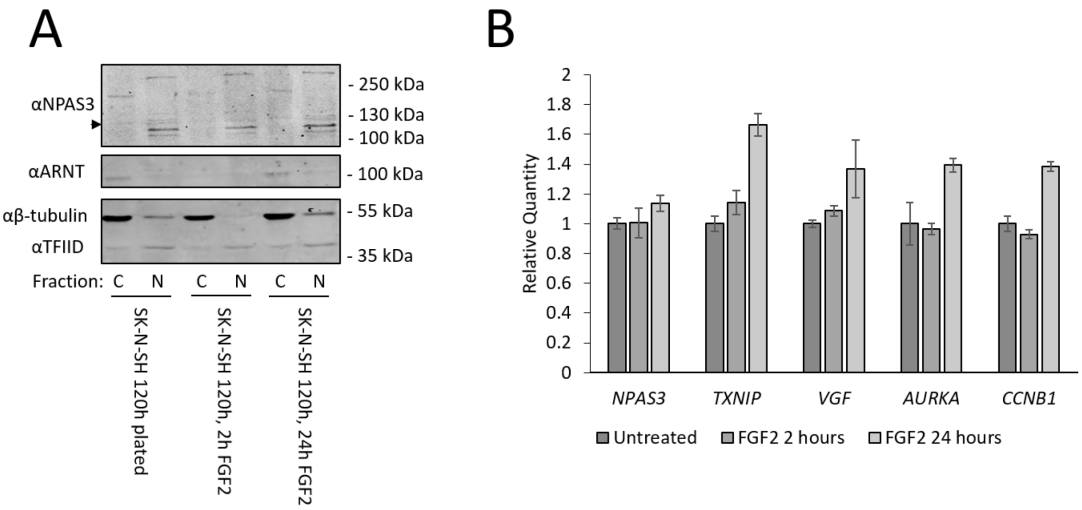



**Figure 5-7: NPAS3 expression is unaffected by treatment with bFGF.** (A) Western blot showing no change in NPAS3 protein detection with treatment with mouse recombinant bFGF and heparin for 2 or 24 hours. (B) qPCR data showing weak increase in markers of proliferation and no change in *NPAS3* expression or regulation of target genes *TXNIP* and *VGF* in a direction suggestive of NPAS3 function.

## 5.6 Circadian cycle

A subset of genes regulated by NPAS3 are differentially regulated over the circadian cycle, although *NPAS3* was not found to be regulated in a circadian manner at the two timepoints (+12 hours and +24 hours) assessed (Sha et al. 2012). Whether the observed differential regulation is due to NPAS3 interacting with a circadian partner, such as the general class 2 bHLH-PAS circadian partner BMAL1, or is due to regulation of NPAS3 targets by circadian factors independently, is unclear (McIntosh, Hogenesch, Bradfield 2010). We induced circadian cycling in SK-N-SH cells to confirm whether NPAS3 is expressed in a circadian manner. Induction of circadian cycling was deemed successful, as evidenced by rhythmic expression of the circadian genes *BMAL1*, *CRY1*, and *NPAS2* (Figure 5-8A) (Balsalobre, Damiola, Schibler 1998; Dudley et al. 2003; McNamara et al. 2001). *NPAS3* and *ARNT* were both found not to be expressed differentially over the circadian cycle. As non-circadian bHLH-PAS proteins can be hypothesized to interact with circadian bHLH-PAS proteins, the ability for NPAS3 to interact with BMAL1 was assessed. NPAS3 was able to weakly interact with BMAL1 in one replicate of the HaloTag pull down assay (Figure 5-8 B), however this result was not replicated and interaction with BMAL1 cannot be concluded or excluded.

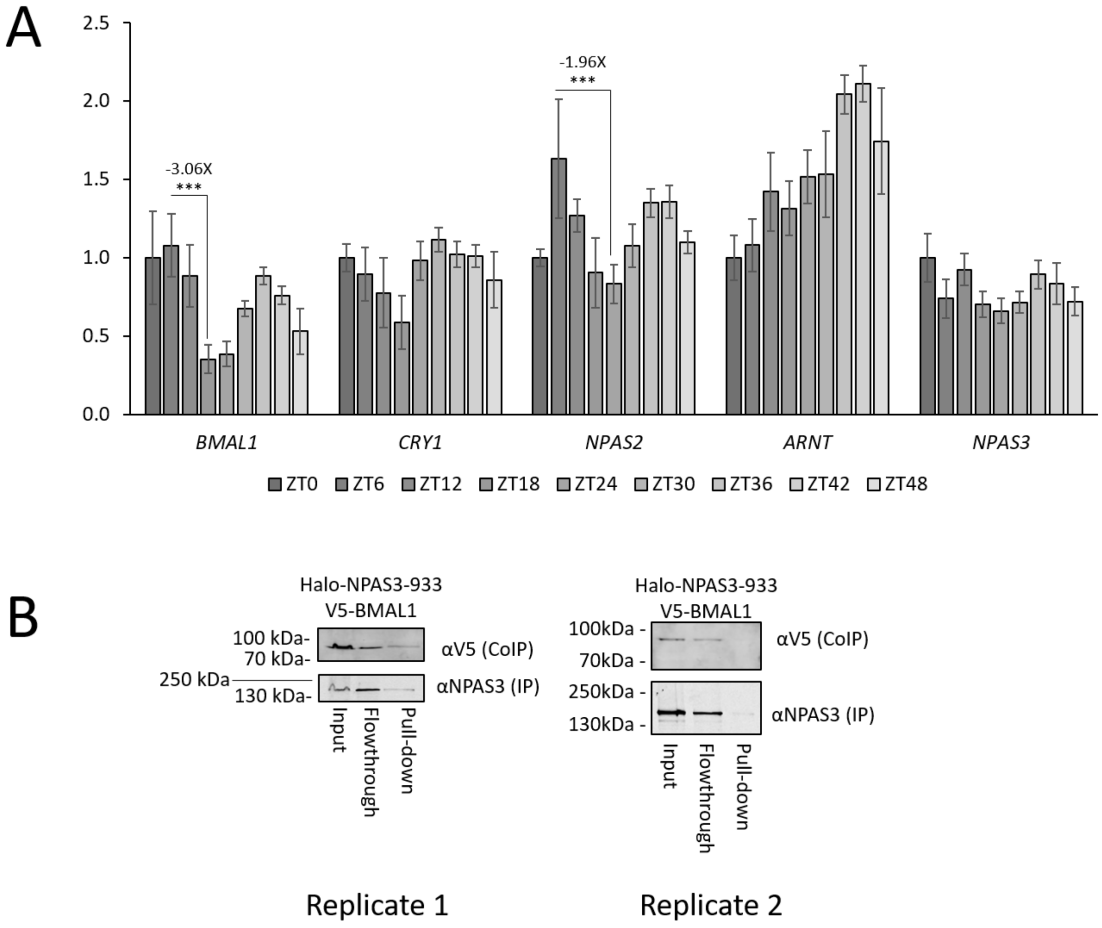



**Figure 5-8: *NPAS3* is not rhythmically expressed in response to the circadian cycle.**
(A) qPCR data on SK-N-SH cells treated with horse serum to induce circadian cycling, showing rhythmic expression of circadian genes, but not the non-circadian *ARNT* or *NPAS3*. (B) HaloTag pull downs of BMAL1 by NPAS3 were inconsistent. **p*<0.05, ***p*<0.01, ****p*<0.001

## 5.7 shRNA knockdown of *NPAS3*

In order to assess the role of endogenous *NPAS3,* cell lines stably expressing shRNAs to *NPAS3* were developed to knockdown *NPAS3* expression. Three shRNAs from The RNAi Consortium (TRC) (Moffat et al. 2006) were obtained and packaged into lentiviral vectors for transduction of SK-N-SH cells at high multiplicity of infection (MOI). Cells were pooled and plated in a new dish and selected for puromycin resistance. For experiments cells were plated with and without puromycin, as indicated. Western blots did not demonstrate clear knockdown at the protein level (Figure 5-9A, Figure 5-10A, Figure 5-11A), however one shRNA construct (TRC76, Table 6) resulted in consistent down-regulation of *NPAS3* mRNA. When plated with puromycin, *NPAS3* was found to be down-regulated with the eGFP (non-targeting) shRNA, which should be non-specific (Figure 5-9). As a slight knockdown of *NPAS3* was observed over the non-specific shRNA, with a coordinate response of *TXNIP,* as well as a response of *VGF*, albeit in the opposite direction of what was expected, this construct was probed further. Cells plated without puromycin resulted in normal levels of *NPAS3* mRNA expression with the non-specific shRNA and a clear 50% knockdown was consistently observed with the TRC76 *NPAS3*-specific shRNA (Figure 5-10). No difference was observed on western blots, however downstream genes were probed and inconsistent results were obtained. *VGF* was found to be up-regulated in one replicate experiment, however no effect was observed in the second replicate.

In addition to the shRNAs from the RNAi consortium, we assessed four shRNAs from the Origene HuSH *NPAS3* panel (Origene) in order to replicate our findings. Pooled clonal stable integrants were tested without puromycin for experiments. No

knockdown was observed by western blot of nuclear lysates, or at the RNA level by qPCR (Figure 5-11).

As the best evidence for knockdown was found for *NPAS3* shRNA TRC76, these samples were probed further with a panel of genes identified as direct targets of Npas3 in the hippocampus of mice, which was recently published (Michaelson et al. 2017). The direct targets of NPAS3 that were screened include: *FMR1*, *NOTCH1*, *NOTCH2* and *UBE3A*. These four genes were all found to be directly bound by Npas3, differentially regulated in hippocampi of $Npas3^{-/-}$ mice and have functions relevant to neurogenesis and/or neuropsychiatric disorders (Michaelson et al. 2017). Figure 5-10 shows that up-regulation of *NOTCH1* and *NOTCH2* was observed in one replicate of the *NPAS3* shRNA TRC76 knockdown study, which is the expected reponse, but not the second. Interestingly, in this replicate, up-regulation of *VGF* was also observed, which is opposite of what is expected. Up-regulation of *VGF* was also observed in the puromycin-treated *NPAS3* shRNA experiments, when a response by *TXNIP* was also observed (Figure 5-9). These preliminary data are intriguing but difficult to interpret, due to the inconsistency of the effects and technical limitations of the western blot precluding determination of relative NPAS3 protein levels. *NPAS3* may be expressed in these cells and acting specifically in response to an uncontrolled variable, resulting in occasional responses, however no conclusions can be drawn.

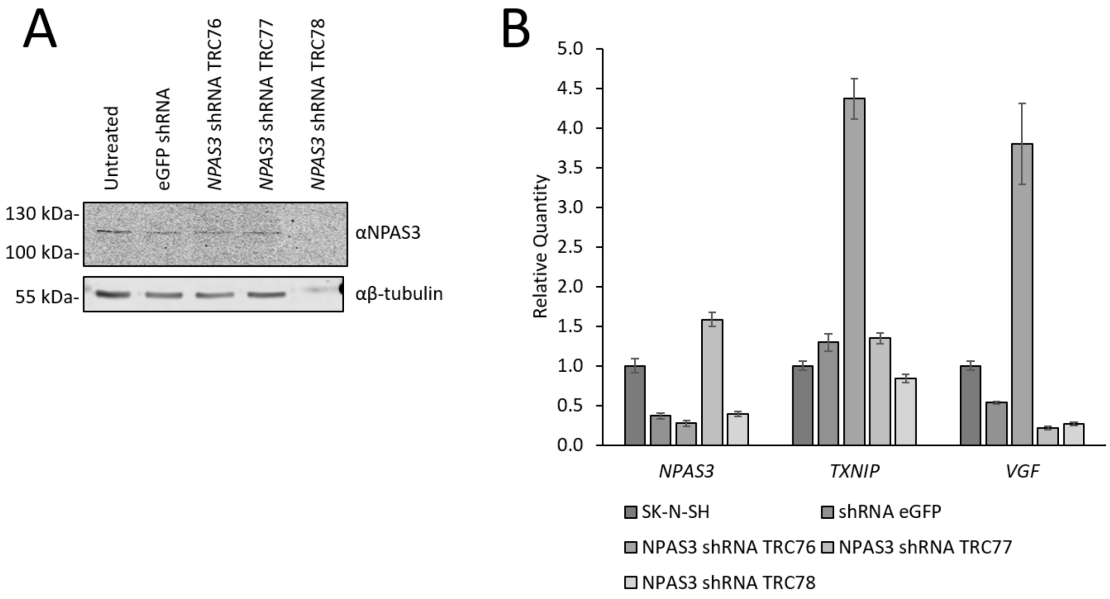



**Figure 5-9: Screening TRC shRNAs for *NPAS3* knockdown with puromycin.**
(A) Western blot of total cell lysates of SK-N-SH cells stably transduced with eGFP (non-targeting) and *NPAS3*-targeting shRNAs. (B) qPCR data showing the results of the pilot shRNA probe.

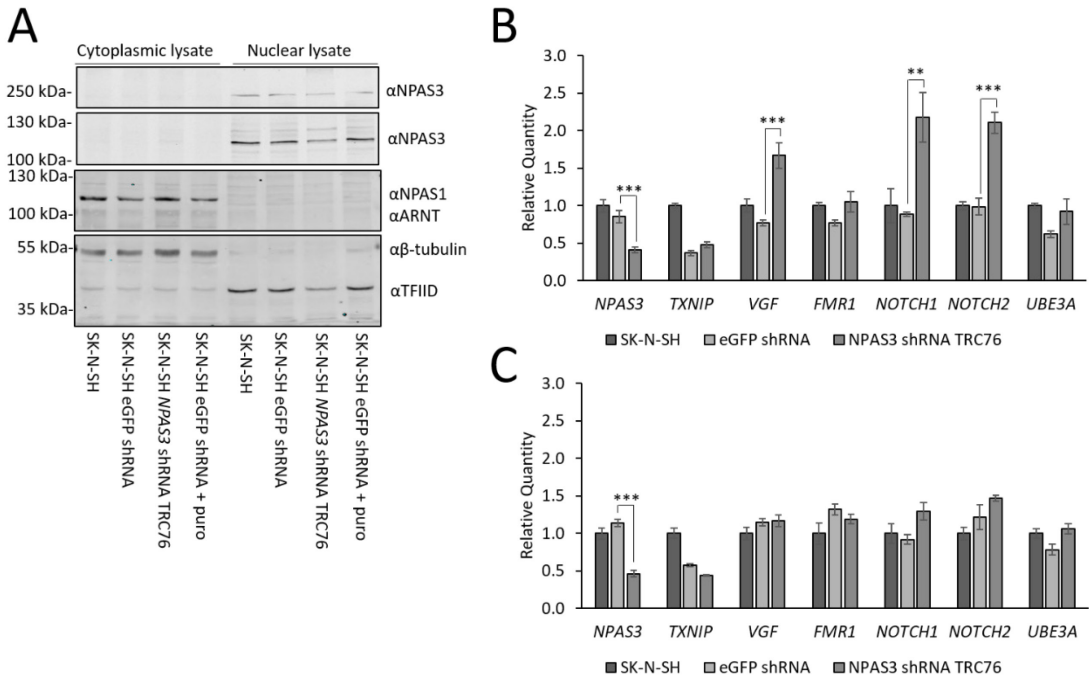



**Figure 5-10: Assessment of TRC shRNAs without puromycin.**
(A) Western blots of SK-N-SH parental lines and the stably transduced eGFP (non-targeting) and *NPAS3* specific shRNA TRC76 showing nuclear bands detected by the NPAS3 antibody and no knockdown of any band relative to the TFIID loading control. (B-C) Two replicate experiments looking at NPAS3 expression and known direct targets with NPAS3 shRNA knockdown. *$p$<0.05, **$p$<0.01, ***$p$<0.001

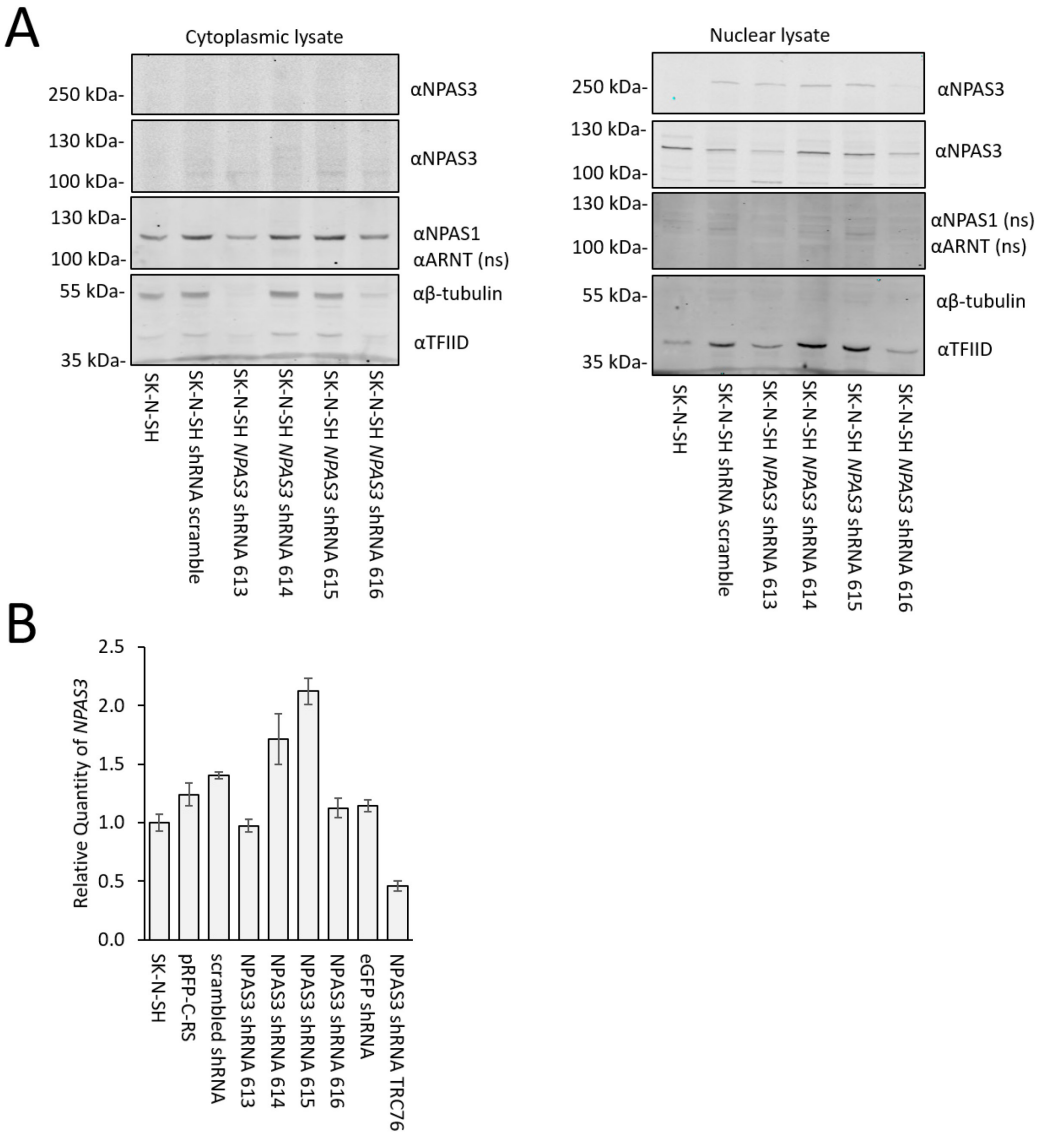



**Figure 5-11: Assessment of knockdown by *NPAS3* shRNAs purchased from Origene.**

(A) Western blots of nuclear and cytoplasmic lysates of the Origene shRNAs (613-616) probed for NPAS3 demonstrated no commonly observed 130 kDa band in most samples, including the shRNA- control. No difference was observed in the expression of the 100 kDa band when normalized to the TFIID band. (B) qPCR data showing no knockdown by the Origene shRNA constructs. Knockdown could not be confirmed in any of these constructs.

## 5.8 Inducible constructs of *NPAS3*

In order to complement our findings using shRNAs to NPAS3, stable integrants of an inducible construct of NPAS3 (pLIX-402-*NPAS3* 933 aa isoform) were generated. NPAS3 was expressed in SK-N-SH cells by treatment with doxycycline on a construct shown to be inducible when transiently transfected into HEK 293T cells (Figure 5-12A). Transient transfections of the HA-*NPAS3* clone, used previously by us for molecular characterization of *NPAS3,* were also used to improve our ability to interpret western blot data. Endogenously expressed NPAS3 protein could not be detected at the size of ectopically expressed HA-NPAS3 933 aa in SK-N-SH (Figure 5-12B,C). A large molecular weight band is resolved above the highest molecular weight marker, but its intensity does not completely correlate with *NPAS3* expression/induction status.

In order to further our understanding of whether the pLIX402-*NPAS3*-933 aa construct is inducible in SK-N-SH cells, immunofluorescence on the parental SK-N-SH cell line and induced and uninduced stably transfected SK-N-SH cell lines was performed. As can be seen in Figure 5-13, *NPAS3* expression as detected by the NPAS3 antibody appears to be increased in the nuclei of SK-N-SH pLIX402-NPAS3 933 aa stable cell lines induced with 1 μg/mL doxycyline, relative to the uninduced or parental cell lines. These data suggest that *NPAS3* expression can be induced in these cell lines and there may be a limitation in the western blots precluding consistent detection of NPAS3 protein in SK-N-SH.

As *NPAS3* can be induced in the stably transduced cell lines, qPCR was performed to confirm this induction and to determine if downstream genes are differentially regulated. Doxycycline treatment of SK-N-SH pLIX402-NPAS3-933 aa

cells was found to significantly increase the expression of *NPAS3* at the RNA level (Figure 5-14A). *TXNIP* was found to not be affected by *NPAS3* induction. *VGF* expression was found to differ in the uninduced and induced SK-N-SH pLIX402-NPAS3-933 aa cells, however the uninduced condition was found to have relatively less expression of *VGF* than the parental cell line and in the induced cell line the expression returned to the parental untransduced cell level. This may be due to up-regulation driven by increased NPAS3 expression, however, the reason for the baseline reduction in VGF expression relative to the parental cell line is unclear.

In summary, the results of experiments on NPAS3 function in SK-N-SH cells are inconclusive, with many possible explanations. Expression of cloned *NPAS3* by transient transfection, as well as inducible constructs, suggest that there may be a technical issue with western blotting precluding consistent detection of NPAS3. This may be due to antibody issues, or potentially related to a real biological effect in response to a factor not controlled for in our experiments.

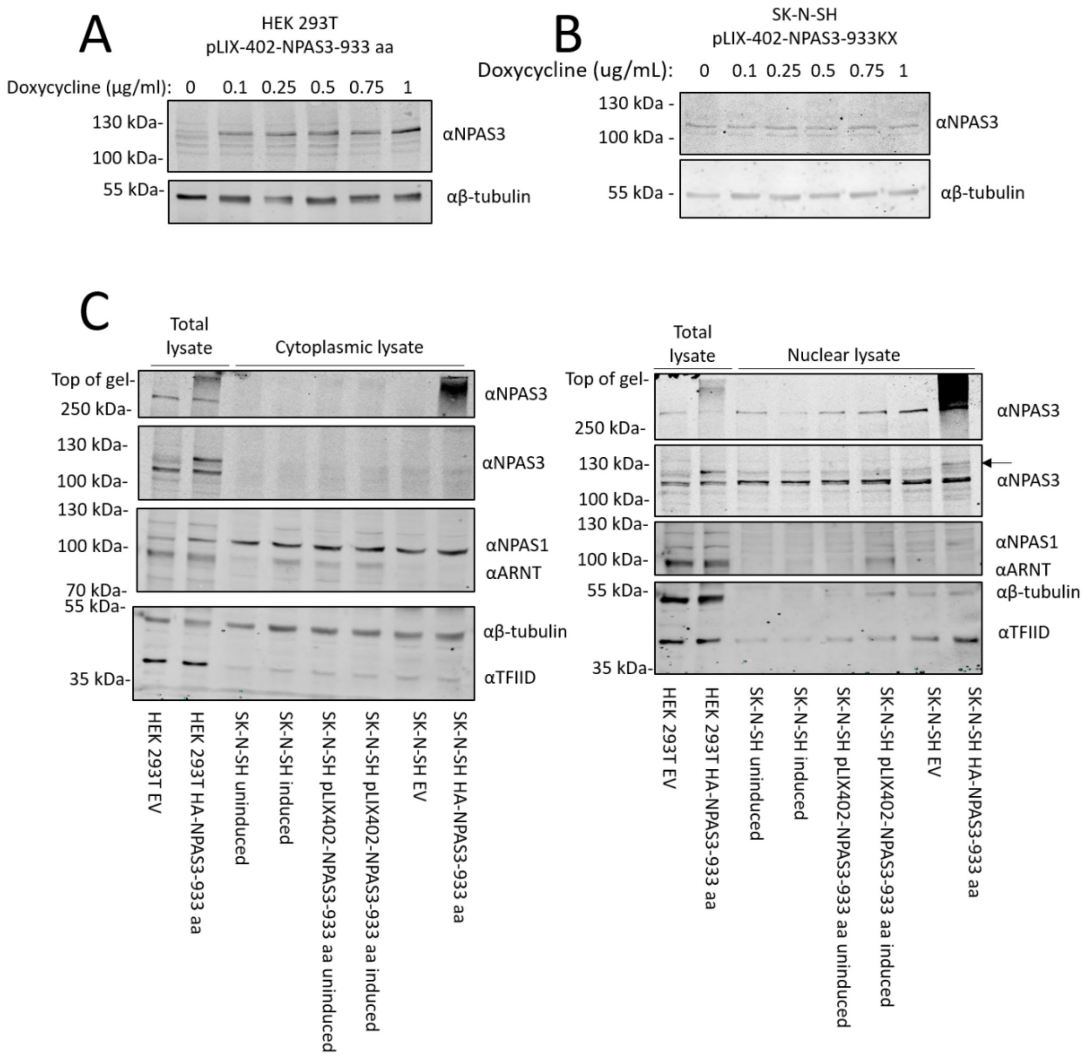



**Figure 5-12: Western blot of NPAS3 induction in SK-N-SH cells**
(A) Western blot of HEK 293T cells transiently transfected with pLIX-402-NPAS3 933 aa, followed by induction with doxycycline, showing that the construct can be induced. (B) Western blot of total cell lysates of SK-N-SH cells stably transduced with pLIX-402-NPAS3 933 aa, induced with the same range of doxycycline, with no apparent induction. (C) Cytoplasmic and nuclear lysates of the untransduced SK-N-SH cells and SK-N-SH cells stably transduced with pLIX-402-NPAS3 933 aa, both uninduced or induced with 1 µg/mL doxycycline, showing no apparent induction. Total cell lysates of HEK 293T cells transfected (left two lanes) with empty vector or HA-NPAS3-933 and SK-N-SH transfected with empty vector or HA-NPAS3-933 (right two lanes) are included for size standards in order to determine which band(s) represent NPAS3. Arrow indicates 120 kDa band.

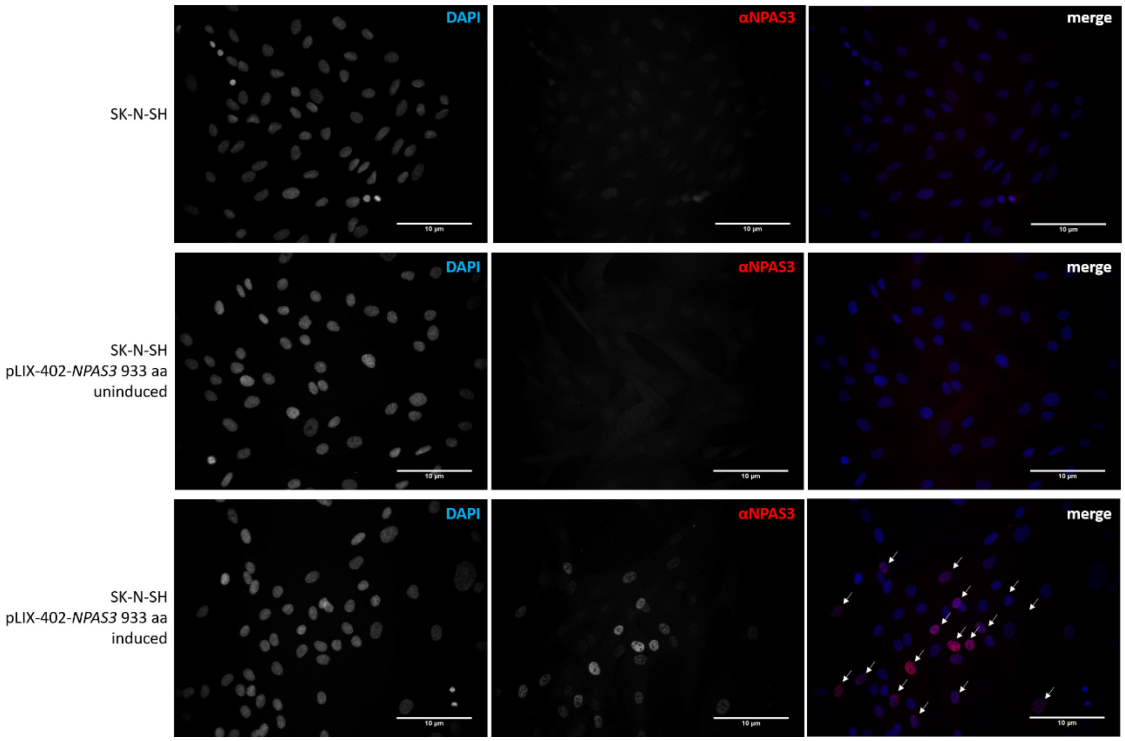



**Figure 5-13: Immunofluorescence showing inducibility of the pLIX-402-NPAS3 933 aa construct**

Parental SK-N-SH cells, and uninduced and induced (1 µg/mL) SK-N-SH pLIX402-*NPAS3*-933 aa cells were probed with αNPAS3 to determine whether *NPAS3* expression was increased with doxycycline treatment. Arrows: nuclei with increased NPAS3 signal. Magnification: 400X, scale bar: 10 µm.

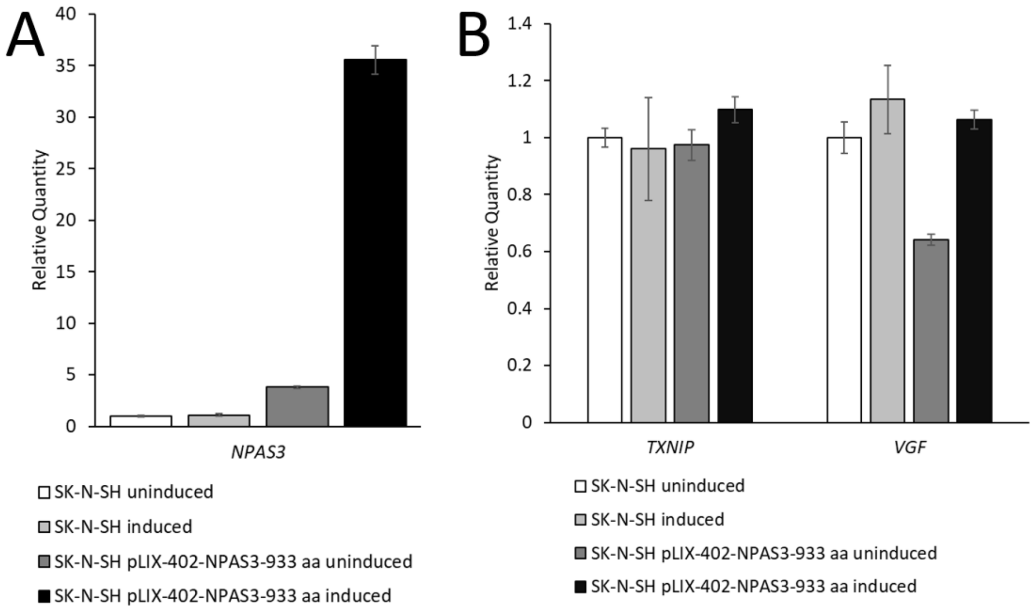



**Figure 5-14: NPAS3 is induced in SK-N-SH pLIX402-NPAS3-933 aa but has no obvious effect on *TXNIP***

(A) qPCR demonstrates that *NPAS3* is induced by 1 μg/mL doxycycline treatment in SK-N-SH pLIX402-NPAS3-933 aa cells, and not in untransduced SK-N-SH cells. (B) *TXNIP* was not differentially regulated in doxycycline induced SK-N-SH pLIX402-NPAS3-933 aa cells. *VGF* was found to be differentially regulated between the uninduced and induced SK-N-SH pLIX402-NPAS3-933 aa cells.

## 5.9 ARNT and ARNT2

Western blot analysis of SK-N-SH cells found that the signal of endogenous ARNT was relatively low and primarily cytoplasmic (Figure 5-10A, Figure 5-11A, Figure 5-12C). A relative lack of ARNT, or any other heterodimeric partner, may result in a decrease in NPAS3 functionality, regardless of protein expression. *ARNT2* is an *ARNT* paralogue that has been shown to be expressed reciprocally with *ARNT* in differentiating neurons with *ARNT2* expression increasing as neurons differentiate (Hao et al. 2013). The ratio of *ARNT* to *ARNT2* mRNA expression normalized to β-actin approximated 2 in SK-N-SH cells (1.78, 2.08 in two replicates of SK-N-SH plated independently as negative controls for two experiments), with Ct values in the range of 24 to 25 suggesting that both *ARNT* and *ARNT2* are readily expressed in SK-N-SH cells. However, given western blot results not consistently detecting ARNT, ARNT2 protein should be pursued further, as it is a potential binding partner for NPAS3 and its expression and localization may affect NPAS3 function.

## 5.10 Assessment of ENStem-A human neuroprogenitor cells for *NPAS3* expression

As *NPAS3* is expressed in neuroprogenitors cells in multiple regions throughout the developing and adult CNS, neuroprogenitors are the ideal cell type to assess NPAS3 function (Erbel-Sieler et al. 2004; Gould and Kamnasaran 2011; Kamm et al. 2013b; Pieper et al. 2005; Stanco et al. 2014; Wong et al. 2012). ENStem-A neuroprogenitor cells (Millipore), derived from H9 human embryonic stem cells, were used for

assessment of *NPAS3* expression. *NPAS3* and newly identified target genes of NPAS3 were found to be expressed in the ENStem-A human neuroprogenitor cells at a higher level than in SK-N-SH neuroblastoma cells (Figure 5-15). Unfortunately, the cells did not behave in the expected manner with culture using the manufacturer's protocol which precluded further experimentation as the results may not be reliable. Human neuroprogenitor cell lines remain an exciting cell type with which to explore the function of NPAS3 and should be pursued further. Furthermore, in these cells, *ARNT2* was found to be expressed at a level 3-fold higher than *ARNT*, supporting a possible role for ARNT2 as a potential partner for NPAS3 that should be explored in these cells.

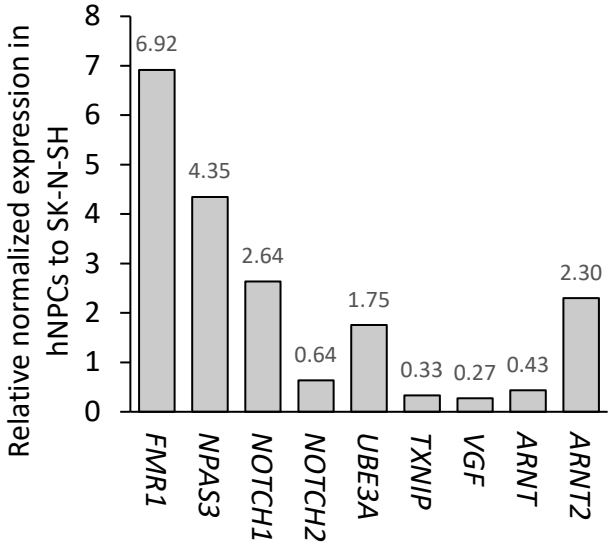



**Figure 5-15: Expression of *NPAS3* and its direct targets in human neuroprogenitors**
The average Ct value of each gene was normalized to β-actin Ct value by the formula Expression=$2^{(Ct\beta\text{-actin-}Ctgene)}$. The ratio of expression observed in ENStem-A was plotted as a ratio of expression in SK-N-SH hNPCs was determined and plotted.

## 5.11 Summary of results

Through these experiments I have identified another direct target of NPAS3 and ARNT, *ANKRD37*, although no regulatory effect was observed, possibly due to insufficient induction of hypoxic gene regulation, or altered hypoxic response in HEK 293T cells. The regulation of *TXNIP* and *VGF* by endogenous levels of *NPAS3* expression was assessed in HEK 293T and SK-N-SH cells. *NPAS3* expression at the mRNA level appears to be regulated in response to cell contact and soluble signals in HEK 293T, however changes of NPAS3 protein and direct contribution of NPAS3 to *TXNIP* regulation could not be confirmed. ARNT was found to be primarily cytoplasmic in all conditions assessed. In SK-N-SH cells, inconsistent detection of NPAS3 protein expression and regulatory function were obtained, potentially suggestive of either technical limitations in detection of NPAS3 protein, or confounding factors regulating *NPAS3* expression, stability and action.

Seven shRNAs were assessed and one, which causes siRNA-mediated silencing was found to have a consistent effect on *NPAS3* mRNA, causing down-regulation of *NPAS3* mRNA, however effects on NPAS3 protein could not be assessed, and effects on known NPAS3 target genes were inconsistent. Complementary expression studies using inducible *NPAS3* constructs confirmed the confounding effects observed by shRNA knockdown studies. ARNT protein expression was found to be relatively weaker in SK-N-SH than HEK 293T cells, raising the possibility of another bHLH-PAS protein being the primary interacting partner for NPAS3 in these cells. *ARNT2* was found to be expressed at the mRNA level in SK-N-SH cells and appears to be predominant at the mRNA level in ENStem-A human neuroprogenitor cells and is worth exploring as a

potential binding partner for NPAS3. Finally, human neuroprogenitor cells were assessed and found to be suitable for future experiments on NPAS3 function.

# Chapter 6: Discussion

## 6.1 NPAS3 behaves as a bHLH-PAS transcription factor

Through the studies described I have functionally characterized the domains of NPAS3 and how they contribute to regulation of NPAS3 target genes. NPAS3 has been characterized as a transcription factor of the bHLH-PAS family since its discovery and analysis of its protein sequence revealed that it contains a bHLH motif and two tandem PAS motifs (Brunskill et al. 1999). Previously, Npas3 and Arnt have been found to synergystically regulate expression of Npas3-regulated genes related to lung morphogenesis (Zhou et al. 2009). However, the study of Npas3 and Arnt driven regulation was performed using luciferase driven by promoter constructs, and neither direct interaction between Npas3 and Arnt, nor direct binding to DNA were demonstrated. Further, the Npas3 construct which was used was not a complete construct, missing 51 residues N-terminal to the bHLH DNA binding domain relative to the full length isoform, and may not be fully functional. Although ARNT has been presumed to be the obligate heterodimeric binding partner for NPAS3 (Brunskill et al. 1999; Pieper et al. 2005), the interaction had not been experimentally confirmed until recent studies of murine Npas3 and Arnt (Moen et al. 2017; Wu et al. 2016). Experimental determination of the functions of the predicted domains of NPAS3 had not been undertaken prior to my experiments. I have demonstrated that ARNT and NPAS3 physically interact (Figure 4-2). Further, I have demonstrated that the bHLH domain is critical for DNA binding specificity and interaction with ARNT; that the PAS domain is critical for

heterodimerization; and that the C-terminus of NPAS3 encodes a potent transactivation domain.

I have assessed two isoforms each of NPAS3 and ARNT. For NPAS3 I assessed interactivity of the longest isoform (933 aa, isoform 1), as well as the splice variant isoform 2 (901 aa, annotated as isoform 1 at the outset of this project) (Figure 1-1). For ARNT, I assessed interactivity of the longest isoform (isoform 1), as well as the splice variant isoform 6. All isoforms of NPAS3 and ARNT were found to be able to interact, supporting the hypothesis that the transactivation domain variation between ARNT isoforms 1 and 6 would not affect interactivity. Furthermore, the 30 aa and 2 aa deletions flanking the bHLH domain that vary between NPAS3 isoforms 1 and 2 did not affect interactivity, indicating that both these isoforms encode proteins that are fully capable of interacting through their bHLH and PAS domains to regulate gene expression. NPAS3 isoforms 1 and 2 were not found to differ significantly in their ability to transactivate expression of the target promoter *VGF,* suggesting that the residues that differ between these isoforms is not critical for NPAS3 function in transactivation at this locus (Figure 4-8). Finally, expression of NPAS3 isoform 1 was sufficient to pull down endogenous ARNT, supporting the interaction as being high affinity and functional in HEK 293T cells (Figure 4-2C). Taken together, these data demonstrate that NPAS3 and ARNT interact robustly and that variant isoforms of these genes can also effectively interact indistinguishably from the longest isoforms.

The molecular nature of the interaction between NPAS3 and ARNT was further characterized by analysis of the function of NPAS3 domains indicated in Figure 1-1. I have found that the bHLH and PAS domains individually are not sufficient for robust

interaction, but both together are required for ARNT and NPAS3 to stably interact (Figure 4-12). This finding is consistent with the literature characterizing other bHLH-PAS heterodimers. In direct support of our finding that both bHLH and PAS domains are required for the NPAS3::ARNT heterodimer (Figure 4-12), the crystal structure of the bHLH-PAS domains of Npas3 and Arnt demonstrated that all three of the bHLH, PAS(A) and PAS(B) domains are involved in the interaction, with similar architecture to other bHLH-PAS protein crystals, such as Npas1::Arnt, Hif1α::Arnt and Clock::Bmal1 (Huang et al. 2012; Wu et al. 2015; Wu et al. 2016). Mutational analysis using variants of murine Arnt known to affect interaction interfaces demonstrated that the interface between the bHLH domains, as well as PAS(A) domains, of Arnt and Npas3 appear to be the most crtitical for interaction, relative to the PAS(B) domain, where variants have relatively less effect on the interaction, similar to other bHLH-PAS proteins (Huang et al. 2012; Wu et al. 2015; Wu et al. 2016). Of note, the murine Npas3::Arnt complex was only able to be stably crystallized in the presence of a double stranded oligonucleotide encoding the hypoxia response element (HRE), suggesting that DNA binding may also be involved in stabilizing the Npas3::Arnt complex (Wu et al. 2016). My studies were performed in whole cell lysates, where genomic DNA was present, so I cannot comment on the necessity of DNA for the interaction between NPAS3 and ARNT.

Structural analysis of bHLH domain proteins, such as MyoD, Max and E47, found that the bHLH domain is involved in the interaction of bHLH domain containing proteins of multiple families (Ellenberger et al. 1994; Kikuchi et al. 2003; Ma et al. 1994). The basic region of the bHLH motif fits in the major groove of the target DNA sequence and interacts with the target nucleotides through the basic (aromatic) residues to

confer sequence specificity, while the helix-loop-helix fold is involved in the interaction between the HLH-containing proteins (Ellenberger et al. 1994; Kikuchi et al. 2003; Ma et al. 1994). Our data demonstrate that the bHLH domain is required (but not sufficient) for high affinity interaction with ARNT (Figure 4-12), and is also involved in conferring sequence specificity to the activity of the NPAS3 transactivation domain, as its absence results in the domains containing the C-terminal transactivation domain being able to activate a promoterless reporter, a non-specific activation (Figure 4-15 B).

PAS domains contain degenerate repeats of a 70 aa PAS motif, PAS(A) and PAS(B) which are involved in interaction between proteins containing these domains (McIntosh, Hogenesch, Bradfield 2010). Both the PAS(A) and PAS(B) domains have been shown to be involved in the interaction between bHLH-PAS proteins (Fukunaga et al. 1995; Pongratz et al. 1998; Reisz-Porszasz et al. 1994). Deletion constructs of the bHLH-PAS protein AhR have demonstrated that the PAS domains without the bHLH domain are able to interact with its known heterodimeric partner, ARNT (Lindebro, Poellinger, Whitelaw 1995). Other studies have observed that although both bHLH and PAS domains of AhR can interact with ARNT independently, both the bHLH and the PAS-A domains are critical for high-affinity interaction and DNA binding (Chapman-Smith, Lutwyche, Whitelaw 2004; Pongratz et al. 1998). The PAS-B is involved in conferring specificity to the regulatory activity of bHLH-PAS heterodimers (McGuire et al. 2001). As such, the presence of bHLH and PAS domains are required to confer specificity to the observed heterodimeric interactions.

Although I was unable to detect interactions between the individual bHLH and PAS(A)-PAS(B) domains by pull down assays, evidence from luciferase reporter assays

suggest that interaction may be occurring at low affinity. When expressed alongside ARNT, the bHLH domain of NPAS3 appears to inhibit its ability to activate expression of the *VGF* promoter, resulting in relative repression (Figure 4-15A). When the PAS domain is expressed in isolation, it has highly variable effects on ARNT transactivation function, suggesting potentially weak effects, or potentially toxic effects given the relatively low transfection efficiency of this construct (Figure 4-10, Figure 4-15). Variable transfection efficiency is controlled for in luciferase assays by the normalization of reporter driven luminescence to the constitutively expressed renilla luciferase, however toxic effects cannot be completely excluded.

I have experimentally demonstrated that the C-terminal domain of NPAS3 is a true transactivation domain, as only NPAS3 constructs containing the transactivation domain are able to activate expression of the known direct target *VGF* promoter (Figure 4-15 A). In a complementary finding, when the bHLH-PAS construct is expressed, which can robustly interact with ARNT, I observed a reduction in the activation of the target *VGF* promoter-luciferase reporter relative to *ARNT* expressed in isolation, as well as relative to *ARNT* co-expressed with *NPAS3*, supporting the finding that the bHLH-PAS domain is able to stably interact with ARNT. This results in inhibition of the inherent transactivation function of the ARNT homodimer. As this domain does localize to the nucleus with ARNT despite lacking a NLS, this downregulation is likely due to the absence of the NPAS3 transactivation domain (Figure 4-15). This conclusion is supported by previous studies, which have found that although ARNT contains a potent transactivation domain, with residues well characterized as critical for homodimeric transactivation activity, the regulatory effects are modulated by the

transactivation/repression domain of its heterodimeric partner (Jain et al. 1994; Jang et al. 2005; Li, Dong, Whitlock 1994; Sogawa et al. 1995; Teh et al. 2006; Whitelaw, Gustafsson, Poellinger 1994; Woods and Whitelaw 2002). In this case I have found that the transactivation domain of NPAS3 is critical for full transactivation function of the NPAS3::ARNT heterodimer. As a final piece of evidence supporting the transactivation domain function, I found that it acts as a weak transactivator of non-targeted gene expression when expressed in isolation (Figure 4-15). These promoterless luciferase data also support the function of the bHLH domain in conferring sequence specificity, as the PAS-TAD construct as well as the TAD in isolation can also non-specifically activate expression. Taken together, these data confirm the predicted functional domains of NPAS3 behave as is prototypical of a class I bHLH-PAS transcriptional activator that acts as a heterodimeric partner of ARNT to regulate target gene expression.

bHLH and PAS domains are generally found to interact, whereby the PAS domain regulates the inherent ability of the bHLH domains to homodimerize and interact with target DNA, and the PAS domain is not required for DNA binding (Pongratz et al. 1998; Uchida et al. 2012). However, in some bHLH-PAS heterodimers, contributions of the PAS domain to DNA binding have been observed, where interaction of the PAS domains affects affinity of DNA binding, either by conformational changes, or direct contact with DNA (Chapman-Smith, Lutwyche, Whitelaw 2004; Rutter et al. 2001; Seok et al. 2017; Wu et al. 2015). Our data do not directly assess the contribution of the PAS domain to DNA binding, although our luciferase data indicate that the bHLH domain appears to be critical for the sequence specificity of NPAS3 regulatory action, so any effects of the PAS domain(s) on DNA binding would be secondary to the bHLH domain.

The crystal structure of the Npas3 and Arnt bHLH-PAS domains did not detect interaction between the PAS domain and DNA, suggesting that the PAS domains of the NPAS3::ARNT heterodimer do not contact DNA (Wu et al. 2016).

The crystal structure of the Npas3::Arnt heterodimer identified mutliple pockets in the PAS domains that may allow for binding of ligands, both endogenous and exogenous, to affect interaction and gene regulatory function (Wu et al. 2016). As PAS domains of bHLH-PAS proteins have been shown to functionally bind cofactors to affect interaction and gene regulatory function, these ligand pockets provide an interesting avenue for further research on the nature of the NPAS3::ARNT interaction, as well as potential targets for pharmacological agents to treat disorders contributed to by *NPAS3* variants (Fukunaga et al. 1995; Rutter et al. 2001; Whitelaw, Gustafsson, Poellinger 1994). Stabilization or activation of the NPAS3::ARNT heterodimer can be predicted to be neuroprotective, given previous data showing that NPAS3 expression is required for proper neurogenesis by decreasing rates of apoptosis of newly born neurons (Pieper et al. 2005; Pieper et al. 2010). Furthermore, the direct targets of the NPAS3::ARNT heterodimeric complex that I have identified can be seen to have relevant functions in neurodevelopment, neurogenesis and neuroprotection, and will be discussed further below.

## 6.2 NPAS3 localization is affected by *ARNT* expression

As a transcription factor containing a predicted nuclear localization sequence in its transactivation domain, NPAS3 is predicted to have predominantly nuclear localization (Figure 1-1) (Macintyre et al. 2010). Previous immunohistochemistry studies

of mouse tissues have identified NPAS3 localization as being predominantly nuclear, or a mix of nuclear and cytoplasmic localization (Pieper et al. 2010; Sha et al. 2012; Stanco et al. 2014). Ectopically expressed NPAS3 was found to be predominantly nuclear in culture (Kamnasaran et al. 2003). These data may indicate that NPAS3 localization to the nucleus can vary based on intra- or extracellular signals, or may be due to variation in detection, due to technical issues, such as secondary signals detected by antibodies. Western blot assessment of endogenous NPAS3 in nuclear and cytoplasmic lysates of SK-N-SH and HEK 293T cells using our antibody detected almost exclusive signal in the nuclear fraction (for example, Figure 5-12), demonstrating that NPAS3 localization is enriched in the nucleus. Although the technical validation of the specificity of the antibody is not complete, pending assessment of signal relative to an *NPAS3*-deleted cell line, our antibody does have high affinity to expressed NPAS3 constructs, can immunoprecipitate them, and the predominant bands detected are of the expected range (~100-130 kDa), relative to clones of NPAS3 (size close to 130 kDa) and predicted molecular weight (101 kDa). Immunoprecipitation demonstrates the ability of the antibody to interact with native NPAS3 (Figure 4-1D). Immunofluorescence of endogenous NPAS3 in SK-N-SH cells using our antibody demonstrated predominantly nuclear signal, with cytoplasmic signal of variable intensity, which is consistent with the previously observed localization (Figure 4-1E). Immunofluorescence using a commercial antibody demonstrated both nuclear and cytoplasmic staining (Figure 5-5D). These data recapitulate the findings in the literature, whereby NPAS3 is variably detected: as nuclear, or both nuclear and cytoplasmic in localization.

In order to assess the regulation of subcellular localization of NPAS3, expression of *NPAS3* constructs was used. In these studies NPAS3 localization was found to be predominantly nuclear, and this nuclear localization was enhanced with co-transfection with ARNT (Figure 4-13,Figure 4-14). The effect of ARNT on NPAS3 localization is similar to that observed with NPAS1, however NPAS1 is excluded from the nucleus in the absence of ARNT, and is not exclusively nuclear in the presence of ARNT, as I observe with NPAS3 (Teh et al. 2006). ARNT has largely been considered to be stably expressed and localized to the nucleus, whereas its heterodimeric partners are differentially expressed and localized (Chilov et al. 1999; Hord and Perdew 1994; Pollenz 1996). Our data show that expressed ARNT was almost invariably localized to the nucleus (Figure 4-10, Figure 4-13). A study of heterodimers of ARNT::HIF1α demonstrated that heterodimeric complexes appear to be more stably associated with the nuclear compartment than when expressed in isolation (Chilov et al. 1999). These data suggest that the observed enhancement of NPAS3 localization to the nucleus by co-expression of ARNT may be due to their interaction stabilizing NPAS3 localization to the nucleus.

Incidentally, when assessing the expressivity of NPAS3 domain constructs, I identified differential localization of the NPAS3 domains in the presence of ARNT. While any NPAS3 domain construct containing the transactivation domain, which contains the only predicted nuclear localization signal of NPAS3, but not the bHLH domain, which contains a predicted nuclear export sequence, was found to be localized to the nucleus irrespective of ARNT (Figure 4-14). These data suggest that the predicted nuclear localization signal of NPAS3 is functional. This should be confirmed by

mutagenesis of the sequence and assessment of subcellular localization. The PAS domain of NPAS3 was not observed to be specifically localized to the nucleus until *ARNT* was expressed. This observation is of interest, given that I do not observe robust interaction between the PAS domains and ARNT in isolation, however I see an effect at the level of localization. As the PAS domain does not contain a predicted nuclear localization signal, its localization being ubiquitous in the nucleus and cytoplasm is not unexpected, however the effect of *ARNT* expression on its targeting of the nucleus is unexpected. A similar effect has been observed whereby ARNT stabilizes nuclear localization of a NLS-deleted AHRR (aryl hydrocarbon receptor repressor, another bHLH-PAS ARNT interacting partner), although the mechanism is unknown (Kanno et al. 2007).

The bHLH domain of NPAS3 also contains a predicted nuclear export signal and appears to have an opposite response to *ARNT* expression relative to other domains, which may suggest that this domain is active. The bHLH domain appears to be involved in the localization of full length NPAS3, given the differences in its responsiveness to ARNT expression relative to the PAS-TAD construct, suggesting export function (Figure 4-14). Nuclear export sequences of AhR have been shown to function in regulating its localization in response to cell::cell contact signals (Ikuta, Kobayashi, Kawajiri 2004). The localization of NPAS3 may also be regulated in part by a similar NES-dependent mechanism. The bHLH-PAS domain was found to behave similarly to full-length NPAS3 in response to ARNT, suggesting that NPAS3 localization to the nucleus can be facilitated both by the transactivation domain independently of ARNT, as well as in a NLS-independent, but ARNT-dependent mechanism, such as by the stabilization of bHLH-PAS heterodimers in the nucleus relative to monomers (Chilov et al. 1999).

## 6.3 NPAS3 and ARNT cooperatively bind to proximal promoter regions of *TXNIP*, *VGF* and *ANKRD37*

A screen of all genes identified as regulated by NPAS3 expression, co-registered with genes with robust ChIP-seq peaks from experiments on endogenous ARNT, was performed. I have identified *VGF* and *TXNIP* as direct regulatory targets of the NPAS3::ARNT heterodimer. I confirmed that NPAS3 and ARNT can affect the regulation of these genes through both regulation of their expression from the genomic loci (Figure 4-3), as well as through luciferase analysis of the proximal promoter regions to the ChIP positive primers (Figure 4-5, Figure 4-8). Through this screen *ANKRD37* was also found to be a direct target of *NPAS3* and *ARNT* (Figure 5-1), however differential regulation by co-expression of NPAS3 and ARNT was not detected (Figure 5-2).

Recently, a study has been published in which RNAseq and ChIP-seq was performed to assess the gene regulatory function of *Npas3* in wild-type and *Npas3* knockout mouse hippocampi (Michaelson et al. 2017). Deletion of *Npas3* was found to differentially regulate 1141 genes, primarily through activation, consistent with our characterization of NPAS3 as having a true transactivation domain. Synthesis of RNAseq and ChIP-seq data identified 461 genes that were directly regulated by NPAS3. *Npas3* knockout mice were used to confirm that the detected peaks were associated with Npas3 binding specifically, and not a secondary target of the antibody. *Txnip* and *Vgf* were not found to be differentially regulated, nor directly bound, by Npas3 in mouse hippocampi. *Ankrd37* was also not identified as bound, or regulated, by Npas3. These discordant results may be due to several factors, in part owing to the vastly different

methodological factors. For example, NPAS3 in humans may regulate overlapping but distinct target genes relative to mice. Alternately, as *Npas3* is functionally expressed in non-neuronal cells, it may regulate different genes in different tissues (Zhou et al. 2009). Furthermore, NPAS3 may target different genes in different cells within the same tissue, such as the observation for different target genes regulated by Npas4, another neural bHLH-PAS transcription factor not otherwise functionally related to Npas3, in different neuronal subtypes (Spiegel et al. 2014). Finally, the regulation of *TXNIP* and *VGF* by NPAS3 may be regulated in response to environmental factors, which, as discussed above, is a common feature of bHLH-PAS proteins and relevant to the functions of *TXNIP* and *VGF,* which are discussed further below. Despite this non-replication, *VGF* and *TXNIP* appear to be true targets of NPAS3 and ARNT regulation.

## 6.4 Transcriptional regulation of *TXNIP*

*TXNIP* expression is rapidly and dynamically regulated in response to multiple environmental inputs, notably oxidative stress, glucose/glycolytic balance and hypoxia (Baker et al. 2008; Chai et al. 2011; Cha-Molstad et al. 2009; Chen et al. 2008; Gao et al. 2015; Kanari et al. 2013; Kim et al. 2012; Minn, Hafele, Shalev 2005). I have cloned the -867 bp proximal to the TSS of *TXNIP* which includes sequences critical for this regulation. The 250 bp directly upstream of the *TXNIP* transcription start site contains two carbohydrate response elements (ChoRE), each of which contains two tandem E-boxes spaced five nucleotides apart, as well as two inverted CCAAT motifs and a FOXO1 binding site, which have been shown to be critical for rapid regulation of *TXNIP* in response to intracellular stimuli (Minn, Hafele, Shalev 2005; Yu and Luo 2009).

*TXNIP* is regulated in response to glucose via the concentration of precursors for glycolysis: phosphorylated hexose sugars, such as glucose-6-phosphate (G6P), which signal glycolytic rate (Yu et al. 2010). High levels of glycolytic precursors, such as G6P, promote nuclear localization of the MondoA bHLH-ZIP transcription factor, which heterodimerizes with its partner bHLH-ZIP protein Mlx to bind the ChoRE E-boxes within the 250 bp of the proximal *TXNIP* promoter and activate transcription (Peterson et al. 2010; Stoltzman et al. 2008; Stoltzman et al. 2011; Yu and Luo 2009). NF-Y complex binding to the tandem inverted CCAAT motifs between the repeated ChOREs has been shown to facilitate assembly of an activation complex, which may involve chromatin looping, whereby the two MondoA::MLX heterodimers interact to activate transcription (Yu and Luo 2009). Similarly, ChREBP (ChoRE binding protein, another bHLH-ZIP family transcription factor) has been shown to bind the *TXNIP* promoter via the ChoRE E-box sequences in response to high glucose, resulting in up-regulation of transcription (Cha-Molstad et al. 2009; Chau et al. 2017).

*TXNIP* expression is regulated in hypoxia through both HIF1α-dependent and HIF1α-independent mechanisms. There is an initial transient down-regulation of *TXNIP* in hypoxia in a HIF1α-independent mechanism, followed by up-regulation of *TXNIP* by long-term hypoxia, which is HIF1α-dependent (Baker et al. 2008; Chai et al. 2011). The transient down-regulation may be due to sequestration of MondoA by mTOR, which has been shown to occur in response to oxidative stress resulting in MondoA being unable to activate *TXNIP* expression (Kaadige et al. 2015). Activation of *TXNIP* in response to glucose stress has been shown to be disrupted through inhibition of MondoA::Mlx and ChREBP::Mlx function, supporting this hypothesis (Chau et al. 2017; Kanari et al. 2013).

These data demonstrate that rapid regulation of *TXNIP* expression is commonly mediated by regulation at the ChoREs in the *TXNIP* promoter. FOXO1, a transcription factor involved in the regulation of glucose metabolism, has also been shown to be involved in glucose-mediated regulation of *TXNIP* by competing for ChORE binding resulting in inhibition, further supporting this assertion (de Candia et al. 2008; Kibbe et al. 2013). NPAS3 may repress *TXNIP* expression by inhibiting the transcriptional regulation mediated by these motifs.

As described above, ChoREs such as those present in the *TXNIP* promoter, contain E-box sequences known to be bound by bHLH domain containing proteins, including bHLH-ZIP and bHLH-PAS proteins. As such, analysis of the 1000 bp upstream of the *TXNIP* transcription start site by ConTra v2.0 for target sequences of ARNT complexes (which bind E-box and E-box like sequences), identified one of the E-box sequences of the ChoREs as a potential target of ARNT-mediated regulation of *TXNIP,* as well as an E-box like sequence 20-25 nt upstream of the predicted transcription start site, and two less conserved sites further upstream (Figure 4-5). I assessed transcription of a reporter gene driven by the *TXNIP* promoter and found that NPAS3 and ARNT are able to affect transcription driven by this construct. The resulting activation of reporter expression is contrary to the repressive effect of *NPAS3* and *ARNT* expression on the expression of the *TXNIP* gene, as well as to the same repressive effect observed with *NPAS3* expression by another group (Sha et al. 2012). Although the contribution of the variants present in our construct to the responsiveness of our clone of the *TXNIP* promoter to ARNT and NPAS3 cannot be completely excluded, the variants are not localized to predicted ARNT binding sites nor glucose response elements, nor are they in

sequences found to be highly conserved, suggesting that they should not affect regulatory function significantly. Furthermore, the construct is responsive to glucose (Figure 4-5 EV RLUs), suggesting that the variants do not compromise transcriptional responsiveness.

An alternate explanation for this result is that the full genomic region/context required for repression is not encoded in our construct. Chromatin looping mediated by NF-Y complex binding to the CCAAT sequences to form an activation complex between the ChOREs of the *TXNIP* promoter may contribute to regulation of expression in response to glucose (Yu and Luo 2009). A similar complex formed between ARNT heterodimer binding sites flanking the *TXNIP* transcription start site may be involved in full regulatory response to ARNT::NPAS3 (Figure 4-5A). Interactions between ARNT heterodimers binding at the promoter and exon 1 may inhibit transcription through steric hinderance or exclusion of factors required for transcriptional activation/maintenance. Alternately, other factors binding to more distal sites may be involved in the full repressive complex.

The exonic ARNT ChIP-seq region was not included in our construct for several reasons: firstly, NPAS3 was found to consistently interact directly with the proximal promoter region of *TXNIP* relative to the inconsistent pull-down of the exon 1 region (Figure 4-4), secondly, ConTra v2.0 analysis did not identify any potential ARNT binding sites in the 5'UTR of *TXNIP*, which is also a region of relatively low conservation. Transcription factor binding sites were not able to be predicted in the coding sequence of exon 1 due to conservation. Intron 1 was predicted to have several potential ARNT binding sites, however their relative conservation was low. Designing a

*TXNIP*-luciferase reporter with an intron was not undertaken, due to complexity in design and issues in obtaining a non-mutated construct of the initial region assessed.

Taken together, these data suggest that NPAS3 and ARNT act through both the proximal promoter of *TXNIP,* as well as other more distal loci to affect regulation of *TXNIP*. These data help reconcile the observation of repression of *TXNIP* by NPAS3, despite its potent transactivation domain, however, further characterization of the mechanism of repression is required. Deletion mapping of the *TXNIP* promoter construct tested to identify the specific region by which NPAS3 and ARNT regulate *TXNIP* expression should be pursued to determine which region of the promoter is bound by NPAS3 and ARNT, which may allow for refinement of the promoter region required to design the larger construct. Once the minimal interval is determined, constructs can be designed to include the exonic ARNT ChIP-seq peak to test the hypothesis that both sites are required for repression. Mutagenesis of these sites can then be used to confirm the specific sequences bound by NPAS3 and ARNT to affect the repression of *TXNIP* expression.

## 6.5 Transcriptional regulation of *VGF*

*VGF* (non-acronymic) was initially identified as a rapidly induced gene in response to treatment with nerve growth factor (NGF) (Ferri, Levi, Possenti 1992; Salton, Fischberg, Dong 1991; Salton 1991). Expression of *VGF* has since been shown to be induced in neurons by other neutrophins, such as neutrophin 3 (NT-3) and brain-derived neurotrophic factor (BDNF), but not induced as robustly by other growth factors, such as FGF (Alder et al. 2003; Hawley, Scheibe, Wagner 1992; Possenti et al. 1992; Salton,

Fischberg, Dong 1991). Activation of *VGF* by NGF signalling through receptor tyrosine kinase Trk/Ras pathways was found to be dependent on both the CCAAT element and the cAMP response element (CRE) in the 250 bp upstream of the transcription start site, as well as binding by the histone acetyl transferase p300 and EGR1 (nerve growth factor-induced protein A, NGFI-A), but not the E-box present in this interval (D'Arcangelo and Halegoua 1993; D'Arcangelo et al. 1996). Rather, the E-box appears to play a role in restriction of *VGF* expression into neuronal cell types through binding of bHLH proteins, such as MASH1 and TCF12 (Di Rocco et al. 1997; Mandolesi et al. 2002). *VGF* expression in the suprachiasmatic nucleus is regulated by light and the circadian clock, possibly through the E-box element (Wisor and Takahashi 1997).

These sequences are included in our *VGF* promoter-luciferase reporter, from which I observed up-regulation of *VGF* promoter driven reporter expression by NPAS3 and ARNT independently and cooperative regulation when they are co-expressed. These data are somewhat different from those observed with regulation of the endogenous *VGF* gene, whereby only co-transfection resulted in differential expression. This may be due to other factors contributing to the regulation of the endogenous *VGF* promoter region encoded in our promoter construct. Alternately, the luciferase reporter may have a higher signal:noise ratio, resulting in more sensitive detection of transcriptional regulation.

NPAS3 and ARNT may affect regulation of *VGF* in HEK 293T cells by activating a gene that should be silent in non-neuronal cells, by signalling neuronal cell type at the E-box upstream of the CCAAT and CRE. It is worth noting that the true cell type of HEK 293T is unknown, as the embryonic kidney from which it was derived is a mixture of cell types, and it has been proposed that HEK 293T cells are neural crest in

origin (Lin et al. 2014). NPAS3 and ARNT may up-regulate *VGF* expression by a mechanism overlapping with the observed circadian regulation of *VGF* in the suprachiasmatic nucleus, however, *VGF* was not observed to be differentially regulated by NPAS3 over the circadian cycle and NPAS3 was not found to be expressed in a circadian manner (Figure 5-8) (Sha et al. 2012).

I performed preliminary analysis as to whether NPAS3 is able to affect gene regulation in a circadian manner by assessing whether it can interact with BMAL1 to effect differential regulation over the circadian cycle and I did not observe a strong, or consistent, interaction (Figure 5-8). Recently, the regulation of *VGF* expression by NPAS3 has been characterized by another group with similar effects observed, although they did not assess the contribution of ARNT to the regulation (Yang et al. 2016). Given their western blot data, where they assess expression and knockdown of NPAS3 by intensity of a 29 kDa band, their expression data are difficult to interpret. However, they observe a similar effect size with expression of NPAS3 (approximately two-fold activation), further, shRNA knockdown resulted in relatively reduced expression. Their construct is over 2 kb in size and contains the entire 750 bp promoter region assessed in this study, as such our findings are generally consistent. In sum, NPAS3 and ARNT are able to cooperatively bind and regulate expression of the *VGF* promoter through the 750 bp construct, including the proximal *VGF* promoter. Identification of the direct target site should be undertaken using deletion constructs, as well as site-directed mutagenesis, notably as a putative Npas3 binding site has been recently predicted (Michaelson et al. 2017).

## 6.6 Functional assessment of *NPAS3* variants

Multiple coding variants to *NPAS3* have been identified, with varying linkage to neuropsychiatric dysfunction. Three such variants were assessed in this project for relative function in molecular assays developed for the characterization of NPAS3 function as a bHLH-PAS transcription factor. The variant c.910G>A (p.Val304Ile) was identified in a small family as co-segregating with mental illness and resulted in altered neurite outgrowth when expressed in mouse primary cortical neurons (Yu et al. 2014). This variant was found to affect regulation of a known NPAS3 target promoter, *VGF*, and appeared to promote the aggregation of both variant and wild-type endogenously expressed NPAS3 into the insoluble fraction of COS-1 (green monkey kidney) and 293FT (derived from HEK 293T) cells (Nucifora et al. 2016).

My initial hypothesis was that this variant could affect the ability for NPAS3 to interact with ARNT, however, I found no significant effect of this variant on ARNT interaction (Figure 4-9). Furthermore, I found NPAS3 p.Val304Ile was normally localized to the nucleus (Figure 4-10), and able to transactivate expression of a *VGF* promoter driven luciferase reporter, both in the presence and absence of ARNT (Figure 4-11B). These data suggest that this variant is not functionally severe. The assay system used was not amenable to assessment of NPAS3 solubility, so I could not perform confirmatory assays. However, luciferase reporter assays demonstrate no reduction in NPAS3 function, which would be predicted by protein sequesteration into aggregates, and was observed in the previous study (Nucifora et al. 2016). My data do not support the pathogenicity of this variant in NPAS3 function, however further characterization is required.

The schizophrenia-associated variant c.1654G>C (p.Ala552Pro) was assessed for expression, protein localization, interactivity and regulatory function. Bioinformatic analysis performed previously predicted that the variant is benign due to relative non-conservation at the site, notably as proline is the ancestral residue in orthologous sequences (Macintyre et al. 2010). This variant was functionally assessed as I demonstrated that c.1654G>C is associated with altered working memory function (Table 8). The nature of the protein variant is also notable, an alanine to proline change represents a large amino acid substitution that restricts backbone flexibility and is often poorly tolerated (Gray, Hause, Fowler 2017).

NPAS3 p.Ala552Pro was found to be expressed and localized normally relative to wild-type NPAS3. Furthermore, interactivity with ARNT was not affected (Figure 4-9Figure 4-10). As this variant is localized to the transactivation domain of NPAS3, its ability to regulate expression of a reporter driven by the target *VGF* promoter was assessed and found to be functionally equivalent to wild-type NPAS3, both with and without co-expression of ARNT (Figure 4-11A). These data indicate that this variant is benign. A functional consequence of this variant may exist, however I was unable to detect any effect in my assays. Given that this variant has now been found to be present in the normal population at similar rates to those originally identified in the schizophrenia patient cohort, the finding that this variant does not severely affect NPAS3 function is unsurprising (Lek et al. 2016; National Center for Biotechnology Information, National Library of Medicine 2017). This variant has also been predicted to affect splicing of *NPAS3*, an effect that would not be observed in the performed assays, and should be assessed otherwise (Macintyre et al. 2010). As such, further assessment of the

contribution of this variant to *NPAS3* expression and protein function is warranted, both in isolation and in combination with the rest of the haplotype with which it co-segregates.

The final variant tested was c.2089G>A (p.Gly697Ser), which is within the first poly-glycine repeat in the transactivation domain (Macintyre et al. 2010). This variant has been identified in normal populations as a rare variant, and has not been associated with neuropsychiatric dysfunction (Lek et al. 2016; National Center for Biotechnology Information, National Library of Medicine 2017). Poly-amino acid repeats are found to be expanded in humans relative to other animals (Alba and Guigo 2004). Although the function of poly-glycine repeats is largely unknown, poly-glycine repeat length variation as small as one repeat has been associated with transactivation domain function, as well as with altered disease risk (Brito et al. 2005; Brockschmidt, Nothen, Hillmer 2007). Variants affecting poly-glycine repeat length in *MECP2* have been identified as associated with intellectual disability (Harvey et al. 2007). Expanding glycine repeats in transcription factors have been noted to be involved in increasing distance between functional domains and to enhance interactions with other proteins (Cummings and Zoghbi 2000; Mojsin et al. 2010). Although large poly-glycine repeats have been shown to aggregate and contribute to the pathogenicity of FXTAS (Fragile X Tremor Ataxia Syndrome), a neurodegenerative disorder, smaller poly-glycine repeats (30 aa repeats) have not been found to aggregate significantly (Alba and Guigo 2004; Sellier et al. 2017; Todd et al. 2013). However, this may suggest that poly glycine repeats are involved in protein interaction.

Assessment of expression and localization of the NPAS3 isoform 1 p.Gly697Ser variant found it to be normally expressed and localized to the nucleus (Figure 4-9Figure

4-10). Transactivation assays demonstrate that, when expressed on its own, this variant is normally able to activate target gene expression, indicating that transactivation function is preserved (Figure 4-11A). When co-transfected with ARNT, however, decreased transactivation function was observed. These data are suggestive of a deficit in transactivation function, however the lack of deficit in singly-transfected conditions is contradictory to this finding. As the functional conservation of this domain is relatively less than other regions, and the polyglycine repeat is expanded in humans (Pickard et al. 2006), interpretation of this finding is limited by the understanding of the functional role of the poly-glycine repeat in transactivation function. A potential explanation is that NPAS3 may be interacting with the target promoter when singly transfected with a different heterodimeric partner than ARNT, whereby this variant has a different effect on the transactivation by NPAS3 with a different heterodimeric partner. Subcellular fractionation of HEK 293T cells does detect nuclear ARNT (Figure 5-3C), indicating that ARNT is available endogenously for interaction with singly-expressed NPAS3. These data do not exclude a separate heterodimeric binding partner also being present and functionally available. Alternately, if NPAS3 is interacting with ARNT in both cases, the increased transactivation when both NPAS3 and ARNT are expressed can be hypothesized to be due to increased availability of ARNT to heterodimerize with the expressed NPAS3, allowing for maximal activation of the *VGF* promoter-luciferase reporter, potentially rendering the assay more sensitive to functional deficits of NPAS3. Further assessment of the p.Gly697Ser variant should be performed in other assays to characterize the significance of this observed deficit in transactivation. Also,

characterization of the residues critical for NPAS3 transactivation function would expand the understanding of the functional consequences of this variant.

## 6.7 NPAS3 is capable of regulating genes involved in cellular response to hypoxia

NPAS3 expression has been shown to strongly inhibit *ANKRD37* expression under normoxic conditions (Sha et al. 2012). In my initial screen for co-targets of NPAS3 and ARNT, I did not observe differential regulation of *ANKRD37* with expression of NPAS3 and/or ARNT in normoxia (21% oxygen, Figure 5-2). Despite the non-finding of regulatory changes to *ANKRD37* expression by NPAS3 or ARNT, the *ANKRD37* promoter was found to be bound by NPAS3, when co-expressed with ARNT, as well as by ARNT regardless of expression of NPAS3 (Figure 5-1). Although *ANKRD37* was found to be expressed in HEK 293T cells cultured under normoxic conditions, it has been characterized as a part of the core response to hypoxia and directly up-regulated by HIF1A (Benita et al. 2009). Given these data, further studies were undertaken using models of hypoxia to determine if the previously observed repression of *ANKRD37* by NPAS3 is related to hypoxic gene regulation.

HIF1A is a bHLH-PAS protein that, when stabilized, interacts with ARNT to drive hypoxia-specific gene expression (Wang et al. 1995; Wood et al. 1996). *HIF1A* is constantly transcribed and translated, but rapidly degraded, in the presence of oxygen by the ubiquitin-proteasome system to prevent hypoxic response in normoxia and sustain normal oxygen tension (Maxwell et al. 1999; Salceda and Caro 1997). Culture of HEK

293T cells in 3% oxygen was performed to induce cellular hypoxic response. Down-regulation of *TXNIP* was observed, consistent with its known transient repression in hypoxia, confirming that culture in 3% oxygen is sufficient to elicit a hypoxic response of HEK 293T (Figure 5-2) (Chai et al. 2011; Wong and Hagen 2013). *ANKRD37* expression was not found to be induced by culture in 3% oxygen, nor did expression of NPAS3 and/or ARNT effect its expression (Figure 5-2), conflicting with the apparently normal response of *TXNIP* to hypoxia. The regulation of *TXNIP* by hypoxia has been shown to be regulated both by HIF1A-dependent and -independent mechanisms, where the HIF1A independent mechanism results in rapid down-regulation of *TXNIP* that is transient until HIF1A stabilization causes up-regulation of *TXNIP* (Baker et al. 2008; Chai et al. 2011). As such, the down-regulation of *TXNIP* I have observed cannot be used to exclude the possibility that I do not see *ANKRD37* up-regulation because HIF1A is either not fully stabilized, or not functionally active in HEK 293T cells incubated at 3% oxygen. Although HIF1A has been shown to be stabilized and able to activate transcription driven by the hypoxia response element when HEK 293T cells are cultured in 3% oxygen, full stabilization is not observed until <1% oxygen, which may mute HIF1A-dependent hypoxic gene regulatory effects (Bracken et al. 2006).

In order to specifically assess *ANKRD37* regulation by stabilization of HIF1A, HEK 293T cells expressing a combination of HIF1A, ARNT and NPAS3 constructs were treated with $CoCl_2$ to stabilize HIF1A and enhance hypoxic gene regulation. $CoCl_2$ inhibits the interaction between HIF1A and VHL (Von-Hippel Lindau protein), preventing ubiquitination and proteasomal degradation of HIF1A in normoxia, resulting in hypoxic gene regulation (Yuan et al. 2003). Expressed V5-tagged HIF1A was found to

be stabilized by the standard concentration of 100 μM $CoCl_2$ (Figure 5-2), however *ANKRD37* expression was only weakly induced with no additive effect of HIF1A expression. A potential explanation for this muted transcriptional response despite HIF1A stabilization could be due to activity of FIH-1 (Factor Inhibiting HIF-1), which under normoxic conditions catalyzes a secondary hydroxylation of the HIF1A transactivation domain. The hydroxylation prevents p300 co-activator recruitment for transcriptional activation. However, FIH-1 has been shown to be inhibited by $CoCl_2$, restoring interaction with p300 (Lando et al. 2002a; Lando et al. 2002b). Hyperactivity of FIH-1 may explain the observation of stabilization of HIF1A without coordinate transactivation. Increased concentration of $CoCl_2$ was not pursued due to observed toxicity in the cells. The effect of NPAS3 on the observed weak up-regulation of *ANKRD37* in 100 μM $CoCl_2$ was characterized, despite the relatively muted response, given the previously observed strongly repressive effect of NPAS3 on *ANKRD37* expression (Sha et al. 2012). When NPAS3 was expressed in the presence and absence of ARNT or HIF1A, the trend was for activation of *ANKRD37* expression by NPAS3 (Figure 5-2B), the opposite of the observed strong repression of *ANKRD37* by NPAS3 observed previously (Sha et al. 2012).

As a result of these conflicting data, combined with the lack of a transcriptional response in 3% oxygen where HIF1A should be stabilized and transcriptionally active, further experiments were not conducted, despite the ability of NPAS3 and ARNT to bind the promoter region of *ANKRD37*. These data suggest that NPAS3 is capable of directly regulating *ANKRD37* in an ARNT-dependent manner, however, the environmental signals required to elicit a regulatory response is not present under the conditions tested.

*ANKRD37* expression is associated with proper angiogenesis during pregnancy (Marshall et al. 2016; Trifonova et al. 2014). *ANKRD37* encodes a 158 aa protein that contains three ankyrin domains, helix-loop-helix motifs that are involved in protein::protein interactions, and a nuclear localization signal, however its function is unknown (Li, Mahajan, Tsai 2006; Shi et al. 2011). Given these limited data, assessment of the interaction between NPAS3 and *ANKRD37* expression was not pursued further.

## 6.8 TXNIP, inflammation and the brain

TXNIP (thioredoxin interacting protein) is a member of the α-arrestin, arrestin domain-containing protein family, a family of proteins with diverse interaction domains that allow for interaction with multiple classes of proteins to regulate intracellular signalling pathways (Figure 6-1A) (Spindel, World, Berk 2012). TXNIP was characterized as an interacting partner of TRX (thioredoxin, an intracellular antioxidant protein) and is capable of inhibiting the reducing activity and expression of TRX, resulting in increased reactive oxygen species within the cell (Junn et al. 2000; Li et al. 2009; Nishiyama et al. 1999; Patwari et al. 2006). This increase in oxidative stress is accompanied by increased apoptosis and expression of proinflammatory cytokines (Junn et al. 2000; Minn, Hafele, Shalev 2005; Oslowski et al. 2012; Zhou et al. 2010). Maintenance of appropriate expression of *TXNIP* is critical for cellular homeostasis and metabolic balance in the body. Loss of *TXNIP* expression in mice results in increased triglycerides, ketones and fatty acids, as well as the inability to adapt metabolic pathways to altered nutritional state, resulting in selective insulin resistance and hypoglycemia,

suggesting a role for TXNIP in metabolic balance and diabetes (Donnelly et al. 2004; Hui et al. 2004).

*TXNIP* expression is induced by multiple stimuli, including glucose, glucocorticoids, vitamin D and oxidative stress (Junn et al. 2000; Schulze et al. 2004; Shalev et al. 2002; Wang et al. 2006). Up-regulation of *TXNIP* is associated with increased rates of mitochondrially mediated apoptosis, while anti-apoptotic effects have been observed with loss of *TXNIP*, notably in response to high glucose conditions, demonstrating that it is a key node in the regulation of apoptosis in response to cellular stress (Chen et al. 2008; Minn, Hafele, Shalev 2005; Wang et al. 2006). Oxidative stress has been shown to cause translocation of TXNIP into the mitochondria where it causes release of the pro-apoptotic kinase ASK-1 and allows initiation of the mitochondrial apoptotic signalling cascade, resulting in cell death (Bhattacharyya et al. 2003; Saxena, Chen, Shalev 2010). Furthermore, TXNIP function has been associated with a balance between energy metabolism and growth, whereby TXNIP has been found to regulate the catalytic activity of PTEN, where loss of TXNIP results in increased signalling through the PI3K/AKT signalling pathway and reduced cell growth (Chen et al. 2008; Hui et al. 2008; Kaadige et al. 2015). As such, TXNIP may be a key node for the integration of cell growth and apoptotic pathways in response to metabolic and cellular stress signals that, when dysregulated could cause the known carcinogenic, metabolic and inflammatory disorders with which its dysregulation is associated (Chong et al. 2014; Jin and Flavell 2010; Zhou and Chng 2013).

Alongside its role in regulation of apoptosis, TXNIP is involved in regulation of cellular inflammatory responses to intracellular and extracellular stress in multiple cell

types and organ systems. TXNIP has been shown to interact with NLRP3, a component of the NLRP3 inflammasome, resulting in the production and secretion of proinflammatory cytokines, including interleukin 1β (Oslowski et al. 2012; Zhou et al. 2010). Interleukin 1β is involved in the initiation and regulation of local and systemic inflammatory processes and is involved in a wide range of inflammatory disorders, as reviewed in (Dinarello 2011). The secretion of inflammatory cytokines in response to cellular stress, including high glucose and endoplasmic reticulum stress, was shown to be dependent on induction of TXNIP, occurring concurrent with the induction of apoptotic pathways (Oslowski et al. 2012; Zhou et al. 2010). In summary, regulation of TXNIP expression is critical for proper inflammatory and apoptotic responses to cellular stress caused by both intra- and extra-ceullar factors. These responses affect both the cellular apoptotic response, but also the inflammatory/apoptotic tone of neighbouring cells. As such, regulation of TXNIP must be maintained in order to ensure proper cellular responses to metabolic and inflammatory stress.

NPAS3 may form a negative feedback loop on expression of *TXNIP*, reducing the apoptotic and inflammatory response of *NPAS3*-expressing cells. In stress conditions, I found that NPAS3 inhibited the induction of *TXNIP* (Figure 4-3). Furthermore, in conditions where *TXNIP* was induced and has been shown to cause secretion of inflammatory markers, I found that *NPAS3* expression was also induced (Figure 5-4). Although I was not able to characterize whether NPAS3 protein was expressed and able to regulate *TXNIP*, the induction of *NPAS3* expression by soluble factors, potentially including inflammatory cytokines can be predicted to limit the expression of *TXNIP* and thus modulate the cellular inflammatory and apoptotic response. This would result in

increased survival of NPAS3+ cells, which is consistent with the observed enhancement of apoptosis of hippocampal neuroprogenitors in mice deficient in *Npas3* (Pieper et al. 2010).

TXNIP function is conserved in neuronal cells, supporting the relevance of NPAS3 binding to *TXNIP* in neuronal function. *TXNIP* expression is induced in neuronal cell types in response to oxidative agents and altered glucose metabolism, as well as removal of nerve growth factor, coordinate with increased rates of apoptosis (Kristiansen et al. 2011; Price et al. 2006; Zaragoza-Campillo and Moran 2017). Study of TXNIP in the CNS *in vivo* has largely been performed to assess its regulation and effects on brain injury. *TXNIP* expression is up-regulated in response to ischemia/reperfusion injury and excitotoxicity, as well as subarachnoid hemorrhage, concordant with rates of apoptosis (Kim et al. 2012; Li et al. 2015; Zhao et al. 2017). Reduction of TXNIP levels has been shown to be neuroprotective in these models, and also involved in the neuroprotective effects of compounds such as curcumin and resveratrol (Kim et al. 2012; Li et al. 2015; Zhao et al. 2017). Similar effects were observed with TXNIP-dependent release of inflammatory markers by endothelial cells of the blood brain barrier in response to ischemia. Treatment of mice with the neuroprotective agent ruscogenin after ischemic injury resulted in reduction of endothelial dysfunction, inflammation and *TXNIP* expression (Cao et al. 2016). These data demonstrate that *in vivo* rodent models support a neuroprotective role for down-regulation of *TXNIP* upon brain injury and environmental stress. Further, these data show that many agents with known neuroprotective or neurotoxic effects act through TXNIP, supporting its importance in neuroprotection.

Finally, a study of post-mortem human brains identified *TXNIP* expression as up-regulated in the hippocampi of alcoholic individuals, along with other markers of stress response pathways (McClintick et al. 2013). As such, *TXNIP* expression does appear to respond to cellular stress in loci relevant to the known function of NPAS3. Furthermore, these data indicate that the observed down-regulation of *TXNIP* by NPAS3 in relatively stressed HEK 293T cells ('aged' media vs fresh, Figure 4-3) may be relevant to neuronal cells *in vivo* and can be seen to be neuroprotective. Down-regulation of *TXNIP* by NPAS3 during oxidative stress or injury in NPAS3+ neurons/neuroprogenitors may promote the survival of neurons by reducing the responsiveness of the ASK1-mediated mitochondrial apoptosis pathway. NPAS3-mediated reduction in *TXNIP* may also decrease the inflammatory tone of neighboring cells by reducing the amount of pro-inflammatory cytokine IL-1β. Further, inflammatory markers released in the CNS in response to stress may induce the expression of *NPAS3* to modulate the cellular response in a TXNIP-dependent manner.

In order to test these hypotheses, relative responsiveness of neuroprogenitor cells and/or neuronal cells normally expressing *NPAS3* or deficient for *NPAS3* to neuronally relevant stresses (such as excitotoxic agents, high glucose and ischemia) should be assessed. These experiments can be performed in human neuroprogenitor cells, primary neuron cultures and in mice, including the *Npas3* knockout mouse. Relative expression of *TXNIP* should be assessed and confirmation should be performed by knockdown of *TXNIP*. These data will allow us to confirm whether NPAS3 regulates *TXNIP* in response to stress in neuroprogenitors and will further the understanding of the role of NPAS3

regulation of *TXNIP* in neuronal stress response which can be seen to contribute to variation in neuropsychological function associated with NPAS3.

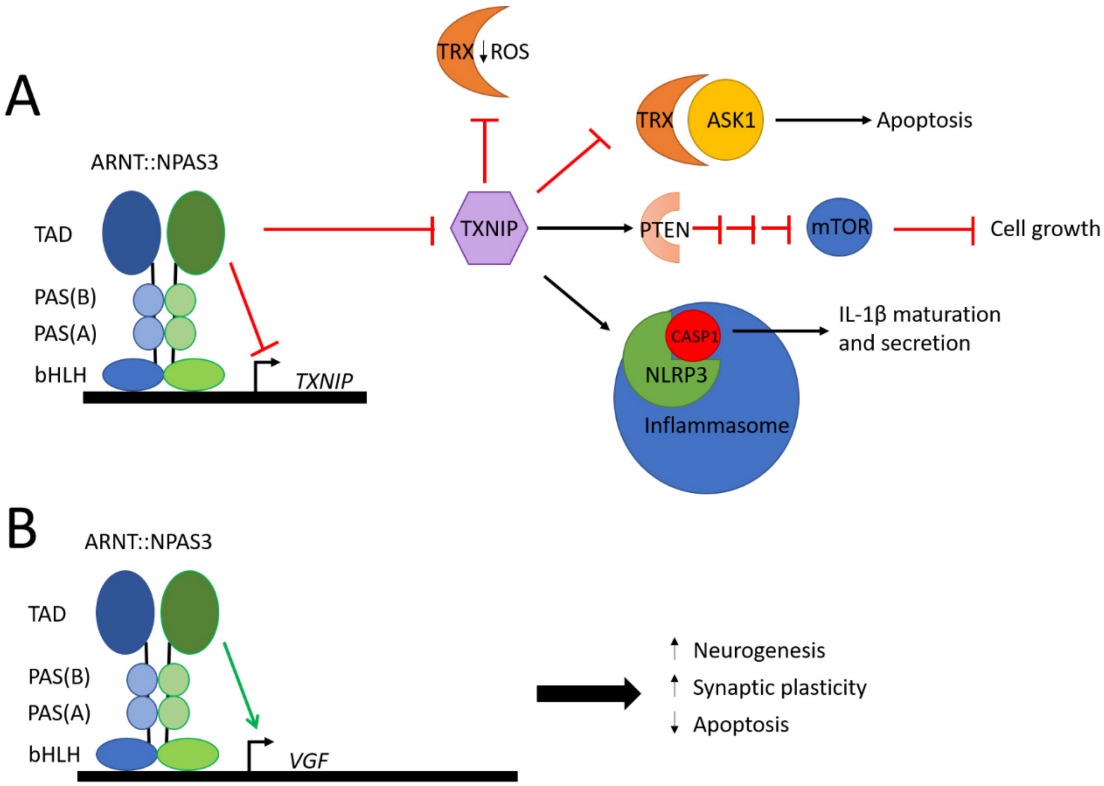



**Figure 6-1 Pathways affected by NPAS3::ARNT.**
(A) Regulatory consequences of NPAS3::ARNT regulation of TXNIP. Inhibition of TXNIP will result in inhibition of apoptotic pathways driven by ASK1, inhibition of mTOR mediated regulation of cell growth and inhibition of the secretion of pro-inflammatory cytokines. (B) Regulatory consequences of NPAS3::ARNT regulation of *VGF* include enhanced neurogenesis, increased synaptic plasticity and reduced apoptosis.

## 6.9 VGF, neurogenesis and neuroprotection

I have demonstrated that *VGF* is a direct target for NPAS3- and ARNT-mediated regulation. *VGF* encodes a neuropeptide which is cleaved to produce multiple secreted peptides with antidepressant actions and a role in energy homeostasis and metabolism (Saderi et al. 2014; Thakker-Varia et al. 2014). The full length VGF polypeptide is cleaved by prohormone convertase-1/3 and -2 into multiple peptides which are then trimmed and processed into mature neuropeptides by other enzymes (Ferri et al. 2011; Pan et al. 2006; Trani et al. 2002; Wardman et al. 2010). Peptides of VGF are stored in vesicles for release when stimulated by depolarization of neurons (Benson and Salton 1996; Possenti et al. 1989). VGF-derived neuropeptides are named by their first four amino acids, followed by their length; for example TLQP-62 is a peptide with sequence starting with Thr-Leu-Gln-Pro, and is 62 aa long (Ferri et al. 2011). The receptors of VGF peptides are largely unknown and different VGF peptides do not interact with the same receptors (Chen et al. 2013; Hannedouche et al. 2013). Although the mechanism by which VGF mediates its effects is complex and poorly understood, the contribution of VGF to neuronal function and neuropsychiatric disorders has been explored in depth.

VGF is a known neuroprotective molecule expressed in hippocampal neurons that is induced by neuropeptides such as NGF and BDNF (Hawley, Scheibe, Wagner 1992; Sakamoto et al. 2015; Salton 1991; Severini et al. 2008; Snyder, Pintar, Salton 1998). In humans, dysregulation of VGF has been associated with psychological disorders: its expression has been shown to be reduced in individuals with major depression, which is rescued by antidepressant treatment, and further, down-regulation of *DISC1*, a gene whose disruption is associated with schizophrenia results in reduction of *VGF* expression

(Cattaneo et al. 2010; Cattaneo et al. 2013; Jiang et al. 2016; Ramos et al. 2014). SNPs in *VGF* have been shown to be associated with increased social anhedonia, an endophenotype of psychosis, as well as decreased maturation and arborization of neurons (Behnke et al. 2017; Ramos et al. 2014). *VGF* expression is also reduced in the hippocampus and dorsolateral prefrontal cortex in individuals with schizophrenia, bipolar disorder and depression, demonstrating widespread dysregulation of *VGF* in neuropsychiatric disorders in a manner consistent with similar associations observed for *NPAS3* (Ramos et al. 2014; Thakker-Varia et al. 2010; Wong et al. 2012).

*VGF* expression is regulated in a manner potentially consistent with regulation by NPAS3 *in vivo. VGF* expression is observed in the developing rat CNS starting around E13.5, temporally correlating with differentiation of the neural cell populations where *VGF* expression was observed, rather than with proliferation (Snyder, Pintar, Salton 1998). *VGF* expression was also observed in developing neuroendocrine cells, consistent with expression in regions known to respond to neutrophins and its role in energy and metabolism regulation (Snyder et al. 2003). Mice deficient in *VGF* are grossly neurologically normal, with marked metabolic dysregulation resulting in pups being 10-20% smaller than wild-type littermate at postnatal day 3 and 50-70% lighter post-weaning. VGF-deficient mice also have abnormal neuropeptide gene expression in response to nutritional cues, along with hyperactivity (Hahm et al. 1999). *Npas3*$^{-/-}$ mice were also observed to be smaller than wild-type littermates and had noted hyperactivity (Brunskill et al. 2005; Erbel-Sieler et al. 2004). Behavioural characterization of *Vgf*$^{-/-}$ mice identified phenotypes similar to those associated with *Npas3* deletion, such as reduced synaptic plasticity and altered neurotransmission in the hippocampus and

reduced performance in memory dependent tasks, including spatial learning and contextual fear conditioning (Bozdagi et al. 2008).

VGF and NPAS3 have similar effects on hippocampal neurogenesis and memory function. Where NPAS3 deficiency is correlated with reduced performance of tasks dependent on hippocampal function, *VGF* expression is up-regulated by hippocampal-dependent learning paradigms (Alder et al. 2003; Brunskill et al. 2005; Erbel-Sieler et al. 2004). Electroconvulsive seizure has been shown to partially rescue *Npas3* knockout-mediated loss of adult neurogenesis and has separately been shown to up-regulate VGF (Newton et al. 2003; Pieper et al. 2005; Snyder et al. 1998). Finally, VGF and Npas3 are both important for adult neurogenesis of hippocampal cells (Pieper et al. 2005; Pieper et al. 2010; Thakker-Varia et al. 2007; Thakker-Varia et al. 2014). The VGF-derived neuropeptide TLQP-62 is best characterized for its role in hippocampal function, synaptic plasticity and antidepressant effects exerted in conjunction with BDNF, another proneurogenic neurotrophic factor (Behnke et al. 2017; Li et al. 2017; Thakker-Varia et al. 2014). Infusion of TLQP-62 into mouse hippocampi results in enhanced neurogenesis, dendritic arborization, synaptic activity and apparent antidepressant effects (Alder et al. 2003; Behnke et al. 2017; Li et al. 2017; Lin et al. 2015; Thakker-Varia et al. 2007; Thakker-Varia et al. 2014). Furthermore, environmental stimulation associated with enhanced neurogenesis and antidepressant effects have been shown to up-regulate *Vgf* in the hippocampus of mice, including exercise and antidepressant medications (Lu et al. 2014; Solvsten et al. 2017; Wang et al. 2015).

NPAS3 and VGF have overlapping functions in the hippocampus, supporting a functional connection between NPAS3 and VGF (Figure 6-1B). The observed activation

of *VGF* by NPAS3 can be predicted to be neuroprotective and to promote neurogenesis. A reduction in *VGF* expression may contribute to the variation in normal and abnormal neuropsychiatric function in individuals carrying variants, deletions and normally expressing *NPAS3*. Furthermore, SNPs identified in *VGF* as associated with neuropsychiatric disorder may modify the effects of variants of *NPAS3*, contributing to the variability of neuropsychological function and neuropsychiatric dysfunction (Behnke et al. 2017).

Although recent ChIP-seq studies in mouse hippocampi did not replicate the finding of *VGF* as a direct target of NPAS3, the effects of *NPAS3* expression on *VGF* expression in neuroprogenitors should be pursued (Michaelson et al. 2017). Studies to confirm how NPAS3 regulates *VGF* should be undertaken in differentiating neuroprogenitors to determine whether NPAS3 regulates *VGF* over the course of differentiation, and whether *VGF* expression correlates with *NPAS3* expression. Further, mass spectrometry should be performed to determine whether expression of *NPAS3* results in production of VGF-derived neuropeptides relevant to the neurogenic and antidepressant functions of VGF. These data will serve to delineate the contribution of NPAS3 regulation of *VGF* to the observed common outcomes in neurogenesis, cognitive functioning and neuropsychiatric disorders.

## 6.10 Endogenous function of NPAS3

*NPAS3* expression was observed to be differentially regulated in response to soluble factors and cell density in HEK 293T cells. Replication of the effects was attempted in SK-N-SH cells, where media exchange was used to stimulate cell growth

instead of quiescence. Assessment of cells during proliferative (72 hours) vs quiescent (120 hours) stages was performed. As can be seen in Figure 5-6, changing media resulted in differential growth trajectory with coordinate effects on markers of proliferation, including the markers of G2/M phases, *CCNB1* and *AURKA*, without affecting apoptosis. *NPAS3* expression was not found to vary consistently over culture time, or in response to exchange of media, at the mRNA or protein level. Furthermore, the expression of *TXNIP* and *VGF* were not found to be affected in a consistent manner.

As NPAS3 has been associated with FGF-mediated neurogenesis, as well as regulation of the FGF/FGFR1 pathway, I assessed the effect of bFGF treatment of cells that are entering quiescence (Pieper et al. 2005; Zhou et al. 2009). I found that there was no change in *NPAS3* expression, nor a coordinate effect on target gene expression suggestive of NPAS3 function. I observed a small increase in markers of proliferative cells suggesting that treatment with mouse bFGF was sufficient to cause a response in the treated cells. However, the possibility remains that this treatment was not sufficient to cause a full response in these cells. Alternately, NPAS3 may act upstream of bFGF, being required in advance of bFGF stimuli, similar to *Drosophila* trachea development, where the NPAS3 orthologue trachealess is required for the initial specification of tracheal cells that undergo tubulogenesis and branching morphogenesis mediated by breathless, the FGFR1 orthologue (Isaac and Andrew 1996; Klambt, Glazer, Shilo 1992; Lee et al. 1996; Ohshiro and Saigo 1997; Reichman-Fried et al. 1994).

As the assessment of NPAS3 function in SK-N-SH cells was hampered by confounding factors, I quantified the relative expression of *NPAS3* and known target genes in neuroprogenitor cells derived from human embryonic stem cells. I found that

*NPAS3* is expressed in these cells, as well as multiple direct targets of NPAS3 (Figure 5-15). These cells appear to be a promising model to assess the function of endogenous NPAS3, once technical limitations are overcome. Human neuroprogenitor cells have been shown to be able to differentiate into multiple functional neuronal cell types through treatment with various neuropeptides and growth factors (Axell, Zlateva, Curtis 2009; Maroof et al. 2013; Woo et al. 2009). As such, assessment of the expression of *NPAS3* during differentiation into multiple neuronal subtypes can be assessed to further understand whether *NPAS3* is expressed in neuroprogenitors that contribute to multiple neuronal cell types, or selective lineages, such as interneurons. Furthermore, direct targets of NPAS3 and the nature of its regulatory effect in neuroprogenitors could be confirmed using neuroprogenitor cells. The hypothesis that NPAS3 binds certain target genes only under stress conditions can also be relatively simply assessed with direct biological relevance to its known role in neurogenesis and neurodevelopment.

## 6.11 Molecular study limitations

The major limitation of the findings of this thesis is the lack of endogenous validation of data obtained by expression of constructs of *NPAS3* and *ARNT*. Although these studies were attempted, characterization of the endogenous function of NPAS3 was hampered by technical limitations. The newly generated NPAS3 antibody was not able to be fully characterized, due to a lack of a cell line known to be deficient for NPAS3 (e.g. an *NPAS3* knockout cell line, which could be assessed in parallel with the parental *NPAS3* expressing line). This partial validation confounded the interpretation of the endogenous expression of *NPAS3* as well as the knockdown of NPAS3 by multiple

shRNA constructs, notably due to variable responsiveness of multiple bands on western blots. This resulted in the inability to conclude with confidence that the bands I am seeing on western blots are, in fact, NPAS3. This antibody should be characterized using a known NPAS3-deficient sample, such as *Npas3* knockout mice and wild-type littermates, or cell lines where *NPAS3* has been deleted.

Althoung we were not able to characterize NPAS3 knockdown at the protein level, we ended up pursuing shRNA construct TRC76 as it is predicted to knockdown *NPAS3* via siRNA mediated silencing, which will result in mRNA degradation that can be quantified by qPCR (Rana 2007). Consistent knockdown of *NPAS3* mRNA was observed with TRC76, and was confirmed with an independent primer pair against *NPAS3* cDNA (Figure 5-10). Despite consistent results at the mRNA level, protein level analysis remained uninterpretable/inconsistent. Further confounding interpretation, the effect of *NPAS3* knockdown resulted in the inability to validate direct targets of NPAS3, including those identified by us, as well as newly-identified NPAS3 targets discussed below.

In a final attempt to characterize the expression of NPAS3 at the protein level, induction of *NPAS3* expression by a doxycyline inducible NPAS3 isoform 1 construct stably integrated into the SK-N-SH cell line genome was performed. Although expression was observed at the mRNA level and protein level based on immunofluorescence, western blots proved difficult to interpret, with no obvious band at the expected ~120 kDa consistently detected with transient expression of the same isoform (Figure 5-12,Figure 5-14). Furthermore, downstream effects of expression of *NPAS3* did not align with the expected up-regulation of *VGF* and down-regulation of *TXNIP*. In

summary, these data suggest significant confounding of some factor in SK-N-SH cells affecting NPAS3 protein or that my antibody does not recognize endogenous NPAS3.

The consistent issue with detection of NPAS3 by western blot may be related to the recent observation of NPAS3 sequestration into an insoluble fraction that was associated with the p.Val304Ile variant (Nucifora et al. 2016). A finding that I observed as imperfectly correlating with the variability in the detection of the ~120 kDa band I predict is endogenous NPAS3, is an ultra-high molecular weight band that is often present, more often or qualitatively more intensely, when the ~120 kDa band is absent (Figure 5-6A, Figure 5-7A, Figure 5-10A, Figure 5-11A, Figure 5-12C). This may represent aggregation of NPAS3. Nuclear aggregations have been observed as pathogenic in neurodegenerative disease, where proteins that oligomerize normally are disrupted resulting in pathological aggregation, is associated with disease (Afroz et al. 2017; DiFiglia et al. 1997). However, whether the observed ultra-high molecular weight bands are biological aggregates or due to technical limitations in preparation of western blots is not clear.

## 6.12 ARNT and ARNT2

The regulatory effects of NPAS3 and ARNT in HEK 293T cells could not be confirmed using endogenous *NPAS3*, expression of an inducible *NPAS3* construct, or knockdown of *NPAS3* in SK-N-SH. Furthermore, I did not consistently detect ARNT in SK-N-SH, and usually more intensely in the cytoplasm than nucleus, suggesting that it is not accessible for interaction with NPAS3 (100 kDa band in NPAS1/ARNT panels of Figure 5-10A, Figure 5-11A, Figure 5-12C). Given that I have demonstrated that *NPAS3*

is expressed in SK-N-SH (Figure 5-5), it may be interacting with a different bHLH-PAS protein partner.

An interaction between NPAS3 and ARNT2 was identified in a yeast two-hybrid (Corominas et al. 2014). This has recently been replicated using mass spectrometry analysis of proteins co-immunoprecipitated with expressed Npas3 in cultured murine neural stem cells (Moen et al. 2017). ARNT and ARNT2 share 81% amino acid sequence identity among their bHLH and PAS domains, with 95% conservation in residues involved in homo- and heterodimerization (Hirose et al. 1996; Kewley, Whitelaw, Chapman-Smith 2004; Wu et al. 2016). Both function as class II bHLH-PAS general heterodimeric partner proteins which can result in different gene-regulatory effects depending on the combination of factors (Hirose et al. 1996; Kewley, Whitelaw, Chapman-Smith 2004; Sekine et al. 2006). For example, both ARNT and ARNT2 are able to interact with AhR and bind target DNA, however, they have differential ability to activate target gene expression (Dougherty and Pollenz 2008; Hirose et al. 1996; Sekine et al. 2006). Alternately, ARNT and ARNT2 can have the same regulatory effect on target promoters, such as when interacting with HIF1A (Drutel et al. 1996; Sekine et al. 2006).

*ARNT* is expressed throughout that body, including in the brain, while *ARNT2* is expressed in a more restricted manner throughout the body and is expressed in the brain of developing mice at a higher level than ARNT over the course of development (Aitola and Pelto-Huikko 2003; Jain et al. 1998). ARNT2 is preferentially expressed in differentiating neurons whereas ARNT is found at higher levels in undifferentiated cells. *ARNT2* levels increase with neuronal differentation (Hao et al. 2013). Given the relative

expression of *ARNT2* in the developing CNS, the differential expression of *ARNT* and *ARNT2* over time, and the detection of an interaction between NPAS3 and ARNT based on a high throughput screen, ARNT2 is an interesting potential partner of NPAS3 which may affect differential regulation by NPAS3 in different tissue types and/or over development and neurodifferentiation. I have found *ARNT2* to be expressed in SK-N-SH cells and human neuroprogenitor cells with approximately two-fold more *ARNT2* mRNA than *ARNT*, as assessed by qPCR. The interaction between NPAS3 and ARNT2 should be directly confirmed in human cells. Further, the relative gene regulatory mechanism/output should be assessed. These data will further our understanding of the role of NPAS3 in neurodevelopment, neurogenesis and neuronal function.

## 6.13 NPAS3 function in neurogenesis

The recent paper reporting RNA seq and ChIP-seq analysis in the hippocampi of *Npas3* knockout mice compared to wild-type littermates, provides the most comprehensive view of the effects of NPAS3 *in vivo* (Michaelson et al. 2017). Direct targets of Npas3 were enriched for genes involved in neurodevelopmental and neuronal processes, as well as genes associated with neuropsychiatric disorders. Npas3 binding sites were found to be enriched upstream of genes up to 290 kb away, suggesting that Npas3 commonly binds enhancer regions. Of note, they identified *Notch1* and *Notch2* as direct regulatory targets of Npas3, and observed altered regulation of other members of the Notch signalling pathway as well. Furthermore the autism- and intellectual disability-associated genes *Fmr1* and *Ube3a* were identified as activated by Npas3 (Chamberlain and Lalande 2010; Geschwind 2011; Michaelson et al. 2017; Suhl and Warren 2015).

My attempts to replicate these findings in SK-N-SH cells were inconsistent, where I observed the expected up-regulation of *NOTCH1* and *NOTCH2* with knockdown of *NPAS3* in one of two replicates, but no other targets tested were found to be regulated (Figure 5-10 B vs C). Notch signalling is important for the repression of neurodifferentiation in neuroprogenitors and astrocytes and has been shown to repress hippocampal neurogenesis in favour of stem cell maintenance (Magnusson et al. 2014; Wilhelmsson et al. 2012; Woo et al. 2009). Loss of Npas3 was associated with up-regulation of *Notch1/2*, consistent with its role in promoting neurogenesis.

Single-cell RNA seq has identified Npas3 as expressed in neural stem cells and a subset of neural precursor cells in the dentate gyrus (Shin et al. 2015). As neural stem cells differentiated into neuronal progenitors, up-regulation of *Txnip* along with the down-regulation of *Npas3* was observed, consistent with Npas3 acting as a repressor of *Txnip* (Shin et al. 2015). Glycolytic enzymes were found to be regulated over the course of neuroprogenitor activation and early differentiation (Shin et al. 2015). Some of these enzymes are differentially regulated by *NPAS3* expression in HEK 293T cells, suggesting that the observed metabolic alterations correlated with deletion of *Npas3* are relevant to neurogenic processes (Sha et al. 2012; Shin et al. 2015). These data support a role for Npas3 in the proliferation of neuroprogenitors in the subgranular zone of the dentate gyrus. Through regulating *Txnip*, Npas3 may contribute to the coordinate effects on cellular metabolism and growth in this neurogenic niche (Pieper et al. 2010; Sha et al. 2012). A similar trend was observed in neural stem cells derived from the mouse subventricular zone, where Npas3 was found to be up-regulated upon neural stem cell proliferation (Dulken et al. 2017). In this study *Txnip* expression was also found to be

reciprocally regulated to *Npas3*. These data indicate that Npas3 is commonly regulated during neural stem cell proliferation in multiple neurogenic niches in mice, however, further assessment of its regulation and function over the course of neurogenesis is required.

These findings should be replicated in human cells to confirm that similar effects occur in human neuroprogenitors. To this end, I have characterized ENStem-A neuroprogenitor cells derived from human embryonic stem cells purchased from Millipore as being *NPAS3*-expressing and thus a logical avenue for assessment of NPAS3 function in neurogenesis. As these cells can be differentiated into multiple neuronal cell types, assessment of the expression of NPAS3 under different neurodifferentiation paths could be undertaken (Axell, Zlateva, Curtis 2009; Maroof et al. 2013; Ozolek et al. 2010; Schulz et al. 2003). The regulation of *TXNIP* and *VGF* by NPAS3 can be assessed in these cells, including assessment of regulation under stress conditions such as hypoxia, oxidative stress, glucose stress and inflammatory factors. Given the limitations of passage number of these cells before spontaneous neurodifferentiation occurs (Millipore 2012), it may be advantageous to generate CRISPR-mediated knockouts for *NPAS3* expression in human embryonic stem cells and generate a neural stem cell population that is *NPAS3*-deficient in order to assess the role of NPAS3 in the stem cell maintenance, differentiation potential and cellular stress response of the generated NPAS3 deleted human neuroprogenitors (Cong et al. 2013). CRISPR has also been used to insert various types of mutations into the genome, which would facilitate assessment of the contribution of the variants assessed in this study, as well as those otherwise associated with

neuropsychiatric dysfunction, to the endogenous function of NPAS3 in neuroprogenitors (Wojtal et al. 2016).

The potential regulation of *NPAS3* by cell autonomous and soluble factors, which may include IL-1β and/or IL-6, may be relevant for hippocampal neurogenesis. *NPAS3* genotype has previously been associated with circulating IL-1β and IL-6 levels in the context of metabolic syndrome in obese individuals (Zhang et al. 2013b). I identified up-regulation of *NPAS3* expression as a function of cell contact and soluble factors related to high glucose concentration and cellular stress, indicated by up-regulation of *TXNIP*. As *TXNIP* up-regulation is known to cause secretion of the pro-inflammatory cytokine IL-1β in high glucose conditions, and IL-6 under oxidative stress, the potential link between inflammatory cytokines and *NPAS3* expression is of great interest (Gao et al. 2015; Zhou et al. 2010). Furthermore, astrocytes have been shown to promote neurogenesis in adult neuroproliferative zones, through soluble and cell-contact mediated factors (Song, Stevens, Gage 2002; Wilhelmsson et al. 2012). Specifically, release of IL-1β and IL-6 from astrocytes have been shown to promote neurodifferentiation in neuroprogenitor cells, and as such, the relationship between IL-1β and IL-6, *NPAS3* expression and neurodifferentiation should be explored (Barkho et al. 2006). In summary, the expression of *NPAS3* in response to secreted markers of inflammation and apoptosis is of interest, as they may affect the rates of neuroproliferation and differentiation of neuroprogenitor cells through regulation of *NPAS3* expression.

## 6.14 *NPAS3* and its contribution to neuropsychological function

In order to connect the molecular function of NPAS3 to the genetic association of *NPAS3* with neuropsychiatric disorders with intellectual disability as a common feature, we undertook a pilot study to assess the contribution of *NPAS3* variation to normal variation in human cognition. Through this study we have identified the first association between *NPAS3* genotype and altered neurocognition, an endophenotype of neuropsychiatric disorders. Specifically, the minor allele of a coding SNP of *NPAS3*, c.1654G>C (p.Ala552Pro) was found to be associated with reduced performance on a verbal working memory task (Table 8). This association was also confirmed for the linked non-coding SNPs c.2208C>T and c.2262C>G (Table 9).

This pilot study is limited by its small sample size with relatively low statistical power. This contributed to our inability to assess dosage effects of the minor allele, as only two individuals homozygous for the minor allele were identified. Confirmation of the association between *NPAS3* genotype and cognitive function should be pursued in a larger population. We were surprised to find that the observed frequency of the minor (previously schizophrenia-associated) alleles of c.1654G>C, c.2208C>T and c.2262C>G in our cohort were not significantly different from the schizophrenia cohort of the index study (Table 7). However, subsequent analysis has determined that our observed allele frequencies are not significantly different from the expected population frequency (Lek et al. 2016; Macintyre et al. 2010; National Center for Biotechnology Information, National Library of Medicine 2017). Genotyping analysis of *NPAS3* by another group has also found similar genotype frequencies in a cohort of autism patients and controls (Stanco et al. 2014). A potential explanation for this may be due to age effects, as the current sample

is much younger than the original sample, which was past the onset age of many neuropsychiatric disorders and thus can be predicted to have a reduced load for psychiatric illness-associated variants, as the population was selected for unaffected individuals (Macintyre et al. 2010). Furthermore, the current population is largely female, whereas the previous population was largely male, which may contribute to differences observed. These data were collected with the intention of co-registering them with other populations, both of cross-sectional populations, as well as psychiatrically ill populations to replicate the results in larger samples and determine if this association with cognitive function is concordant regardless of disorder status (state-independent) and thus a true endophenotype. As such, these data are intriguing preliminary findings that should be replicated and studied further in order to expand our understanding of the role of variation to *NPAS3* in neuropsychological function in both normal and psychiatrically ill individuals.

The assessed SNPs have been hypothesized to affect NPAS3 function at both the mRNA and protein level. In this study, I have molecularly characterized the protein-level effects of c.1654G>C (p.Ala552Pro) on NPAS3 function and have found that there are no severe effects of this variant on protein function, however, this does not preclude effects of this variant on functions I have not assessed, nor that the assays used do not have the power to assess (discussed in a previous section). For example c.1654G>C is predicted to create a novel splice enhancer element and the linked non-coding variants are predicted to affect splicing function, with some reduction in splice enhancer function predicted for the c.2208C>T variant and significant reduction in splicing enhancer activity predicted

for the c.2262C>G variant (Macintyre et al. 2010). This may affect the relative expression of functional NPAS3, or production of alternate isoforms.

Reduction in *NPAS3* expression, or altered function of its target genes, can be connected to the observed effects on working memory function by analysis of its roles in cellular function, neurodevelopment and neurogenesis. I have identified *VGF* and *TXNIP* as direct target genes of NPAS3 with critical roles in energy homeostasis, metabolic balance, cellular response to environmental stimuli, neurogenesis, and signalling of inflammation (Chong et al. 2014; Ferri et al. 2011; Thakker-Varia and Alder 2009). Based on the function of *TXNIP* and my finding that NPAS3::ARNT represses *TXNIP*, reduction in NPAS3 function is expected to cause an increase in *TXNIP* expression that would lead to an enhanced response to environmental stresses, such as hyperglycemia, hypoxia and excitotoxicity, resulting in altered metabolic function, increased rates of apoptosis and increased inflammatory tone (Chen et al. 2008; Gao et al. 2015; Kim et al. 2012; Kristiansen et al. 2011; Li et al. 2015; Price et al. 2006; Zaragoza-Campillo and Moran 2017). Further, given the functions of VGF and my finding that NPAS3::ARNT directly activates *VGF* expression, reduction in *NPAS3* expression or function may cause reduced levels of VGF-derived neuroprotective neuropeptides, resulting in increased apoptosis, contribute to decreased neurogenesis in the hippocampus associated with depression, as well as learning and memory deficits (Alder et al. 2003; Behnke et al. 2017; Bozdagi et al. 2008; Severini et al. 2008). These findings can be directly linked to the observed effects of deletion of *Npas3* in mice, where altered behavioral tone (mood) was observed, alongside reduced learning and memory which was associated with an

almost complete loss of FGF-mediated neurogenesis in the dentate gyrus of the hippocampus (Brunskill et al. 2005; Erbel-Sieler et al. 2004; Pieper et al. 2005).

Working memory function has been shown to be associated with activity in the frontal and parietal cortexes in both healthy and psychiatrically ill populations (Callicott et al. 2003; Drapier et al. 2008; Glahn et al. 2005; Klingberg, Forssberg, Westerberg 2002; Seidman et al. 2006). Proper balance of excitatory (dopaminergic and glutamatergic) and inhibitory (GABAergic) signalling is critical for optimal working memory performance (Karlsgodt et al. 2011; Schwarz, Tost, Meyer-Lindenberg 2015). As such, variation in the tone of neuronal circuits in the frontal and parietal cortical regions required for working memory processes can be predicted to affect its functioning. Given that *Npas3* is involved in the generation of interneurons during embryogenesis, and that it co-stains with interneuronal markers throughout the cortex postnatally, I can hypothesize that variation in *NPAS3* that affect its expression and/or function may contribute to altered working memory function by affecting the number of interneurons that are generated and able to migrate from the MGE and/or CGE to cortical regions critical for proper working memory function (Stanco et al. 2014). *Npas3* knockout mice display altered dopaminergic signalling, as assessed by treatment with pharmacologic modulators of dopamine receptors D1R- and D2R-mediated neurotransmission, which may be another symptom of reduced interneuron incorporation into the cortex, resulting in hyperexcitability of response to the dopaminergic and/or glutamatergic neurotransmitters (Brunskill et al. 2005; Stanco et al. 2014). As *NPAS3* expression has been detected in the cortex of humans from infancy through adulthood and altered cortical NPAS3 expression has been associated with schizophrenia, the role of NPAS3 in

generation of neurons that incorporate into the cortex appears to be conserved between mice and humans, suggesting common function and contributions to neuropsychological function (Wong et al. 2012).

Genome-wide association studies assessing the contribution of genetic variation to cognitive function are starting to identify loci associated with intellectual achievement and cognitive functions. Preliminary studies were hampered by small sample sizes and unable to identify significantly associated genetic loci (Benyamin et al. 2014; Davies et al. 2011; Lencz et al. 2014). In a meta-analysis including a subset of these studies a SNP within *AKAP6*, 105 kb centromeric to *NPAS3* was identified as significantly associated with cognitive function at a genome-wide level (Davies et al. 2015). Other large meta-analyses have continued to identify significant loci as associated with general cognitive function (Trampush et al. 2017) and of educational attainment as a proxy measure found to share genetic correlation with cognitive function (Okbay et al. 2016; Rietveld et al. 2013; Rietveld et al. 2014). Genes associated with cognitive function are enriched for neuronal processes and previously identified neuropsychiatric disorder associated genes, however linkage to *NPAS3* has not been replicated in these studies (Davies et al. 2016; Okbay et al. 2016).

Meta-analyses of specific cognitive measures are limited due to variation in cognitive testing paradigms between studies. In one large meta-analysis, genetic associations have been found for verbal-numberical reasoning and reaction time but not for memory (Davies et al. 2016). In order to assess specific cognitive functions the selection of cognitive measures and assays were found to be critical for success in identifying linkages to genetic loci (Davies et al. 2016; Trampush et al. 2017).

Identification of loci implicated in specific cognitive function at the genome-wide level will require coordination of large studies using rigorous and consistent cognitive testing batteries that are administered in a consistent fashion (Greenwood et al. 2007; Greenwood et al. 2013).

As *NPAS3* has been associated with neuropsychiatric disorders and intellectual disability through both common variants present in normal and psychiatrically ill individuals have relatively subtle effects (Macintyre et al. 2010; Nurnberger et al. 2014; Pickard et al. 2008; Weber et al. 2011) as well as by rare *de novo* events that cause marked neurological dysfunction (Kamnasaran et al. 2003; Kamnasaran et al. 2005; Phelps et al. 2016; Piccione et al. 2012; Pickard et al. 2005; Rosenfeld et al. 2010; Visser et al. 2010), aberrations in *NPAS3* function can be seen to have dosage effects. This is consistent with the model that neuropsychiatric disorders and variation in normal cognition are due to a combination of common variation and rare variants of large effect sizes (Burmeister, McInnis, Zollner 2008).

The specific contribution of *NPAS3* disruption, or dysregulation, in these individuals can be linked to both alterations in neurogenesis during development and adult neurogenic potential. Complete disruption of *NPAS3* could cause severe deficits in the neuroproliferative potential of *NPAS3*-expressing neural stem cells during embryogenesis, which would contribute to reduced neuronal cell numbers in far-reaching regions of the brain, including the cortex, basal ganglia, midbrain and hindbrain resulting in the severe neuropsychiatric, cognitive and neuroanatomical deficits observed in individuals with deletions affecting *NPAS3*. Variants that cause minor deficits in NPAS3 function would result in relatively more minor, but widespread alterations in neuronal

circuit tone, as well as reduction in the relative rates of hippocampal neurogenesis, which have been associated with reduced memory processes as well as altered mood, and can be seen to cause alterations in neurocognitive function in individuals carrying these variants. Furthermore, variation in *NPAS3* may correlate with environmental factors as its target gene *TXNIP* is involved in cellular response to stress. As such *NPAS3* genotype may interact with environmental factors to cause differential effects on neuropsychological outcomes.

Here, I have linked altered working memory with *NPAS3* genotype. Reduced working memory is associated with multiple neuropsychiatric disorders and multiple genes associated with neuropsychiatric disorders have been shown to contribute to deficits in working memory function (Schwarz, Tost, Meyer-Lindenberg 2015). As such, variation in *NPAS3* expression and function can be seen to affect both the variability in normal and disordered neuropsychological functioning, by affecting the survival and responsiveness of neuroprogenitors of inhibitory neurons and others during neurodevelopment affecting cortical tone, which can be seen to affect function of wide-ranging neuropsychological functions, including working memory. When combined with altered survival of proliferating neuroprogenitor cells in the hippocampus over childhood and young adulthood, variation in NPAS3 function can result in additive effects leading to altered cognitive performance and potentially psychiatric disorders, depending on the cumulative load of genetic and environmental risk factors for neuropsychiatric disease. Combined with environmental and other genetic factors, *NPAS3* variation could contribute to the normal and clinical variation in neuropsychological function.

# Chapter 7: Concluding remarks

Through the studies undertaken in this thesis project, I aimed to further our understanding of the function of NPAS3 at the molecular and cellular level. Furthermore, our collaboration attempted to determine how variation in *NPAS3* contributes to both the severe neurodevelopmental disorders and psychiatric disorders with which it has been linked, by assessing its contribution to variation in normal cognitive functioning. We have identified the first cognitive endophenotype associated with variation of *NPAS3* present in humans. These data provide insight into how common variants linked to and/or affecting expression and function of *NPAS3* can affect the neuropsychological functioning of the individuals who carry them, both those with neuropsychiatric disease and those with normal neuropsychological function.

I have also demonstrated that human NPAS3 and ARNT interact and cooperatively regulate target genes. I have identified two direct targets of NPAS3 and ARNT, *TXNIP* and *VGF,* and have explored how their function and dysregulation might correlate with neuropsychological function and neuropsychiatric dysfunction, among other disorders, most notably metabolic dysregulation. Furthermore, I have functionally characterized the domains of NPAS3, confirming that the bHLH domain binds to DNA and participates in protein::protein interaction, that the PAS domain is involved in protein::protein interaction, and that the C-terminus of NPAS3 does encode a potent transactivation domain and a functional nuclear localization signal. In addition, I have characterized three variants of *NPAS3* that have been identified in humans, and have found one variant to have a potential deficit in transactivation function, which warrants

further characterization. These data are limited by my inability to confirm the findings in a system endogenously expressing *NPAS3*.

Finally, I undertook characterization of NPAS3 regulation and function at the cellular level as it relates to the regulation of the target gene *TXNIP*. *NPAS3* was found to be differentially regulated by secreted factors associated with high glucose, a stress condition. This up-regulation correlated with an increase of *TXNIP*, which is responding to the glycolytic stress of the cell. However, whether or not NPAS3 affects that regulation remains unclear. In neuroblastoma cells, *NPAS3* was found to be expressed, however technical difficulties prevented characterization of knockdown and expression constructs, and hampered interpretation of experiments. Characterization of human embryonic stem cell derived neuroprogenitor cells was undertaken and *NPAS3* found to be expressed, supporting its role in human neurodevelopment and neurogenesis, and a promising avenue for further assessment of NPAS3 function in human neurogenesis.

Through these studies, I have expanded our understanding of the molecular function of NPAS3 and opened up many lines of future questioning, including the function of NPAS3 in human neuroprogenitors, its ability to interact with other bHLH-PAS proteins, as well as finer characterization of its functional domains. Development of cells that do not express *NPAS3*, such as CRISPR-mediated *NPAS3* deletions should be undertaken, to support and advance the data presented here and will help characterize the antibody to NPAS3 that I have generated, as well as potentially allow for re-examination of the data and samples obtained for further assessment of NPAS3 function and its role in neuropsychological function and dysfunction.

# References


Addgene: pLKO.1 - TRC Cloning Vector [Internet]. Cambridge, Mass: Addgene; c2006. Available from: https://www.addgene.org/tools/protocols/plko .

Adesida AB, Grady LM, Khan WS, Hardingham TE. 2006. The matrix-forming phenotype of cultured human meniscus cells is enhanced after culture with fibroblast growth factor 2 and is further stimulated by hypoxia. Arthritis Res Ther 8(3):R61.

Afroz T, Hock EM, Ernst P, Foglieni C, Jambeau M, Gilhespy LAB, Laferriere F, Maniecka Z, Pluckthun A, Mittl P, et al. 2017. Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. Nat Commun 8(1):45,017-00062-0.

Aitola MH and Pelto-Huikko MT. 2003. Expression of arnt and Arnt2 mRNA in developing murine tissues. J Histochem Cytochem 51(1):41-54.

Alaerts M and Del-Favero J. 2009. Searching genetic risk factors for schizophrenia and bipolar disorder: Learn from the past and back to the future. Hum Mutat 30(8):1139-52.

Alba MM and Guigo R. 2004. Comparative analysis of amino acid repeats in rodents and humans. Genome Res 14(4):549-54.

Alder J, Thakker-Varia S, Bangasser DA, Kuroiwa M, Plummer MR, Shors TJ, Black IB. 2003. Brain-derived neurotrophic factor-induced gene expression reveals novel actions of VGF in hippocampal synaptic plasticity. J Neurosci 23(34):10800-8.

Anderson SA, Eisenstat DD, Shi L, Rubenstein JL. 1997. Interneuron migration from basal forebrain to neocortex: Dependence on dlx genes. Science 278(5337):474-6.

Anderson SA, Kaznowski CE, Horn C, Rubenstein JL, McConnell SK. 2002. Distinct origins of neocortical projection neurons and interneurons in vivo. Cereb Cortex 12(7):702-9.

Ando J, Ono Y, Wright MJ. 2001. Genetic structure of spatial and verbal working memory. Behav Genet 31(6):615-24.

Aviner R, Shenoy A, Elroy-Stein O, Geiger T. 2015. Uncovering hidden layers of cell cycle regulation through integrative multi-omic analysis. PLoS Genet 11(10):e1005554.

Axell MZ, Zlateva S, Curtis M. 2009. A method for rapid derivation and propagation of neural progenitors from human embryonic stem cells. J Neurosci Methods 184(2):275-84.

Baddeley A. 2003. Working memory: Looking back and looking forward. Nat Rev Neurosci 4(10):829-39.

Baek ST, Copeland B, Yun EJ, Kwon SK, Guemez-Gamboa A, Schaffer AE, Kim S, Kang HC, Song S, Mathern GW, et al. 2015. An AKT3-FOXG1-reelin network underlies defective migration in human focal malformations of cortical development. Nat Med 21(12):1445-54.

Baker AF, Koh MY, Williams RR, James B, Wang H, Tate WR, Gallegos A, Von Hoff DD, Han H, Powis G. 2008. Identification of thioredoxin-interacting protein 1 as a hypoxia-inducible factor 1alpha-induced gene in pancreatic cancer. Pancreas 36(2):178-86.

Balsalobre A, Damiola F, Schibler U. 1998. A serum shock induces circadian gene expression in mammalian tissue culture cells. Cell 93(6):929-37.
Bao BY, Lin VC, Yu CC, Yin HL, Chang TY, Lu TL, Lee HZ, Pao JB, Huang CY, Huang SP. 2016. Genetic variants in ultraconserved regions associate with prostate cancer recurrence and survival. Sci Rep 6:22124.
Barkho BZ, Song H, Aimone JB, Smrt RD, Kuwabara T, Nakashima K, Gage FH, Zhao X. 2006. Identification of astrocyte-expressed factors that modulate neural stem/progenitor cell differentiation. Stem Cells Dev 15(3):407-21.
Batista-Brito R, Machold R, Klein C, Fishell G. 2008. Gene expression in cortical interneuron precursors is prescient of their mature function. Cereb Cortex 18(10):2306-17.
Bavarva JH, Tae H, Settlage RE, Garner HR. 2013. Characterizing the genetic basis for nicotine induced cancer development: A transcriptome sequencing study. PLoS One 8(6):e67252.
Behnke J, Cheedalla A, Bhatt V, Bhat M, Teng S, Palmieri A, Windon CC, Thakker-Varia S, Alder J. 2017. Neuropeptide VGF promotes maturation of hippocampal dendrites that is reduced by single nucleotide polymorphisms. Int J Mol Sci 18(3):10.3390/ijms18030612.
Benita Y, Kikuchi H, Smith AD, Zhang MQ, Chung DC, Xavier RJ. 2009. An integrative genomics approach identifies hypoxia inducible factor-1 (HIF-1)-target genes that form the core response to hypoxia. Nucleic Acids Res 37(14):4587-602.
Benson DL and Salton SR. 1996. Expression and polarization of VGF in developing hippocampal neurons. Brain Res Dev Brain Res 96(1-2):219-28.
Benyamin B, Pourcain B, Davis OS, Davies G, Hansell NK, Brion MJ, Kirkpatrick RM, Cents RA, Franic S, Miller MB, et al. 2014. Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. Mol Psychiatry 19(2):253-8.
Bhattacharyya A, Pathak S, Basak C, Law S, Kundu M, Basu J. 2003. Execution of macrophage apoptosis by mycobacterium avium through apoptosis signal-regulating kinase 1/p38 mitogen-activated protein kinase signaling and caspase 8 activation. J Biol Chem 278(29):26517-25.
Blumenthal MN, Langefeld CD, Beaty TH, Bleecker ER, Ober C, Lester L, Lange E, Barnes KC, Wolf R, King RA, et al. 2004. A genome-wide search for allergic response (atopy) genes in three ethnic groups: Collaborative study on the genetics of asthma. Hum Genet 114(2):157-64.
Bozdagi O, Rich E, Tronel S, Sadahiro M, Patterson K, Shapiro ML, Alberini CM, Huntley GW, Salton SR. 2008. The neurotrophin-inducible gene vgf regulates hippocampal function and behavior through a brain-derived neurotrophic factor-dependent mechanism. J Neurosci 28(39):9857-69.
Bracken CP, Fedele AO, Linke S, Balrak W, Lisy K, Whitelaw ML, Peet DJ. 2006. Cell-specific regulation of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha stabilization and transactivation in a graded oxygen environment. J Biol Chem 281(32):22575-85.
Brito M, Malta-Vacas J, Carmona B, Aires C, Costa P, Martins AP, Ramos S, Conde AR, Monteiro C. 2005. Polyglycine expansions in eRF3/GSPT1 are associated with gastric cancer susceptibility. Carcinogenesis 26(12):2046-9.

Brockschmidt FF, Nothen MM, Hillmer AM. 2007. The two most common alleles of the coding GGN repeat in the androgen receptor gene cause differences in protein function. J Mol Endocrinol 39(1):1-8.
Broos S, Hulpiau P, Galle J, Hooghe B, Van Roy F, De Bleser P. 2011. ConTra v2: A tool to identify transcription factor binding sites across species, update 2011. Nucleic Acids Res 39(Web Server issue):W74-8.
Brown KN, Chen S, Han Z, Lu CH, Tan X, Zhang XJ, Ding L, Lopez-Cruz A, Saur D, Anderson SA, et al. 2011. Clonal production and organization of inhibitory interneurons in the neocortex. Science 334(6055):480-6.
Brunskill EW, Witte DP, Shreiner AB, Potter SS. 1999. Characterization of npas3, a novel basic helix-loop-helix PAS gene expressed in the developing mouse nervous system. Mech Dev 88(2):237-41.
Brunskill EW, Ehrman LA, Williams MT, Klanke J, Hammer D, Schaefer TL, Sah R, Dorn GW,2nd, Potter SS, Vorhees CV. 2005. Abnormal neurodevelopment, neurosignaling and behaviour in Npas3-deficient mice. Eur J Neurosci 22(6):1265-76.
Burmeister M, McInnis MG, Zollner S. 2008. Psychiatric genetics: Progress amid controversy. Nat Rev Genet 9(7):527-40.
Callicott JH, Egan MF, Mattay VS, Bertolino A, Bone AD, Verchinksi B, Weinberger DR. 2003. Abnormal fMRI response of the dorsolateral prefrontal cortex in cognitively intact siblings of patients with schizophrenia. Am J Psychiatry 160(4):709-19.
Cao G, Jiang N, Hu Y, Zhang Y, Wang G, Yin M, Ma X, Zhou K, Qi J, Yu B, et al. 2016. Ruscogenin attenuates cerebral ischemia-induced blood-brain barrier dysfunction by suppressing TXNIP/NLRP3 inflammasome activation and the MAPK pathway. Int J Mol Sci 17(9):10.3390/ijms17091418.
Cattaneo A, Sesta A, Calabrese F, Nielsen G, Riva MA, Gennarelli M. 2010. The expression of VGF is reduced in leukocytes of depressed patients and it is restored by effective antidepressant treatment. Neuropsychopharmacology 35(7):1423-8.
Cattaneo A, Gennarelli M, Uher R, Breen G, Farmer A, Aitchison KJ, Craig IW, Anacker C, Zunsztain PA, McGuffin P, et al. 2013. Candidate genes expression profile associated with antidepressants response in the GENDEP study: Differentiating between baseline 'predictors' and longitudinal 'targets'. Neuropsychopharmacology 38(3):377-85.
Chai TF, Leck YC, He H, Yu FX, Luo Y, Hagen T. 2011. Hypoxia-inducible factor independent down-regulation of thioredoxin-interacting protein in hypoxia. FEBS Lett 585(3):492-8.
Chamberlain SJ and Lalande M. 2010. Angelman syndrome, a genomic imprinting disorder of the brain. J Neurosci 30(30):9958-63.
Cha-Molstad H, Saxena G, Chen J, Shalev A. 2009. Glucose-stimulated expression of txnip is mediated by carbohydrate response element-binding protein, p300, and histone H4 acetylation in pancreatic beta cells. J Biol Chem 284(25):16898-905.
Chapman LJ and Chapman JP. 1987. The search for symptoms predictive of schizophrenia. Schizophr Bull 13(3):497-503.

Chapman-Smith A, Lutwyche JK, Whitelaw ML. 2004. Contribution of the per/arnt/sim (PAS) domains to DNA binding by the basic helix-loop-helix PAS transcriptional regulators. J Biol Chem 279(7):5353-62.
Chau GC, Im DU, Kang TM, Bae JM, Kim W, Pyo S, Moon EY, Um SH. 2017. mTOR controls ChREBP transcriptional activity and pancreatic beta cell survival under diabetic stress. J Cell Biol 216(7):2091-105.
Chen J, Saxena G, Mungrue IN, Lusis AJ, Shalev A. 2008. Thioredoxin-interacting protein: A critical link between glucose toxicity and beta-cell apoptosis. Diabetes 57(4):938-44.
Chen YC, Pristera A, Ayub M, Swanwick RS, Karu K, Hamada Y, Rice AS, Okuse K. 2013. Identification of a receptor for neuropeptide VGF and its role in neuropathic pain. J Biol Chem 288(48):34638-46.
Chilov D, Camenisch G, Kvietikova I, Ziegler U, Gassmann M, Wenger RH. 1999. Induction and nuclear translocation of hypoxia-inducible factor-1 (HIF-1): Heterodimerization with ARNT is not necessary for nuclear accumulation of HIF-1alpha. J Cell Sci 112 ( Pt 8)(Pt 8):1203-12.
Chong CR, Chan WP, Nguyen TH, Liu S, Procter NE, Ngo DT, Sverdlov AL, Chirkov YY, Horowitz JD. 2014. Thioredoxin-interacting protein: Pathophysiology and emerging pharmacotherapeutics in cardiovascular disease and diabetes. Cardiovasc Drugs Ther 28(4):347-60.
Cobos I, Long JE, Thwin MT, Rubenstein JL. 2006. Cellular patterns of transcription factor expression in developing cortical interneurons. Cereb Cortex 16 Suppl 1:i82-8.
Collaborative Study on the Genetics of Asthma (CSGA). 1997. A genome-wide search for asthma susceptibility loci in ethnically diverse populations. the collaborative study on the genetics of asthma (CSGA). Nat Genet 15(4):389-92.
Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. 2013. Multiplex genome engineering using CRISPR/cas systems. Science 339(6121):819-23.
Conners CK. 1994. The conners continuous performance test. Toronto, Canada: Multi-Health Systems.
Corominas R, Yang X, Lin GN, Kang S, Shen Y, Ghamsari L, Broly M, Rodriguez M, Tam S, Trigg SA, et al. 2014. Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. Nat Commun 5:3650.
Cowan N. 2008. What are the differences between long-term, short-term, and working memory? Prog Brain Res 169:323-38.
Cross-Disorder Group of the Psychiatric Genomics Consortium, Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, Mowry BJ, Thapar A, Goddard ME, et al. 2013. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet 45(9):984-94.
Cuesta MJ, Pino O, Guilera G, Rojo JE, Gomez-Benito J, Purdon SE, Franco M, Martinez-Aran A, Segarra N, Tabares-Seisdedos R, et al. 2011. Brief cognitive assessment instruments in schizophrenia and bipolar patients, and healthy control subjects: A comparison study between the brief cognitive assessment tool for schizophrenia (B-CATS) and the screen for cognitive impairment in psychiatry (SCIP). Schizophr Res 130(1-3):137-42.

Cummings CJ and Zoghbi HY. 2000. Trinucleotide repeats: Mechanisms and pathophysiology. Annu Rev Genomics Hum Genet 1:281-328.
Daftuar L, Zhu Y, Jacq X, Prives C. 2013. Ribosomal proteins RPL37, RPS15 and RPS20 regulate the Mdm2-p53-MdmX network. PLoS One 8(7):e68667.
D'Arcangelo G and Halegoua S. 1993. A branched signaling pathway for nerve growth factor is revealed by src-, ras-, and raf-mediated gene inductions. Mol Cell Biol 13(6):3146-55.
D'Arcangelo G, Habas R, Wang S, Halegoua S, Salton SR. 1996. Activation of codependent transcription factors is required for transcriptional induction of the vgf gene by nerve growth factor and ras. Mol Cell Biol 16(9):4621-31.
Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaars SP, Harris SE, Ritchie SJ, Luciano M, Fawns-Ritchie C, Lyall D, et al. 2016. Genome-wide association study of cognitive functions and educational attainment in UK biobank (N=112 151). Mol Psychiatry 21(6):758-67.
Davies G, Tenesa A, Payton A, Yang J, Harris SE, Liewald D, Ke X, Le Hellard S, Christoforou A, Luciano M, et al. 2011. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. Mol Psychiatry 16(10):996-1005.
Davies G, Armstrong N, Bis JC, Bressler J, Chouraki V, Giddaluru S, Hofer E, Ibrahim-Verbaas CA, Kirin M, Lahti J, et al. 2015. Genetic contributions to variation in general cognitive function: A meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). Mol Psychiatry 20(2):183-92.
de Candia P, Blekhman R, Chabot AE, Oshlack A, Gilad Y. 2008. A combination of genomic approaches reveals the role of FOXO1a in regulating an oxidative stress response pathway. PLoS One 3(2):e1670.
de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N. 2006. ScanProsite: Detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34(Web Server issue):W362-5.
Della Sala, S., Gray, C., Baddeley, A., Wilson, L. 1997. Visual patterns test: A test of short-term visual recall. Bury St Edmunds, England: Thames Valley Test Company.
Di Rocco G, Pennuto M, Illi B, Canu N, Filocamo G, Trani E, Rinaldi AM, Possenti R, Mandolesi G, Sirinian MI, et al. 1997. Interplay of the E box, the cyclic AMP response element, and HTF4/HEB in transcriptional regulation of the neurospecific, neurotrophin-inducible vgf gene. Mol Cell Biol 17(3):1244-53.
DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, Vonsattel JP, Aronin N. 1997. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. Science 277(5334):1990-3.
Dinarello CA. 2011. A clinical perspective of IL-1beta as the gatekeeper of inflammation. Eur J Immunol 41(5):1203-17.
Dioum EM, Rutter J, Tuckerman JR, Gonzalez G, Gilles-Gonzalez MA, McKnight SL. 2002. NPAS2: A gas-responsive transcription factor. Science 298(5602):2385-7.
Donnelly KL, Margosian MR, Sheth SS, Lusis AJ, Parks EJ. 2004. Increased lipogenesis and fatty acid reesterification contribute to hepatic triacylglycerol stores in hyperlipidemic txnip-/- mice. J Nutr 134(6):1475-80.

Dougherty EJ and Pollenz RS. 2008. Analysis of ah receptor-ARNT and ah receptor-ARNT2 complexes in vitro and in cell culture. Toxicol Sci 103(1):191-206.
Drapier D, Surguladze S, Marshall N, Schulze K, Fern A, Hall MH, Walshe M, Murray RM, McDonald C. 2008. Genetic liability for bipolar disorder is characterized by excess frontal activation in response to a working memory task. Biol Psychiatry 64(6):513-20.
Drutel G, Kathmann M, Heron A, Schwartz JC, Arrang JM. 1996. Cloning and selective expression in brain and kidney of ARNT2 homologous to the ah receptor nuclear translocator (ARNT). Biochem Biophys Res Commun 225(2):333-9.
Dudley CA, Erbel-Sieler C, Estill SJ, Reick M, Franken P, Pitts S, McKnight SL. 2003. Altered patterns of sleep and behavioral adaptability in NPAS2-deficient mice. Science 301(5631):379-83.
Dulken BW, Leeman DS, Boutet SC, Hebestreit K, Brunet A. 2017. Single-cell transcriptomic analysis defines heterogeneity and transcriptional dynamics in the adult neural stem cell lineage. Cell Rep 18(3):777-90.
Dupret D, Revest JM, Koehl M, Ichas F, De Giorgi F, Costet P, Abrous DN, Piazza PV. 2008. Spatial relational memory requires hippocampal adult neurogenesis. PLoS One 3(4):e1959.
Eisenstat DD, Liu JK, Mione M, Zhong W, Yu G, Anderson SA, Ghattas I, Puelles L, Rubenstein JL. 1999. DLX-1, DLX-2, and DLX-5 expression define distinct stages of basal forebrain differentiation. J Comp Neurol 414(2):217-37.
Ellenberger T, Fass D, Arnaud M, Harrison SC. 1994. Crystal structure of transcription factor E47: E-box recognition by a basic region helix-loop-helix dimer. Genes Dev 8(8):970-80.
Erbel-Sieler C, Dudley C, Zhou Y, Wu X, Estill SJ, Han T, Diaz-Arrastia R, Brunskill EW, Potter SS, McKnight SL. 2004. Behavioral and regulatory abnormalities in mice deficient in the NPAS1 and NPAS3 transcription factors. Proc Natl Acad Sci U S A 101(37):13648-53.
Eriksson PS, Perfilieva E, Bjork-Eriksson T, Alborn AM, Nordborg C, Peterson DA, Gage FH. 1998. Neurogenesis in the adult human hippocampus. Nat Med 4(11):1313-7.
Farrell MS, Werge T, Sklar P, Owen MJ, Ophoff RA, O'Donovan MC, Corvin A, Cichon S, Sullivan PF. 2015. Evaluating historical candidate genes for schizophrenia. Mol Psychiatry 20(5):555-62.
Ferreira MA, O'Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L, Fan J, Kirov G, Perlis RH, Green EK, et al. 2008. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet 40(9):1056-8.
Ferri GL, Levi A, Possenti R. 1992. A novel neuroendocrine gene product: Selective VGF8a gene expression and immuno-localisation of the VGF protein in endocrine and neuronal populations. Brain Res Mol Brain Res 13(1-2):139-43.
Ferri GL, Noli B, Brancia C, D'Amato F, Cocco C. 2011. VGF: An inducible gene product, precursor of a diverse array of neuro-endocrine peptides and tissue-specific disease biomarkers. J Chem Neuroanat 42(4):249-61.

Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztanyi Z, El-Gebali S, Fraser M, et al. 2017. InterPro in 2017-beyond protein family and domain annotations. Nucleic Acids Res 45(D1):D190-9.
Fonseca DJ, Prada CF, Siza LM, Angel D, Gomez YM, Restrepo CM, Douben H, Rivadeneira F, de Klein A, Laissue P. 2012. A de novo 14q12q13.3 interstitial deletion in a patient affected by a severe neurodevelopmental disorder of unknown origin. Am J Med Genet A 158A(3):689-93.
Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, Georgieva L, Rees E, Palta P, Ruderfer DM, et al. 2014. De novo mutations in schizophrenia implicate synaptic networks. Nature 506(7487):179-84.
Fukunaga BN, Probst MR, Reisz-Porszasz S, Hankinson O. 1995. Identification of functional domains of the aryl hydrocarbon receptor. J Biol Chem 270(49):29270-8.
Galbraith MD, Allen MA, Bensard CL, Wang X, Schwinn MK, Qin B, Long HW, Daniels DL, Hahn WC, Dowell RD, et al. 2013. HIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. Cell 153(6):1327-39.
Gao P, He FF, Tang H, Lei CT, Chen S, Meng XF, Su H, Zhang C. 2015. NADPH oxidase-induced NALP3 inflammasome activation is driven by thioredoxin-interacting protein which contributes to podocyte injury in hyperglycemia. J Diabetes Res 2015:504761.
Geschwind DH. 2011. Genetics of autism spectrum disorders. Trends Cogn Sci 15(9):409-16.
Geschwind DH and Flint J. 2015. Genetics and genomics of psychiatric disease. Science 349(6255):1489-94.
Giegling I, Hosak L, Mossner R, Serretti A, Bellivier F, Claes S, Collier DA, Corrales A, DeLisi LE, Gallo C, et al. 2017. Genetics of schizophrenia: A consensus paper of the WFSBP task force on genetics. World J Biol Psychiatry :1-14.
Glahn DC, Ragland JD, Abramoff A, Barrett J, Laird AR, Bearden CE, Velligan DI. 2005. Beyond hypofrontality: A quantitative meta-analysis of functional neuroimaging studies of working memory in schizophrenia. Hum Brain Mapp 25(1):60-9.
Glajch KE, Kelver DA, Hegeman DJ, Cui Q, Xenias HS, Augustine EC, Hernandez VM, Verma N, Huang TY, Luo M, et al. 2016. Npas1+ pallidal neurons target striatal projection neurons. J Neurosci 36(20):5472-88.
Goncalves JT, Schafer ST, Gage FH. 2016. Adult neurogenesis in the hippocampus: From stem cells to behavior. Cell 167(4):897-914.
Gonzalez-Penas J, Arrojo M, Paz E, Brenlla J, Paramo M, Costas J. 2015. Cumulative role of rare and common putative functional genetic variants at NPAS3 in schizophrenia susceptibility. Am J Med Genet B Neuropsychiatr Genet 168(7):528-35.
Gorski JA, Talley T, Qiu M, Puelles L, Rubenstein JL, Jones KR. 2002. Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage. J Neurosci 22(15):6309-14.
Gottesman II and Gould TD. 2003. The endophenotype concept in psychiatry: Etymology and strategic intentions. Am J Psychiatry 160(4):636-45.
Gould E, Beylin A, Tanapat P, Reeves A, Shors TJ. 1999. Learning enhances adult neurogenesis in the hippocampal formation. Nat Neurosci 2(3):260-5.

Gould P and Kamnasaran D. 2011. Immunohistochemical analyses of NPAS3 expression in the developing human fetal brain. Anat Histol Embryol .
Gray VE, Hause RJ, Fowler DM. 2017. Analysis of large-scale mutagenesis data to assess the impact of single amino acid substitutions. Genetics .
Greenwood TA, Lazzeroni LC, Calkins ME, Freedman R, Green MF, Gur RE, Gur RC, Light GA, Nuechterlein KH, Olincy A, et al. 2016. Genetic assessment of additional endophenotypes from the consortium on the genetics of schizophrenia family study. Schizophr Res 170(1):30-40.
Greenwood TA, Braff DL, Light GA, Cadenhead KS, Calkins ME, Dobie DJ, Freedman R, Green MF, Gur RE, Gur RC, et al. 2007. Initial heritability analyses of endophenotypic measures for schizophrenia: The consortium on the genetics of schizophrenia. Arch Gen Psychiatry 64(11):1242-50.
Greenwood TA, Swerdlow NR, Gur RE, Cadenhead KS, Calkins ME, Dobie DJ, Freedman R, Green MF, Gur RC, Lazzeroni LC, et al. 2013. Genome-wide linkage analyses of 12 endophenotypes for schizophrenia from the consortium on the genetics of schizophrenia. Am J Psychiatry 170(5):521-32.
Gulzar ZG, McKenney JK, Brooks JD. 2013. Increased expression of NuSAP in recurrent prostate cancer is mediated by E2F1. Oncogene 32(1):70-7.
Gupta CN, Chen J, Liu J, Damaraju E, Wright C, Perrone-Bizzozero NI, Pearlson G, Luo L, Michael AM, Turner JA, et al. 2015. Genetic markers of white matter integrity in schizophrenia revealed by parallel ICA. Front Hum Neurosci 9:100.
Gur RE, Calkins ME, Gur RC, Horan WP, Nuechterlein KH, Seidman LJ, Stone WS. 2007. The consortium on the genetics of schizophrenia: Neurocognitive endophenotypes. Schizophr Bull 33(1):49-68.
Haffner C, Dettmer U, Weiler T, Haass C. 2007. The nicastrin-like protein nicalin regulates assembly and stability of the nicalin-nodal modulator (NOMO) membrane protein complex. J Biol Chem 282(14):10632-8.
Hahm S, Mizuno TM, Wu TJ, Wisor JP, Priest CA, Kozak CA, Boozer CN, Peng B, McEvoy RC, Good P, et al. 1999. Targeted deletion of the vgf gene indicates that the encoded secretory peptide precursor plays a novel role in the regulation of energy balance. Neuron 23(3):537-48.
Hannedouche S, Beck V, Leighton-Davies J, Beibel M, Roma G, Oakeley EJ, Lannoy V, Bernard J, Hamon J, Barbieri S, et al. 2013. Identification of the C3a receptor (C3AR1) as the target of the VGF-derived peptide TLQP-21 in rodent cells. J Biol Chem 288(38):27434-43.
Hansen DV, Lui JH, Parker PR, Kriegstein AR. 2010. Neurogenic radial glia in the outer subventricular zone of human neocortex. Nature 464(7288):554-61.
Hansen DV, Lui JH, Flandin P, Yoshikawa K, Rubenstein JL, Alvarez-Buylla A, Kriegstein AR. 2013. Non-epithelial stem cells and cortical interneuron production in the human ganglionic eminences. Nat Neurosci 16(11):1576-87.
Hao N, Bhakti VL, Peet DJ, Whitelaw ML. 2013. Reciprocal regulation of the basic helix-loop-helix/per-arnt-sim partner proteins, arnt and Arnt2, during neuronal differentiation. Nucleic Acids Res .
Harvey CG, Menon SD, Stachowiak B, Noor A, Proctor A, Mensah AK, Mnatzakanian GN, Alfred SE, Guo R, Scherer SW, et al. 2007. Sequence variants within exon 1 of

MECP2 occur in females with mental retardation. Am J Med Genet B Neuropsychiatr Genet 144B(3):355-60.
Hawley RJ, Scheibe RJ, Wagner JA. 1992. NGF induces the expression of the VGF gene through a cAMP response element. J Neurosci 12(7):2573-81.
Hernandez-Jimenez M, Martinez-Lopez D, Gabande-Rodriguez E, Martin-Segura A, Lizasoain I, Ledesma MD, Dotti CG, Moro MA. 2016. Seladin-1/DHCR24 is neuroprotective by associating EAAT2 glutamate transporter to lipid rafts in experimental stroke. Stroke 47(1):206-13.
Hirose K, Morita M, Ema M, Mimura J, Hamada H, Fujii H, Saijo Y, Gotoh O, Sogawa K, Fujii-Kuriyama Y. 1996. cDNA cloning and tissue-specific expression of a novel basic helix-loop-helix/PAS factor (Arnt2) with close sequence similarity to the aryl hydrocarbon receptor nuclear translocator (arnt). Mol Cell Biol 16(4):1706-13.
Hoffman EC, Reyes H, Chu FF, Sander F, Conley LH, Brooks BA, Hankinson O. 1991. Cloning of a factor required for activity of the ah (dioxin) receptor. Science 252(5008):954-8.
Horan WP, Braff DL, Nuechterlein KH, Sugar CA, Cadenhead KS, Calkins ME, Dobie DJ, Freedman R, Greenwood TA, Gur RE, et al. 2008. Verbal working memory impairments in individuals with schizophrenia and their first-degree relatives: Findings from the consortium on the genetics of schizophrenia. Schizophr Res 103(1-3):218-28.
Hord NG and Perdew GH. 1994. Physicochemical and immunocytochemical analysis of the aryl hydrocarbon receptor nuclear translocator: Characterization of two monoclonal antibodies to the aryl hydrocarbon receptor nuclear translocator. Mol Pharmacol 46(4):618-26.
Huang J, Perlis RH, Lee PH, Rush AJ, Fava M, Sachs GS, Lieberman J, Hamilton SP, Sullivan P, Sklar P, et al. 2010. Cross-disorder genomewide analysis of schizophrenia, bipolar disorder, and depression. Am J Psychiatry 167(10):1254-63.
Huang N, Chelliah Y, Shan Y, Taylor CA, Yoo SH, Partch C, Green CB, Zhang H, Takahashi JS. 2012. Crystal structure of the heterodimeric CLOCK:BMAL1 transcriptional activator complex. Science 337(6091):189-94.
Hui ST, Andres AM, Miller AK, Spann NJ, Potter DW, Post NM, Chen AZ, Sachithanantham S, Jung DY, Kim JK, et al. 2008. Txnip balances metabolic and growth signaling via PTEN disulfide reduction. Proc Natl Acad Sci U S A 105(10):3921-6.
Hui TY, Sheth SS, Diffley JM, Potter DW, Lusis AJ, Attie AD, Davis RA. 2004. Mice lacking thioredoxin-interacting protein provide evidence linking cellular redox state to appropriate response to nutritional signals. J Biol Chem 279(23):24387-93.
Ikuta T, Kobayashi Y, Kawajiri K. 2004. Cell density regulates intracellular localization of aryl hydrocarbon receptor. J Biol Chem 279(18):19209-16.
Fisher's Exact Test (3x2) Tool [Internet]; c2016 [cited 2017 08/09]. Available from: http://in-silico.net/tools/statistics/fisher_exact_test/3x2 .
Isaac DD and Andrew DJ. 1996. Tubulogenesis in drosophila: A requirement for the trachealess gene product. Genes Dev 10(1):103-17.
Ishihara S, Yasuda M, Ishizu A, Ishikawa M, Shirato H, Haga H. 2015. Activating transcription factor 5 enhances radioresistance and malignancy in cancer cells. Oncotarget 6(7):4602-14.

Jain S, Dolwick KM, Schmidt JV, Bradfield CA. 1994. Potent transactivation domains of the ah receptor and the ah receptor nuclear translocator map to their carboxyl termini. J Biol Chem 269(50):31518-24.
Jain S, Maltepe E, Lu MM, Simon C, Bradfield CA. 1998. Expression of ARNT, ARNT2, HIF1 alpha, HIF2 alpha and ah receptor mRNAs in the developing mouse. Mech Dev 73(1):117-23.
Jang MS, Park JE, Lee JA, Park SG, Myung PK, Lee DH, Park BC, Cho S. 2005. Binding and regulation of hypoxia-inducible factor-1 by the inhibitory PAS proteins. Biochem Biophys Res Commun 337(1):209-15.
Jiang BH, Rue E, Wang GL, Roe R, Semenza GL. 1996. Dimerization, DNA binding, and transactivation properties of hypoxia-inducible factor 1. J Biol Chem 271(30):17771-8.
Jiang H, Chen S, Lu N, Yue Y, Yin Y, Zhang Y, Jiang W, Liang J, Yuan Y. 2016. Reduced serum VGF levels were reversed by antidepressant treatment in depressed patients. World J Biol Psychiatry :1-6.
Jin C and Flavell RA. 2010. Inflammasome activation. the missing link: How the inflammasome senses oxidative stress. Immunol Cell Biol 88(5):510-2.
Junn E, Han SH, Im JY, Yang Y, Cho EW, Um HD, Kim DK, Lee KW, Han PL, Rhee SG, et al. 2000. Vitamin D3 up-regulated protein 1 mediates oxidative stress via suppressing the thioredoxin function. J Immunol 164(12):6287-95.
Kaadige MR, Yang J, Wilde BR, Ayer DE. 2015. MondoA-mlx transcriptional activity is limited by mTOR-MondoA interaction. Mol Cell Biol 35(1):101-10.
Kageyama R, Ohtsuka T, Hatakeyama J, Ohsawa R. 2005. Roles of bHLH genes in neural stem cell differentiation. Exp Cell Res 306(2):343-8.
Kamm GB, Pisciottano F, Kliger R, Franchini LF. 2013a. The developmental brain gene NPAS3 contains the largest number of accelerated regulatory sequences in the human genome. Mol Biol Evol 30(5):1088-102.
Kamm GB, Lopez-Leal R, Lorenzo JR, Franchini LF. 2013b. A fast-evolving human NPAS3 enhancer gained reporter expression in the developing forebrain of transgenic mice. Philos Trans R Soc Lond B Biol Sci 368(1632):20130019.
Kamnasaran D, Muir WJ, Ferguson-Smith MA, Cox DW. 2003. Disruption of the neuronal PAS3 gene in a family affected with schizophrenia. J Med Genet 40(5):325-32.
Kamnasaran D, Chen CP, Devriendt K, Mehta L, Cox DW. 2005. Defining a holoprosencephaly locus on human chromosome 14q13 and characterization of potential candidate genes. Genomics 85(5):608-21.
Kanari Y, Sato Y, Aoyama S, Muta T. 2013. Thioredoxin-interacting protein gene expression via MondoA is rapidly and transiently suppressed during inflammatory responses. PLoS One 8(3):e59026.
Kanatani S, Yozu M, Tabata H, Nakajima K. 2008. COUP-TFII is preferentially expressed in the caudal ganglionic eminence and is involved in the caudal migratory stream. J Neurosci 28(50):13582-91.
Kanno Y, Miyama Y, Takane Y, Nakahama T, Inouye Y. 2007. Identification of intracellular localization signals and of mechanisms underlining the nucleocytoplasmic shuttling of human aryl hydrocarbon receptor repressor. Biochem Biophys Res Commun 364(4):1026-31.

Karlsgodt KH, Bachman P, Winkler AM, Bearden CE, Glahn DC. 2011. Genetic influence on the working memory circuitry: Behavior, structure, function and extensions to illness. Behav Brain Res 225(2):610-22.
Kato T. 2015. Whole genome/exome sequencing in mood and psychotic disorders. Psychiatry Clin Neurosci 69(2):65-76.
Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The human genome browser at UCSC. Genome Res 12(6):996-1006.
Keshavan MS, Morris DW, Sweeney JA, Pearlson G, Thaker G, Seidman LJ, Eack SM, Tamminga C. 2011. A dimensional approach to the psychosis spectrum between bipolar disorder and schizophrenia: The schizo-bipolar scale. Schizophr Res 133(1-3):250-4.
Kewley RJ, Whitelaw ML, Chapman-Smith A. 2004. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. Int J Biochem Cell Biol 36(2):189-204.
Kibbe C, Chen J, Xu G, Jing G, Shalev A. 2013. FOXO1 competes with carbohydrate response element-binding protein (ChREBP) and inhibits thioredoxin-interacting protein (TXNIP) transcription in pancreatic beta cells. J Biol Chem 288(32):23194-202.
Kikuchi Y, Ohsawa S, Mimura J, Ema M, Takasaki C, Sogawa K, Fujii-Kuriyama Y. 2003. Heterodimers of bHLH-PAS protein fragments derived from AhR, AhRR, and arnt prepared by co-expression in escherichia coli: Characterization of their DNA binding activity and preparation of a DNA complex. J Biochem 134(1):83-90.
Kim GS, Jung JE, Narasimhan P, Sakata H, Chan PH. 2012. Induction of thioredoxin-interacting protein is mediated by oxidative stress, calcium, and glucose after brain injury in mice. Neurobiol Dis 46(2):440-9.
Kim JM, Kim DH, Lee Y, Park SJ, Ryu JH. 2014. Distinct roles of the hippocampus and perirhinal cortex in GABAA receptor blockade-induced enhancement of object recognition memory. Brain Res 1552:17-25.
Kinoshita K, Kikuchi Y, Sasakura Y, Suzuki M, Fujii-Kuriyama Y, Sogawa K. 2004. Altered DNA binding specificity of arnt by selection of partner bHLH-PAS proteins. Nucleic Acids Res 32(10):3169-79.
Klambt C, Glazer L, Shilo BZ. 1992. Breathless, a drosophila FGF receptor homolog, is essential for migration of tracheal and specific midline glial cells. Genes Dev 6(9):1668-78.
Klingberg T, Forssberg H, Westerberg H. 2002. Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. J Cogn Neurosci 14(1):1-10.
Knoth R, Singec I, Ditter M, Pantazis G, Capetian P, Meyer RP, Horvat V, Volk B, Kempermann G. 2010. Murine features of neurogenesis in the human hippocampus across the lifespan from 0 to 100 years. PLoS One 5(1):e8809.
Kodama M, Fujioka T, Duman RS. 2004. Chronic olanzapine or fluoxetine administration increases cell proliferation in hippocampus and prefrontal cortex of adult rat. Biol Psychiatry 56(8):570-80.
Kon E, Cossard A, Jossin Y. 2017. Neuronal polarity in the embryonic mammalian cerebral cortex. Front Cell Neurosci 11:163.

Kornack DR and Rakic P. 1995. Radial and horizontal deployment of clonally related cells in the primate neocortex: Relationship to distinct mitotic lineages. Neuron 15(2):311-21.
Kreft L, Soete A, Hulpiau P, Botzki A, Saeys Y, De Bleser P. 2017. ConTra v3: A tool to identify transcription factor binding sites across species, update 2017. Nucleic Acids Res .
Krishnan NM, Dhas K, Nair J, Palve V, Bagwan J, Siddappa G, Suresh A, Kekatpure VD, Kuriakose MA, Panda B. 2016. A minimal DNA methylation signature in oral tongue squamous cell carcinoma links altered methylation with tumor attributes. Mol Cancer Res 14(9):805-19.
Kristiansen M, Menghi F, Hughes R, Hubank M, Ham J. 2011. Global analysis of gene expression in NGF-deprived sympathetic neurons identifies molecular pathways associated with cell death. BMC Genomics 12:551,2164-12-551.
Kuo YM, Jones N, Zhou B, Panzer S, Larson V, Beckendorf SK. 1996. Salivary duct determination in drosophila: Roles of the EGF receptor signalling pathway and the transcription factors fork head and trachealess. Development 122(6):1909-17.
la Cour T, Kiemer L, Molgaard A, Gupta R, Skriver K, Brunak S. 2004. Analysis and prediction of leucine-rich nuclear export signals. Protein Eng Des Sel 17(6):527-36.
Lando D, Peet DJ, Whelan DA, Gorman JJ, Whitelaw ML. 2002a. Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch. Science 295(5556):858-61.
Lando D, Peet DJ, Gorman JJ, Whelan DA, Whitelaw ML, Bruick RK. 2002b. FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. Genes Dev 16(12):1466-71.
Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and clustal X version 2.0. Bioinformatics 23(21):2947-8.
Lavedan C, Licamele L, Volpi S, Hamilton J, Heaton C, Mack K, Lannan R, Thompson A, Wolfgang CD, Polymeropoulos MH. 2009. Association of the NPAS3 gene and five other loci with response to the antipsychotic iloperidone identified in a whole genome association study. Mol Psychiatry 14(8):804-19.
Lee J, Lee Y, Lee MJ, Park E, Kang SH, Chung CH, Lee KH, Kim K. 2008. Dual modification of BMAL1 by SUMO2/3 and ubiquitin promotes circadian activation of the CLOCK/BMAL1 complex. Mol Cell Biol 28(19):6056-65.
Lee T, Hacohen N, Krasnow M, Montell DJ. 1996. Regulated breathless receptor tyrosine kinase activity required to pattern cell migration and branching in the drosophila tracheal system. Genes Dev 10(22):2912-21.
Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536(7616):285-91.
Lencz T, Knowles E, Davies G, Guha S, Liewald DC, Starr JM, Djurovic S, Melle I, Sundet K, Christoforou A, et al. 2014. Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: A report from the cognitive genomics consorTium (COGENT). Mol Psychiatry 19(2):168-74.

Leuner B, Mendolia-Loffredo S, Kozorovitskiy Y, Samburg D, Gould E, Shors TJ. 2004. Learning enhances the survival of new neurons beyond the time when the hippocampus is required for memory. J Neurosci 24(34):7477-81.
Li C, Li M, Yu H, Shen X, Wang J, Sun X, Wang Q, Wang C. 2017. Neuropeptide VGF C-terminal peptide TLQP-62 alleviates lipopolysaccharide-induced memory deficits and anxiety-like and depression-like behaviors in mice: The role of BDNF/TrkB signaling. ACS Chem Neurosci .
Li H, Dong L, Whitlock JP,Jr. 1994. Transcriptional activation function of the mouse ah receptor nuclear translocator. J Biol Chem 269(45):28098-105.
Li J, Mahajan A, Tsai MD. 2006. Ankyrin repeat: A unique motif mediating protein-protein interactions. Biochemistry 45(51):15168-78.
Li S, Sheng J, Hu JK, Yu W, Kishikawa H, Hu MG, Shima K, Wu D, Xu Z, Xin W, et al. 2013. Ribonuclease 4 protects neuron degeneration by promoting angiogenesis, neurogenesis, and neuronal survival under stress. Angiogenesis 16(2):387-404.
Li X, Rong Y, Zhang M, Wang XL, LeMaire SA, Coselli JS, Zhang Y, Shen YH. 2009. Up-regulation of thioredoxin interacting protein (txnip) by p38 MAPK and FOXO1 contributes to the impaired thioredoxin activity and increased ROS in glucose-treated endothelial cells. Biochem Biophys Res Commun 381(4):660-5.
Li Y, Li J, Li S, Li Y, Wang X, Liu B, Fu Q, Ma S. 2015. Curcumin attenuates glutamate neurotoxicity in the hippocampus by suppression of ER stress-associated TXNIP/NLRP3 inflammasome activation in a manner dependent on AMPK. Toxicol Appl Pharmacol 286(1):53-63.
Liang SG and Greenwood TA. 2015. The impact of clinical heterogeneity in schizophrenia on genomic analyses. Schizophr Res 161(2-3):490-5.
Lin WJ, Jiang C, Sadahiro M, Bozdagi O, Vulchanova L, Alberini CM, Salton SR. 2015. VGF and its C-terminal peptide TLQP-62 regulate memory formation in hippocampus via a BDNF-TrkB-dependent mechanism. J Neurosci 35(28):10343-56.
Lin YC, Boone M, Meuris L, Lemmens I, Van Roy N, Soete A, Reumers J, Moisse M, Plaisance S, Drmanac R, et al. 2014. Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. Nat Commun 5:4767.
Lindebro MC, Poellinger L, Whitelaw ML. 1995. Protein-protein interaction via PAS domains: Role of the PAS domain in positive and negative regulation of the bHLH/PAS dioxin receptor-arnt transcription factor complex. Embo J 14(14):3528-39.
Liu Q, Liu L, Zhao Y, Zhang J, Wang D, Chen J, He Y, Wu J, Zhang Z, Liu Z. 2011. Hypoxia induces genomic DNA demethylation through the activation of HIF-1alpha and transcriptional upregulation of MAT2A in hepatoma cells. Mol Cancer Ther 10(6):1113-23.
Long JE, Cobos I, Potter GB, Rubenstein JL. 2009. Dlx1&2 and Mash1 transcription factors control MGE and CGE patterning and differentiation through parallel and overlapping pathways. Cereb Cortex 19 Suppl 1:i96-106.
Lu Y, Wang C, Xue Z, Li C, Zhang J, Zhao X, Liu A, Wang Q, Zhou W. 2014. PI3K/AKT/mTOR signaling-mediated neuropeptide VGF in the hippocampus of mice is involved in the rapid onset antidepressant-like effects of GLYX-13. Int J Neuropsychopharmacol 18(5):10.1093/ijnp/pyu110.

Ma PC, Rould MA, Weintraub H, Pabo CO. 1994. Crystal structure of MyoD bHLH domain-DNA complex: Perspectives on DNA recognition and implications for transcriptional activation. Cell 77(3):451-9.
Macintyre G, Alford T, Xiong L, Rouleau GA, Tibbo PG, Cox DW. 2010. Association of NPAS3 exonic variation with schizophrenia. Schizophr Res 120(1-3):143-9.
Magnusson JP, Goritz C, Tatarishvili J, Dias DO, Smith EM, Lindvall O, Kokaia Z, Frisen J. 2014. A latent neurogenic program in astrocytes regulated by notch signaling in the mouse. Science 346(6206):237-41.
Malhotra AK, Kestler LJ, Mazzanti C, Bates JA, Goldberg T, Goldman D. 2002. A functional polymorphism in the COMT gene and performance on a test of prefrontal cognition. Am J Psychiatry 159(4):652-4.
Mandolesi G, Gargano S, Pennuto M, Illi B, Molfetta R, Soucek L, Mosca L, Levi A, Jucker R, Nasi S. 2002. NGF-dependent and tissue-specific transcription of vgf is regulated by a CREB-p300 and bHLH factor interaction. FEBS Lett 510(1-2):50-6.
Marin O. 2012. Interneuron dysfunction in psychiatric disorders. Nat Rev Neurosci 13(2):107-20.
Marin O, Anderson SA, Rubenstein JL. 2000. Origin and molecular specification of striatal interneurons. J Neurosci 20(16):6063-76.
Maroof AM, Keros S, Tyson JA, Ying SW, Ganat YM, Merkle FT, Liu B, Goulburn A, Stanley EG, Elefanty AG, et al. 2013. Directed differentiation and functional maturation of cortical interneurons from human embryonic stem cells. Cell Stem Cell 12(5):559-72.
Marshall CR, Howrigan DP, Merico D, Thiruvahindrapuram B, Wu W, Greer DS, Antaki D, Shetty A, Holmans PA, Pinto D, et al. 2017. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. Nat Genet 49(1):27-35.
Marshall SA, Ng L, Unemori EN, Girling JE, Parry LJ. 2016. Relaxin deficiency results in increased expression of angiogenesis- and remodelling-related genes in the uterus of early pregnant mice but does not affect endometrial angiogenesis prior to implantation. Reprod Biol Endocrinol 14:11,016-0148-y.
Maxwell PH, Wiesener MS, Chang GW, Clifford SC, Vaux EC, Cockman ME, Wykoff CC, Pugh CW, Maher ER, Ratcliffe PJ. 1999. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature 399(6733):271-5.
McClintick JN, Xuei X, Tischfield JA, Goate A, Foroud T, Wetherill L, Ehringer MA, Edenberg HJ. 2013. Stress-response pathways are altered in the hippocampus of chronic alcoholics. Alcohol 47(7):505-15.
McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron C, Shumilina S, Lasken RS, Vermeesch JR, Hall IM, et al. 2013. Mosaic copy number variation in human neurons. Science 342(6158):632-7.
McGuire J, Okamoto K, Whitelaw ML, Tanaka H, Poellinger L. 2001. Definition of a dioxin receptor mutant that is a constitutive activator of transcription: Delineation of overlapping repression and ligand binding functions within the PAS domain. J Biol Chem 276(45):41841-9.

McGuire J, Whitelaw ML, Pongratz I, Gustafsson JA, Poellinger L. 1994. A cellular factor stimulates ligand-dependent release of hsp90 from the basic helix-loop-helix dioxin receptor. Mol Cell Biol 14(4):2438-46.
McIntosh BE, Hogenesch JB, Bradfield CA. 2010. Mammalian per-arnt-sim proteins in environmental adaptation. Annu Rev Physiol 72:625-45.
McNamara P, Seo SB, Rudic RD, Sehgal A, Chakravarti D, FitzGerald GA. 2001. Regulation of CLOCK and MOP4 by nuclear hormone receptors in the vasculature: A humoral mechanism to reset a peripheral clock. Cell 105(7):877-89.
Michaelson JJ, Shin MK, Koh JY, Brueggeman L, Zhang A, Katzman A, McDaniel L, Fang M, Pufall M, Pieper AA. 2017. Neuronal PAS domain proteins 1 and 3 are master regulators of neuropsychiatric risk genes. Biol Psychiatry .
Millipore. 2012. ENStem-A(tm) human neural progenitor expansion kit product manual. Temecula, CA, USA: Millipore.
Minn AH, Hafele C, Shalev A. 2005. Thioredoxin-interacting protein is stimulated by glucose through a carbohydrate response element and induces beta-cell apoptosis. Endocrinology 146(5):2397-405.
Miyoshi G, Butt SJ, Takebayashi H, Fishell G. 2007. Physiologically distinct temporal cohorts of cortical interneurons arise from telencephalic Olig2-expressing precursors. J Neurosci 27(29):7786-98.
Miyoshi G, Hjerling-Leffler J, Karayannis T, Sousa VH, Butt SJ, Battiste J, Johnson JE, Machold RP, Fishell G. 2010. Genetic fate mapping reveals that the caudal ganglionic eminence produces a large and diverse population of superficial cortical interneurons. J Neurosci 30(5):1582-94.
Moen MJ, Adams HH, Brandsma JH, Dekkers DH, Akinci U, Karkampouna S, Quevedo M, Kockx CE, Ozgur Z, van IJcken WF, et al. 2017. An interaction network of mental disorder proteins in neural stem cells. Transl Psychiatry 7(4):e1082.
Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepfer AM, Hinkle G, Piqani B, Eisenhaure TM, Luo B, Grenier JK, et al. 2006. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. Cell 124(6):1283-98.
Moffett P and Pelletier J. 2000. Different transcriptional properties of mSim-1 and mSim-2. FEBS Lett 466(1):80-6.
Moffett P, Reece M, Pelletier J. 1997. The murine sim-2 gene product inhibits transcription by active repression and functional interference. Mol Cell Biol 17(9):4933-47.
Mojsin M, Kovacevic-Grujicic N, Krstic A, Popovic J, Milivojevic M, Stevanovic M. 2010. Comparative analysis of SOX3 protein orthologs: Expansion of homopolymeric amino acid tracts during vertebrate evolution. Biochem Genet 48(7-8):612-23.
Moreira F, Kiehl TR, So K, Ajeawung NF, Honculada C, Gould P, Pieper RO, Kamnasaran D. 2011. NPAS3 demonstrates features of a tumor suppressive role in driving the progression of astrocytomas. Am J Pathol 179(1):462-76.
Morozova T, Hackett J, Sedaghat Y, Sonnenfeld M. 2010. The drosophila jing gene is a downstream target in the trachealess/tango tracheal pathway. Dev Genes Evol 220(7-8):191-206.

Mueller WA, Hassel M, Grealy M. 2015. Development and reproduction in humans and animal model species. Berlin: Springer.
Nakazawa K, Zsiros V, Jiang Z, Nakao K, Kolata S, Zhang S, Belforte JE. 2012. GABAergic interneuron origin of schizophrenia pathophysiology. Neuropharmacology 62(3):1574-83.
Database of Single Nucleotide Polymorphisms (dbSNP Build ID: 150) [Internet]; c2017 [cited 2017 06/22]. Available from: http://www.ncbi.nlm.nih.gov/SNP/ .
Newton SS, Collier EF, Hunsberger J, Adams D, Terwilliger R, Selvanayagam E, Duman RS. 2003. Gene profile of electroconvulsive seizures: Induction of neurotrophic and angiogenic factors. J Neurosci 23(34):10841-51.
Nishiyama A, Matsui M, Iwata S, Hirota K, Masutani H, Nakamura H, Takagi Y, Sono H, Gon Y, Yodoi J. 1999. Identification of thioredoxin-binding protein-2/vitamin D(3) up-regulated protein 1 as a negative regulator of thioredoxin function and expression. J Biol Chem 274(31):21645-50.
Nobrega-Pereira S, Gelman D, Bartolini G, Pla R, Pierani A, Marin O. 2010. Origin and molecular specification of globus pallidus neurons. J Neurosci 30(8):2824-34.
Noctor SC, Flint AC, Weissman TA, Dammerman RS, Kriegstein AR. 2001. Neurons derived from radial glial cells establish radial units in neocortex. Nature 409(6821):714-20.
Nucifora LG, Wu YC, Lee BJ, Sha L, Margolis RL, Ross CA, Sawa A, Nucifora FC,Jr. 2016. A mutation in NPAS3 that segregates with schizophrenia in a small family leads to protein aggregation. Mol Neuropsychiatry 2(3):133-44.
Nurnberger JI,Jr, Koller DL, Jung J, Edenberg HJ, Foroud T, Guella I, Vawter MP, Kelsoe JR, Psychiatric Genomics Consortium Bipolar Group. 2014. Identification of pathways for bipolar disorder: A meta-analysis. JAMA Psychiatry 71(6):657-64.
Ohsawa S, Hamada S, Kakinuma Y, Yagi T, Miura M. 2005. Novel function of neuronal PAS domain protein 1 in erythropoietin expression in neuronal cells. J Neurosci Res 79(4):451-8.
Ohshiro T and Saigo K. 1997. Transcriptional regulation of breathless FGF receptor gene by binding of TRACHEALESS/dARNT heterodimers to three central midline elements in drosophila developing trachea. Development 124(20):3975-86.
Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, Turley P, Chen GB, Emilsson V, Meddens SF, et al. 2016. Genome-wide association study identifies 74 loci associated with educational attainment. Nature 533(7604):539-42.
O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al. 2016. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44(D1):D733-45.
Oslowski CM, Hara T, O'Sullivan-Murphy B, Kanekura K, Lu S, Hara M, Ishigaki S, Zhu LJ, Hayashi E, Hui ST, et al. 2012. Thioredoxin-interacting protein mediates ER stress-induced beta cell death through initiation of the inflammasome. Cell Metab 16(2):265-73.
Ozolek JA, Jane EP, Esplen JE, Petrosko P, Wehn AK, Erb TM, Mucko SE, Cote LC, Sammak PJ. 2010. In vitro neural differentiation of human embryonic stem cells using a low-density mouse embryonic fibroblast feeder protocol. Methods Mol Biol 584:71-95.

Pan H, Che FY, Peng B, Steiner DF, Pintar JE, Fricker LD. 2006. The role of prohormone convertase-2 in hypothalamic neuropeptide processing: A quantitative neuropeptidomic study. J Neurochem 98(6):1763-77.
Park S and Gooding DC. 2014. Working memory impairment as an endophenotypic marker of a schizophrenia diathesis. Schizophr Res Cogn 1(3):127-36.
Patwari P, Higgins LJ, Chutkow WA, Yoshioka J, Lee RT. 2006. The interaction of thioredoxin with txnip. evidence for formation of a mixed disulfide by disulfide exchange. J Biol Chem 281(31):21884-91.
Pchelintsev NA, Adams PD, Nelson DM. 2016. Critical parameters for efficient sonication and improved chromatin immunoprecipitation of high molecular weight proteins. PLoS One 11(1):e0148023.
Pereira-Terra P, Kholdebarin R, Higgins M, Iwasiow BM, Correia-Pinto J, Keijzer R. 2015. Lower NPAS3 expression during the later stages of abnormal lung development in rat congenital diaphragmatic hernia. Pediatr Surg Int 31(7):659-63.
Pesold C, Impagnatiello F, Pisu MG, Uzunov DP, Costa E, Guidotti A, Caruncho HJ. 1998. Reelin is preferentially expressed in neurons synthesizing gamma-aminobutyric acid in cortex and hippocampus of adult rats. Proc Natl Acad Sci U S A 95(6):3221-6.
Peterson CW, Stoltzman CA, Sighinolfi MP, Han KS, Ayer DE. 2010. Glucose controls nuclear accumulation, promoter binding, and transcriptional activity of the MondoA-mlx heterodimer. Mol Cell Biol 30(12):2887-95.
Phelps R, Tsai A, Hagen A, Pinter J, Smith R, Stein MT. 2016. The curse of the dolphins: Cognitive decline and psychosis. J Dev Behav Pediatr 37(4):343-5.
Piccione M, Serra G, Consiglio V, Di Fiore A, Cavani S, Grasso M, Malacarne M, Pierluigi M, Viaggi C, Corsello G. 2012. 14q13.1-21.1 deletion encompassing the HPE8 locus in an adolescent with intellectual disability and bilateral microphthalmia, but without holoprosencephaly. Am J Med Genet A 158A(6):1427-33.
Pickard BS, Pieper AA, Porteous DJ, Blackwood DH, Muir WJ. 2006. The NPAS3 gene--emerging evidence for a role in psychiatric illness. Ann Med 38(6):439-48.
Pickard BS, Malloy MP, Porteous DJ, Blackwood DH, Muir WJ. 2005. Disruption of a brain transcription factor, NPAS3, is associated with schizophrenia and learning disability. Am J Med Genet B Neuropsychiatr Genet 136B(1):26-32.
Pickard BS, Christoforou A, Thomson PA, Fawkes A, Evans KL, Morris SW, Porteous DJ, Blackwood DH, Muir WJ. 2008. Interacting haplotypes at the NPAS3 locus alter risk of schizophrenia and bipolar disorder. Mol Psychiatry .
Pieper AA, Wu X, Han TW, Estill SJ, Dang Q, Wu LC, Reece-Fincanon S, Dudley CA, Richardson JA, Brat DJ, et al. 2005. The neuronal PAS domain protein 3 transcription factor controls FGF-mediated adult hippocampal neurogenesis in mice. Proc Natl Acad Sci U S A 102(39):14052-7.
Pieper AA, Xie S, Capota E, Estill SJ, Zhong J, Long JM, Becker GL, Huntington P, Goldman SE, Shen CH, et al. 2010. Discovery of a proneurogenic, neuroprotective chemical. Cell 142(1):39-51.
Pollard KS, Salama SR, King B, Kern AD, Dreszer T, Katzman S, Siepel A, Pedersen JS, Bejerano G, Baertsch R, et al. 2006a. Forces shaping the fastest evolving regions in the human genome. PLoS Genet 2(10):e168.

Pollard KS, Salama SR, Lambert N, Lambot MA, Coppens S, Pedersen JS, Katzman S, King B, Onodera C, Siepel A, et al. 2006b. An RNA gene expressed during cortical development evolved rapidly in humans. Nature 443(7108):167-72.
Pollenz RS. 1996. The aryl-hydrocarbon receptor, but not the aryl-hydrocarbon receptor nuclear translocator protein, is rapidly depleted in hepatic and nonhepatic culture cells exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Mol Pharmacol 49(3):391-8.
Pongratz I, Antonsson C, Whitelaw ML, Poellinger L. 1998. Role of the PAS domain in regulation of dimerization and DNA binding specificity of the dioxin receptor. Mol Cell Biol 18(7):4079-88.
Possenti R, Di Rocco G, Nasi S, Levi A. 1992. Regulatory elements in the promoter region of vgf, a nerve growth factor-inducible gene. Proc Natl Acad Sci U S A 89(9):3815-9.
Possenti R, Eldridge JD, Paterson BM, Grasso A, Levi A. 1989. A protein induced by NGF in PC12 cells is stored in secretory vesicles and released through the regulated pathway. Embo J 8(8):2217-23.
Potter GB, Petryniak MA, Shevchenko E, McKinsey GL, Ekker M, Rubenstein JL. 2009. Generation of cre-transgenic mice using Dlx1/Dlx2 enhancers and their characterization in GABAergic interneurons. Mol Cell Neurosci 40(2):167-86.
Pretto DI, Eid JS, Yrigollen CM, Tang HT, Loomis EW, Raske C, Durbin-Johnson B, Hagerman PJ, Tassone F. 2015. Differential increases of specific FMR1 mRNA isoforms in premutation carriers. J Med Genet 52(1):42-52.
Price SA, Gardiner NJ, Duran-Jimenez B, Zeef LA, Obrosova IG, Tomlinson DR. 2006. Thioredoxin interacting protein is increased in sensory neurons in experimental diabetes. Brain Res 1116(1):206-14.
Psychiatric GWAS Consortium Steering Committee. 2009. A framework for interpreting genome-wide association studies of psychiatric disorders. Mol Psychiatry 14(1):10-7.
A Report on Mental Illnesses in Canada: Chapter 3 Schizophrenia [Internet]; c2012 [cited 2017 07/05]. Available from: http://www.phac-aspc.gc.ca/publicat/miic-mmac/chap_3-eng.php .
Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P, O'Dushlaine C, Chambert K, Bergen SE, Kahler A, et al. 2014. A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506(7487):185-90.
Purdon S. 2005. The Screen for Cognitive Impairment in Psychiatry (SCIP): Manual. PNL Inc: Edmonton, Alberta [computer program]. .
R Core Team. 2013. R: A language and environment for statistical computing.[computer program]. Vienna, Austria: R Foundation for Statistical Computing. .
Ramos A, Rodriguez-Seoane C, Rosa I, Trossbach SV, Ortega-Alonso A, Tomppo L, Ekelund J, Veijola J, Jarvelin MR, Alonso J, et al. 2014. Neuropeptide precursor VGF is genetically associated with social anhedonia and underrepresented in the brain of major mental illness: Its downregulation by DISC1. Hum Mol Genet 23(22):5859-65.
Rana TM. 2007. Illuminating the silence: Understanding the structure and function of small RNAs. Nat Rev Mol Cell Biol 8(1):23-36.
Rapisarda A, Kraus M, Tan YW, Lam M, Eng GK, Lee J, Subramaniam M, Collinson SL, Chong SA, Keefe RS. 2014. The continuous performance test, identical pairs:

Norms, reliability and performance in healthy controls and patients with schizophrenia in singapore. Schizophr Res 156(2-3):233-40.
Reichman-Fried M, Dickson B, Hafen E, Shilo BZ. 1994. Elucidation of the role of breathless, a drosophila FGF receptor homolog, in tracheal cell migration. Genes Dev 8(4):428-39.
Reif A, Fritzen S, Finger M, Strobel A, Lauer M, Schmitt A, Lesch KP. 2006. Neural stem cell proliferation is decreased in schizophrenia, but not in depression. Mol Psychiatry 11(5):514-22.
Reisz-Porszasz S, Probst MR, Fukunaga BN, Hankinson O. 1994. Identification of functional domains of the aryl hydrocarbon receptor nuclear translocator protein (ARNT). Mol Cell Biol 14(9):6075-86.
Rey G, Cesbron F, Rougemont J, Reinke H, Brunner M, Naef F. 2011. Genome-wide and phase-specific DNA-binding rhythms of BMAL1 control circadian output functions in mouse liver. PLoS Biol 9(2):e1000595.
Reyes H, Reisz-Porszasz S, Hankinson O. 1992. Identification of the ah receptor nuclear translocator protein (arnt) as a component of the DNA binding form of the ah receptor. Science 256(5060):1193-5.
Rietveld CA, Esko T, Davies G, Pers TH, Turley P, Benyamin B, Chabris CF, Emilsson V, Johnson AD, Lee JJ, et al. 2014. Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. Proc Natl Acad Sci U S A 111(38):13790-4.
Rietveld CA, Medland SE, Derringer J, Yang J, Esko T, Martin NW, Westra HJ, Shakhbazov K, Abdellaoui A, Agrawal A, et al. 2013. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. Science 340(6139):1467-71.
Rosenbloom KR, Dreszer TR, Pheasant M, Barber GP, Meyer LR, Pohl A, Raney BJ, Wang T, Hinrichs AS, Zweig AS, et al. 2010. ENCODE whole-genome data in the UCSC genome browser. Nucleic Acids Res 38(Database issue):D620-5.
Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, Wong MC, Maddren M, Fang R, Heitner SG, et al. 2013. ENCODE data in the UCSC genome browser: Year 5 update. Nucleic Acids Res 41(Database issue):D56-63.
Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M, et al. 2015. The UCSC genome browser database: 2015 update. Nucleic Acids Res 43(Database issue):D670-81.
Rosenfeld JA, Ballif BC, Martin DM, Aylsworth AS, Bejjani BA, Torchia BS, Shaffer LG. 2010. Clinical characterization of individuals with deletions of genes in holoprosencephaly pathways by aCGH refines the phenotypic spectrum of HPE. Hum Genet 127(4):421-40.
Rudy B, Fishell G, Lee S, Hjerling-Leffler J. 2011. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev Neurobiol 71(1):45-61.
Rutter J, Reick M, Wu LC, McKnight SL. 2001. Regulation of clock and NPAS2 DNA binding by the redox state of NAD cofactors. Science 293(5529):510-4.
Saade E, Mechold U, Kulyyassov A, Vertut D, Lipinski M, Ogryzko V. 2009. Analysis of interaction partners of H4 histone by a new proteomics approach. Proteomics 9(21):4934-43.

Saderi N, Buijs FN, Salgado-Delgado R, Merkenstein M, Basualdo MC, Ferri GL, Escobar C, Buijs RM. 2014. A role for VGF in the hypothalamic arcuate and paraventricular nuclei in the control of energy homeostasis. Neuroscience 265:184-95.
Sahakian BJ, Bruhl AB, Cook J, Killikelly C, Savulich G, Piercy T, Hafizi S, Perez J, Fernandez-Egea E, Suckling J, et al. 2015. The impact of neuroscience on society: Cognitive enhancement in neuropsychiatric disorders and in healthy people. Philos Trans R Soc Lond B Biol Sci 370(1677):20140214.
Sakamoto M, Miyazaki Y, Kitajo K, Yamaguchi A. 2015. VGF, which is induced transcriptionally in stroke brain, enhances neurite extension and confers protection against ischemia in vitro. Transl Stroke Res 6(4):301-8.
Salceda S and Caro J. 1997. Hypoxia-inducible factor 1alpha (HIF-1alpha) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions. its stabilization by hypoxia depends on redox-induced changes. J Biol Chem 272(36):22642-7.
Salton SR. 1991. Nucleotide sequence and regulatory studies of VGF, a nervous system-specific mRNA that is rapidly and relatively selectively induced by nerve growth factor. J Neurochem 57(3):991-6.
Salton SR, Fischberg DJ, Dong KW. 1991. Structure of the gene encoding VGF, a nervous system-specific mRNA that is rapidly and selectively induced by nerve growth factor in PC12 cells. Mol Cell Biol 11(5):2335-49.
Sancak Y, Peterson TR, Shaul YD, Lindquist RA, Thoreen CC, Bar-Peled L, Sabatini DM. 2008. The rag GTPases bind raptor and mediate amino acid signaling to mTORC1. Science 320(5882):1496-501.
Saxena G, Chen J, Shalev A. 2010. Intracellular shuttling and mitochondrial function of thioredoxin-interacting protein. J Biol Chem 285(6):3997-4005.
Schizophrenia Working Group of the Psychiatric Genomics Consortium. 2014. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511(7510):421-7.
Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH image to ImageJ: 25 years of image analysis. Nat Methods 9(7):671-5.
Schulz TC, Palmarini GM, Noggle SA, Weiler DA, Mitalipova MM, Condie BG. 2003. Directed neuronal differentiation of human embryonic stem cells. BMC Neurosci 4:27.
Schulze PC, Yoshioka J, Takahashi T, He Z, King GL, Lee RT. 2004. Hyperglycemia promotes oxidative stress through inhibition of thioredoxin function by thioredoxin-interacting protein. J Biol Chem 279(29):30369-74.
Schwarz E, Tost H, Meyer-Lindenberg A. 2015. Working memory genetics in schizophrenia and related disorders: An RDoC perspective. Am J Med Genet B Neuropsychiatr Genet .
Seidman LJ, Thermenos HW, Poldrack RA, Peace NK, Koch JK, Faraone SV, Tsuang MT. 2006. Altered brain activation in dorsolateral prefrontal cortex in adolescents and young adults at genetic risk for schizophrenia: An fMRI study of working memory. Schizophr Res 85(1-3):58-72.
Seiler CY, Park JG, Sharma A, Hunter P, Surapaneni P, Sedillo C, Field J, Algar R, Price A, Steel J, et al. 2014. DNASU plasmid and PSI:Biology-materials repositories:

Resources to accelerate biological research. Nucleic Acids Res 42(Database issue):D1253-60.
Sekine H, Mimura J, Yamamoto M, Fujii-Kuriyama Y. 2006. Unique and overlapping transcriptional roles of arylhydrocarbon receptor nuclear translocator (arnt) and Arnt2 in xenobiotic and hypoxic responses. J Biol Chem 281(49):37507-16.
Sellier C, Buijsen RA, He F, Natla S, Jung L, Tropel P, Gaucherot A, Jacobs H, Meziane H, Vincent A, et al. 2017. Translation of expanded CGG repeats into FMRpolyG is pathogenic and may contribute to fragile X tremor ataxia syndrome. Neuron 93(2):331-47.
Seok SH, Lee W, Jiang L, Molugu K, Zheng A, Li Y, Park S, Bradfield CA, Xing Y. 2017. Structural hierarchy controlling dimerization and target DNA recognition in the AHR transcriptional complex. Proc Natl Acad Sci U S A 114(21):5431-6.
Severini C, Ciotti MT, Biondini L, Quaresima S, Rinaldi AM, Levi A, Frank C, Possenti R. 2008. TLQP-21, a neuroendocrine VGF-derived peptide, prevents cerebellar granule cells death induced by serum and potassium deprivation. J Neurochem 104(2):534-44.
Sha L, Macintyre L, Machell JA, Kelly MP, Porteous DJ, Brandon NJ, Muir WJ, Blackwood DH, Watson DG, Clapcote SJ, et al. 2012. Transcriptional regulation of neurodevelopmental and metabolic pathways by NPAS3. Mol Psychiatry 17(3):267-79.
Shalev A, Pise-Masison CA, Radonovich M, Hoffmann SC, Hirshberg B, Brady JN, Harlan DM. 2002. Oligonucleotide microarray analysis of intact human pancreatic islets: Identification of glucose-responsive genes and a highly regulated TGFbeta signaling pathway. Endocrinology 143(9):3695-8.
Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: The NCBI database of genetic variation. Nucleic Acids Res 29(1):308-11.
Shi YQ, Liao SY, Zhuang XJ, Han CS. 2011. Mouse Fem1b interacts with and induces ubiquitin-mediated degradation of Ankrd37. Gene 485(2):153-9.
Shin J and Kim J. 2013. Novel alternative splice variants of chicken NPAS3 are expressed in the developing central nervous system. Gene 530(2):222-8.
Shin J, Jeong HY, Lee KE, Kim J. 2010. Isolation and characterization of chicken NPAS3. Exp Neurobiol 19(2):71-4.
Shin J, Berg DA, Zhu Y, Shin JY, Song J, Bonaguidi MA, Enikolopov G, Nauen DW, Christian KM, Ming GL, et al. 2015. Single-cell RNA-seq with waterfall reveals molecular cascades underlying adult neurogenesis. Cell Stem Cell 17(3):360-72.
Snyder SE, Pintar JE, Salton SR. 1998. Developmental expression of VGF mRNA in the prenatal and postnatal rat. J Comp Neurol 394(1):64-90.
Snyder SE, Peng B, Pintar JE, Salton SR. 2003. Expression of VGF mRNA in developing neuroendocrine and endocrine tissues. J Endocrinol 179(2):227-35.
Snyder SE, Cheng HW, Murray KD, Isackson PJ, McNeill TH, Salton SR. 1998. The messenger RNA encoding VGF, a neuronal peptide precursor, is rapidly regulated in the rat central nervous system by neuronal activity, seizure and lesion. Neuroscience 82(1):7-19.
Sogawa K, Iwabuchi K, Abe H, Fujii-Kuriyama Y. 1995. Transcriptional activation domains of the ah receptor and ah receptor nuclear translocator. J Cancer Res Clin Oncol 121(9-10):612-20.

Solek CM, Feng S, Perin S, Weinschutz Mendes H, Ekker M. 2017. Lineage tracing of dlx1a/2a and dlx5a/6a expressing cells in the developing zebrafish brain. Dev Biol 427(1):131-47.
Solvsten CAE, Daugaard TF, Luo Y, de Paoli F, Christensen JH, Nielsen AL. 2017. The effects of voluntary physical exercise-activated neurotrophic signaling in rat hippocampus on mRNA levels of downstream signaling molecules. J Mol Neurosci 62(2):142-53.
Song H, Stevens CF, Gage FH. 2002. Astroglia induce neurogenesis from adult neural stem cells. Nature 417(6884):39-44.
Sonnenfeld M, Ward M, Nystrom G, Mosher J, Stahl S, Crews S. 1997. The drosophila tango gene encodes a bHLH-PAS protein that is orthologous to mammalian arnt and controls CNS midline and tracheal development. Development 124(22):4571-82.
Spalding KL, Bergmann O, Alkass K, Bernard S, Salehpour M, Huttner HB, Bostrom E, Westerlund I, Vial C, Buchholz BA, et al. 2013. Dynamics of hippocampal neurogenesis in adult humans. Cell 153(6):1219-27.
Spengler ML, Kuropatwinski KK, Schumer M, Antoch MP. 2009. A serine cluster mediates BMAL1-dependent CLOCK phosphorylation and degradation. Cell Cycle 8(24):4138-46.
Spiegel I, Mardinly AR, Gabel HW, Bazinet JE, Couch CH, Tzeng CP, Harmin DA, Greenberg ME. 2014. Npas4 regulates excitatory-inhibitory balance within neural circuits through cell-type-specific gene programs. Cell 157(5):1216-29.
Spindel ON, World C, Berk BC. 2012. Thioredoxin interacting protein: Redox dependent and independent regulatory mechanisms. Antioxid Redox Signal 16(6):587-96.
Stanco A, Pla R, Vogt D, Chen Y, Mandal S, Walker J, Hunt RF, Lindtner S, Erdman CA, Pieper AA, et al. 2014. NPAS1 represses the generation of specific subtypes of cortical interneurons. Neuron 84(5):940-53.
Stoltzman CA, Kaadige MR, Peterson CW, Ayer DE. 2011. MondoA senses non-glucose sugars: Regulation of thioredoxin-interacting protein (TXNIP) and the hexose transport curb. J Biol Chem 286(44):38027-34.
Stoltzman CA, Peterson CW, Breen KT, Muoio DM, Billin AN, Ayer DE. 2008. Glucose sensing by MondoA:Mlx complexes: A role for hexokinases and direct regulation of thioredoxin-interacting protein expression. Proc Natl Acad Sci U S A 105(19):6912-7.
Suhl JA and Warren ST. 2015. Single-nucleotide mutations in FMR1 reveal novel functions and regulatory mechanisms of the fragile X syndrome protein FMRP. J Exp Neurosci 9(Suppl 2):35-41.
Sullivan PF, Daly MJ, O'Donovan M. 2012. Genetic architectures of psychiatric disorders: The emerging picture and its implications. Nat Rev Genet 13(8):537-51.
Sullivan PF, Kendler KS, Neale MC. 2003. Schizophrenia as a complex trait: Evidence from a meta-analysis of twin studies. Arch Gen Psychiatry 60(12):1187-92.
Swanson HI, Chan WK, Bradfield CA. 1995. DNA binding specificities and pairing rules of the ah receptor, ARNT, and SIM proteins. J Biol Chem 270(44):26292-302.
Tamaru T, Hirayama J, Isojima Y, Nagai K, Norioka S, Takamatsu K, Sassone-Corsi P. 2009. CK2alpha phosphorylates BMAL1 to regulate the mammalian clock. Nat Struct Mol Biol 16(4):446-8.

Tandon R, Keshavan MS, Nasrallah HA. 2008. Schizophrenia, "just the facts": What we know in 2008 part 1: Overview. Schizophr Res 100(1-3):4-19.
Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, Arora VK, Kaushik P, Cerami E, Reva B, et al. 2010. Integrative genomic profiling of human prostate cancer. Cancer Cell 18(1):11-22.
Teh CH, Lam KK, Loh CC, Loo JM, Yan T, Lim TM. 2006. Neuronal PAS domain protein 1 is a transcriptional repressor and requires arylhydrocarbon nuclear translocator for its nuclear localization. J Biol Chem 281(45):34617-29.
Thakker-Varia S and Alder J. 2009. Neuropeptides in depression: Role of VGF. Behav Brain Res 197(2):262-78.
Thakker-Varia S, Krol JJ, Nettleton J, Bilimoria PM, Bangasser DA, Shors TJ, Black IB, Alder J. 2007. The neuropeptide VGF produces antidepressant-like behavioral effects and enhances proliferation in the hippocampus. J Neurosci 27(45):12156-67.
Thakker-Varia S, Jean YY, Parikh P, Sizer CF, Jernstedt Ayer J, Parikh A, Hyde TM, Buyske S, Alder J. 2010. The neuropeptide VGF is reduced in human bipolar postmortem brain and contributes to some of the behavioral and molecular effects of lithium. J Neurosci 30(28):9368-80.
Thakker-Varia S, Behnke J, Doobin D, Dalal V, Thakkar K, Khadim F, Wilson E, Palmieri A, Antila H, Rantamaki T, et al. 2014. VGF (TLQP-62)-induced neurogenesis targets early phase neural progenitor cells in the adult hippocampus and requires glutamate and BDNF signaling. Stem Cell Res 12(3):762-77.
Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen KC, Scaglione KM, Basrur V, et al. 2013. CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron 78(3):440-55.
Torres-Peraza JF, Engel T, Martin-Ibanez R, Sanz-Rodriguez A, Fernandez-Fernandez MR, Esgleas M, Canals JM, Henshall DC, Lucas JJ. 2013. Protective neuronal induction of ATF5 in endoplasmic reticulum stress induced by status epilepticus. Brain 136(Pt 4):1161-76.
Trampush JW, Yang ML, Yu J, Knowles E, Davies G, Liewald DC, Starr JM, Djurovic S, Melle I, Sundet K, et al. 2017. GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: A report from the COGENT consortium. Mol Psychiatry 22(3):336-45.
Trani E, Giorgi A, Canu N, Amadoro G, Rinaldi AM, Halban PA, Ferri GL, Possenti R, Schinina ME, Levi A. 2002. Isolation and characterization of VGF peptides in rat brain. role of PC1/3 and PC2 in the maturation of VGF precursor. J Neurochem 81(3):565-74.
Travan L, Naviglio S, De Cunto A, Pellegrin A, Pecile V, Spinelli AM, Cappellani S, Faletra F. 2017. Phenotypic expression of 19q13.32 microdeletions: Report of a new patient and review of the literature. Am J Med Genet A .
Trifonova EA, Gabidulina TV, Ershov NI, Serebrova VN, Vorozhishcheva AY, Stepanov VA. 2014. Analysis of the placental tissue transcriptome of normal and preeclampsia complicated pregnancies. Acta Naturae 6(2):71-83.
Uchida T, et al. 2012. Effects of the bHLH domain on axial coordination of heme in the PAS-A domain of neuronal PAS domain protein 2 (NPAS2): Conversion from His119/Cys170 coordination to His119/His171 coordination. Journal of Inorganic Biochemistry 108:188-95.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3(7):RESEARCH0034.
Vieland VJ, Walters KA, Lehner T, Azaro M, Tobin K, Huang Y, Brzustowicz LM. 2014. Revisiting schizophrenia linkage data in the NIMH repository: Reanalysis of regularized data across multiple studies. Am J Psychiatry 171(3):350-9.
Visser R, Gijsbers A, Ruivenkamp C, Karperien M, Reeser HM, Breuning MH, Kant SG, Wit JM. 2010. Genome-wide SNP array analysis in patients with features of sotos syndrome. Horm Res Paediatr 73(4):265-74.
Vreugdenhil E and Berezikov E. 2010. Fine-tuning the brain: MicroRNAs. Front Neuroendocrinol 31(2):128-33.
Walker AK, Rivera PD, Wang Q, Chuang JC, Tran S, Osborne-Lawrence S, Estill SJ, Starwalt R, Huntington P, Morlock L, et al. 2015. The P7C3 class of neuroprotective compounds exerts antidepressant efficacy in mice by increasing hippocampal neurogenesis. Mol Psychiatry 20(4):500-8.
Wang GL, Jiang BH, Rue EA, Semenza GL. 1995. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. Proc Natl Acad Sci U S A 92(12):5510-4.
Wang W, Lu Y, Xue Z, Li C, Wang C, Zhao X, Zhang J, Wei X, Chen X, Cui W, et al. 2015. Rapid-acting antidepressant-like effects of acetyl-l-carnitine mediated by PI3K/AKT/BDNF/VGF signaling pathway in mice. Neuroscience 285:281-91.
Wang Z, Rong YP, Malone MH, Davis MC, Zhong F, Distelhorst CW. 2006. Thioredoxin-interacting protein (txnip) is a glucocorticoid-regulated primary response gene involved in mediating glucocorticoid-induced apoptosis. Oncogene 25(13):1903-13.
Wardman JH, Zhang X, Gagnon S, Castro LM, Zhu X, Steiner DF, Day R, Fricker LD. 2010. Analysis of peptides in prohormone convertase 1/3 null mouse brain using quantitative peptidomics. J Neurochem 114(1):215-25.
Weber H, Kittel-Schneider S, Gessner A, Domschke K, Neuner M, Jacob CP, Buttenschon HN, Boreatti-Hummer A, Volkert J, Herterich S, et al. 2011. Cross-disorder analysis of bipolar risk genes: Further evidence of DGKH as a risk gene for bipolar disorder, but also unipolar depression and adult ADHD. Neuropsychopharmacology 36(10):2076-85.
Wei PC, Chang AN, Kao J, Du Z, Meyers RM, Alt FW, Schwer B. 2016. Long neural genes harbor recurrent DNA break clusters in neural stem/progenitor cells. Cell 164(4):644-55.
Wharton KA,Jr, Franks RG, Kasai Y, Crews ST. 1994. Control of CNS midline transcription by asymmetric E-box-like elements: Similarity to xenobiotic responsive regulation. Development 120(12):3563-9.
Whitelaw ML, Gustafsson JA, Poellinger L. 1994. Identification of transactivation and repression functions of the dioxin receptor and its basic helix-loop-helix/PAS partner factor arnt: Inducible versus constitutive modes of regulation. Mol Cell Biol 14(12):8343-55.
Wilhelmsson U, Faiz M, de Pablo Y, Sjoqvist M, Andersson D, Widestrand A, Potokar M, Stenovec M, Smith PL, Shinjyo N, et al. 2012. Astrocytes negatively regulate

neurogenesis through the Jagged1-mediated notch pathway. Stem Cells 30(10):2320-9.
Wilk R, Weizman I, Shilo BZ. 1996. Trachealess encodes a bHLH-PAS protein that is an inducer of tracheal cell fates in drosophila. Genes Dev 10(1):93-102.
Wilkinson G and Robertson G. 2006. Wide range achievement test 4 professional manual. Lutz, FL: Psychological Assessment Resources.
Winocur G, Wojtowicz JM, Sekeres M, Snyder JS, Wang S. 2006. Inhibition of neurogenesis interferes with hippocampus-dependent memory function. Hippocampus 16(3):296-304.
Wisor JP and Takahashi JS. 1997. Regulation of the vgf gene in the golden hamster suprachiasmatic nucleus by light and by the circadian clock. J Comp Neurol 378(2):229-38.
Witte AV and Floel A. 2012. Effects of COMT polymorphisms on brain function and behavior in health and disease. Brain Res Bull 88(5):418-28.
Wojtal D, Kemaladewi DU, Malam Z, Abdullah S, Wong TW, Hyatt E, Baghestani Z, Pereira S, Stavropoulos J, Mouly V, et al. 2016. Spell checking nature: Versatility of CRISPR/Cas9 for developing treatments for inherited disorders. Am J Hum Genet 98(1):90-101.
Wong J, Duncan CE, Beveridge NJ, Webster MJ, Cairns MJ, Shannon Weickert C. 2012. Expression of NPAS3 in the human cortex and evidence of its posttranscriptional regulation by miR-17 during development, with implications for schizophrenia. Schizophr Bull .
Wong RW and Hagen T. 2013. Mechanistic target of rapamycin (mTOR) dependent regulation of thioredoxin interacting protein (TXNIP) transcription in hypoxia. Biochem Biophys Res Commun 433(1):40-6.
Woo SM, Kim J, Han HW, Chae JI, Son MY, Cho S, Chung HM, Han YM, Kang YK. 2009. Notch signaling is required for maintaining stem-cell features of neuroprogenitor cells derived from human embryonic stem cells. BMC Neurosci 10:97,2202-10-97.
Wood SM, Gleadle JM, Pugh CW, Hankinson O, Ratcliffe PJ. 1996. The role of the aryl hydrocarbon receptor nuclear translocator (ARNT) in hypoxic induction of gene expression. studies in ARNT-deficient cells. J Biol Chem 271(25):15117-23.
Woods SL and Whitelaw ML. 2002. Differential activities of murine single minded 1 (SIM1) and SIM2 on a hypoxic response element. cross-talk between basic helix-loop-helix/per-arnt-sim homology transcription factors. J Biol Chem 277(12):10236-43.
Wray NR, Lee SH, Kendler KS. 2012. Impact of diagnostic misclassification on estimation of genetic correlations using genome-wide genotypes. Eur J Hum Genet 20(6):668-74.
Wu D, Potluri N, Kim Y, Rastinejad F. 2013a. Structure and dimerization properties of the aryl hydrocarbon receptor PAS-A domain. Mol Cell Biol 33(21):4346-56.
Wu D, Su X, Potluri N, Kim Y, Rastinejad F. 2016. NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors. Elife 5:10.7554/eLife.18790.
Wu D, Potluri N, Lu J, Kim Y, Rastinejad F. 2015. Structural integration in hypoxia-inducible factors. Nature 524(7565):303-8.

Wu D, Zhang R, Zhao R, Chen G, Cai Y, Jin J. 2013b. A novel function of novobiocin: Disrupting the interaction of HIF 1alpha and p300/CBP through direct binding to the HIF1alpha C-terminal activation domain. PLoS One 8(5):e62014.
Xiong W, Li J, Zhang E, Huang H. 2016. BMAL1 regulates transcription initiation and activates circadian clock gene expression in mammals. Biochem Biophys Res Commun 473(4):1019-25.
Xu Q, Cobos I, De La Cruz E, Rubenstein JL, Anderson SA. 2004. Origins of cortical interneuron subtypes. J Neurosci 24(11):2612-22.
Yang D, Zhang W, Padhiar A, Yue Y, Shi Y, Zheng T, Davis K, Zhang Y, Huang M, Li Y, et al. 2016. NPAS3 regulates transcription and expression of VGF: Implications for neurogenesis and psychiatric disorders. Front Mol Neurosci 9:109.
Yoshihara E, Masaki S, Matsuo Y, Chen Z, Tian H, Yodoi J. 2014. Thioredoxin/txnip: Redoxisome, as a redox switch for the pathogenesis of diseases. Front Immunol 4:514.
Yu FX and Luo Y. 2009. Tandem ChoRE and CCAAT motifs and associated factors regulate txnip expression in response to glucose or adenosine-containing molecules. PLoS One 4(12):e8397.
Yu FX, Chai TF, He H, Hagen T, Luo Y. 2010. Thioredoxin-interacting protein (txnip) gene expression: Sensing oxidative phosphorylation status and glycolytic rate. J Biol Chem 285(33):25822-30.
Yu L, Arbez N, Nucifora LG, Sell GL, Delisi LE, Ross CA, Margolis RL, Nucifora FC,Jr. 2014. A mutation in NPAS3 segregates with mental illness in a small family. Mol Psychiatry 19(1):7-8.
Yuan Y, Hilliard G, Ferguson T, Millhorn DE. 2003. Cobalt inhibits the interaction between hypoxia-inducible factor-alpha and von hippel-lindau protein by direct binding to hypoxia-inducible factor-alpha. J Biol Chem 278(18):15911-6.
Zai G, Robbins TW, Sahakian BJ, Kennedy JL. 2017. A review of molecular genetic studies of neurocognitive deficits in schizophrenia. Neurosci Biobehav Rev 72:50-67.
Zaragoza-Campillo MA and Moran J. 2017. Reactive oxygen species evoked by potassium deprivation and staurosporine inactivate akt and induce the expression of TXNIP in cerebellar granule neurons. Oxid Med Cell Longev 2017:8930406.
Zelzer E, Wappner P, Shilo BZ. 1997. The PAS domain confers target gene specificity of drosophila bHLH/PAS proteins. Genes Dev 11(16):2079-89.
Zhang W, Ni C, Sheng J, Hua Y, Ma J, Wang L, Zhao Y, Xing Y. 2013a. TLQP-21 protects human umbilical vein endothelial cells against high-glucose-induced apoptosis by increasing G6PD expression. PLoS One 8(11):e79760.
Zhang Y, Kent JW,Jr, Olivier M, Ali O, Cerjak D, Broeckel U, Abdou RM, Dyer TD, Comuzzie A, Curran JE, et al. 2013b. A comprehensive analysis of adiponectin QTLs using SNP association, SNP cis-effects on peripheral blood gene expression and gene expression correlation identified novel metabolic syndrome (MetS) genes with potential role in carcinogenesis and systemic inflammation. BMC Med Genomics 6:14,8794-6-14.
Zhang Z, Jones A, Joo HY, Zhou D, Cao Y, Chen S, Erdjument-Bromage H, Renfrow M, He H, Tempst P, et al. 2013c. USP49 deubiquitinates histone H2B and regulates cotranscriptional pre-mRNA splicing. Genes Dev 27(14):1581-95.

Zhao Q, Che X, Zhang H, Tan G, Liu L, Jiang D, Zhao J, Xiang X, Sun X, He Z. 2017. Thioredoxin-interacting protein mediates apoptosis in early brain injury after subarachnoid haemorrhage. Int J Mol Sci 18(4):10.3390/ijms18040854.

Zhou J and Chng WJ. 2013. Roles of thioredoxin binding protein (TXNIP) in oxidative stress, apoptosis and cancer. Mitochondrion 13(3):163-9.

Zhou R, Tardivel A, Thorens B, Choi I, Tschopp J. 2010. Thioredoxin-interacting protein links oxidative stress to inflammasome activation. Nat Immunol 11(2):136-40.

Zhou S, Degan S, Potts EN, Foster WM, Sunday ME. 2009. NPAS3 is a trachealess homolog critical for lung development and homeostasis. Proc Natl Acad Sci U S A 106(28):11691-6.

Zhou YD, Barnard M, Tian H, Li X, Ring HZ, Francke U, Shelton J, Richardson J, Russell DW, McKnight SL. 1997. Molecular characterization of two mammalian bHLH-PAS domain proteins selectively expressed in the central nervous system. Proc Natl Acad Sci U S A 94(2):713-8.